Protein Science (2000), *9*:1873–1877. Cambridge University Press. Printed in the USA.

Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange

LELAND MAYNE AND S. WALTER ENGLANDER

The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

(RECEIVED May 2, 2000; FINAL REVISION July 21, 2000; ACCEPTED July 28, 2000)

Abstract

A direct conflict between the stabilization free energy parameters of cytochrome *c* determined by optical methods and by hydrogen exchange (HX) is quantitatively explained when the partially folded intermediates seen by HX are taken into account. The results support the previous HX measurements of intermediate populations, show how intermediates can elude the standard melting analysis, and illustrate how they confuse the analysis when they are significantly populated within the melting transition region.

Keywords: cytochrome *c*; folding; hydrogen exchange; stability; melting; two-state

Partially folded protein intermediates can be very difficult to detect and study even though they may be important for both kinetic and equilibrium properties. Here, we consider how cryptic intermediates can affect the classical analysis of protein stability.

Classical melting analysis provides the major method for measuring the stabilization free energy of protein molecules and changes in stability imposed by mutations and other perturbants. Denaturant, temperature, or pressure is used to drive proteins through their global unfolding transition where one can measure the population of unfolded (U) and native (N) forms. The U/N ratio measured through the melting transition region is then transformed into stabilization free energy (Equation 1) and extrapolated (Pace, 1986) to obtain the equilibrium stability at native conditions.

$$\Delta G_U = -RT \ln K_U = -RT \ln U/N. \quad (1)$$

One generally assumes two-state unfolding, i.e., that only the fully native and fully unfolded forms contribute to the measured data. This assumption is often justified by the fact that the measured melting behavior does not obviously reveal intermediate forms and fits well to the standard two-state melting equation (Santoro & Bolen, 1988). However, many observations (kinetic folding, molten globule structure, hydrogen exchange) now show that proteins are able to stabilize intermediates that lie between the N and U states (Englander, 2000).

A different method, the native state hydrogen exchange (HX) experiment, can measure equilibrium stability parameters directly under native conditions (Bai et al., 1995a, 1995b). This capability depends on the thermodynamic truism that proteins must cycle between N and U even under native conditions. When a protein provides probe hydrogens that exchange only from the globally unfolded state, their measured HX rate can be used to compute the population of the U state and therefore the free energy of the global unfolding equilibrium at the condition of the HX measurement (Bai et al., 1994; Huyghues-Despointes et al., 1999). When partially unfolded intermediates dominate the HX of the hydrogens that they expose, the identity and free energy of these forms can be determined.

The classical melting extrapolation and the direct HX measurement have found the same global stability values in a number of cases but not in some others. For oxidized equine cytochrome *c* (cyt *c*), stability determined by classical melting and by HX differs dramatically, by almost 3 kcal/mol. We find that this discrepancy is due to a failure of the two-state assumption. Partially unfolded intermediates reach significant population in the melting region even though the protein passes the usual tests for two-state folding. The presence of intermediates reduces the measured slope of the melting transition and therefore the extrapolated free energy between N and U.

Here, we analyze this case and consider some implications.

Results

Two-state melting analysis

The equilibrium unfolding of cyt *c* as a function of guanidinium chloride (GmCl) concentration was measured by fluorescence (Fl) and by circular dichroism (CD) at 222 nm (Fig. 1A). The melting data were fit by the equation of Santoro and Bolen (1988), which

Reprint requests to: Leland Mayne, University of Pennsylvania, Department of Biochemistry and Biophysics, School of Medicine, Philadelphia, Pennsylvania 19104; e-mail: leland@HX2.Med.Upenn.Edu.

Abbreviations: HX, hydrogen exchange; Fl, fluorescence; CD, circular dichroism; GmCl, guanidinium chloride; cyt *c*, oxidized equine cytochrome *c*.

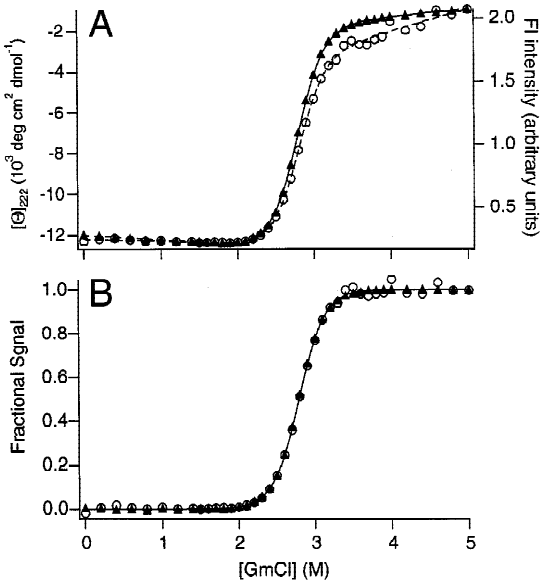

Fig. 1. A: Denaturant melting curves for equine cyt *c*, measured by fluorescence (○) and by circular dichroism (▲) at 222 nm (pDr 7, 30 °C). **B:** Normalized by fitting to the two-state equation of Santoro and Bolen (1988).

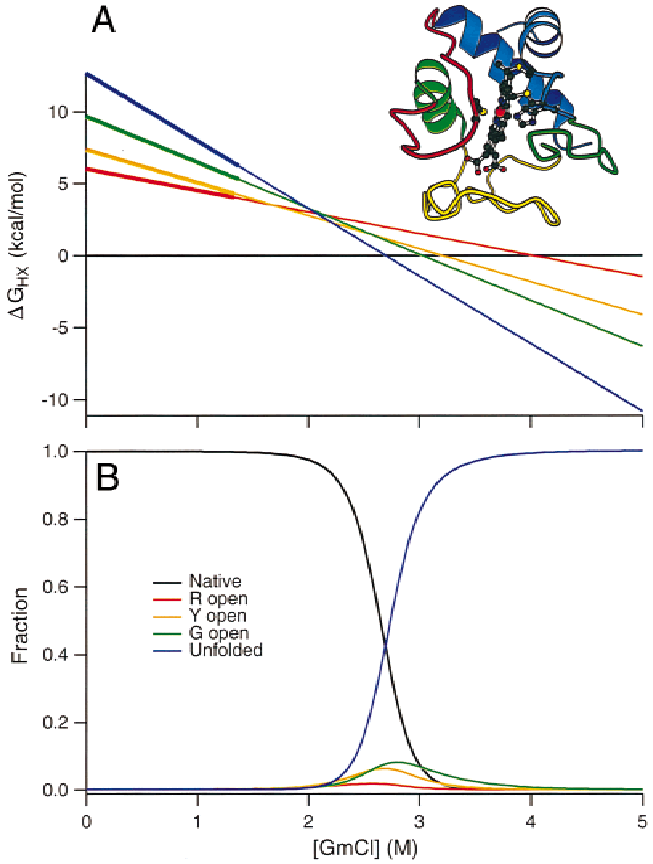

Fig. 2. Unfolding in oxidized cyt *c* measured by native state HX. **A:** Summary of results (Bai et al., 1995b) for the free energy levels, relative to *N*, of the globally unfolded (blue open) and partially unfolded intermediate states (red open, yellow open, green open) as a function of GmCl concentration (pDr 7, 30 °C). HX measurements made below 1.5 M GmCl (shown in bold) are extrapolated forward through the melting region. The free energy levels are color coded to match the molecular segments in cyt *c* (inset) thought to cooperatively unfold to produce each intermediate state. **B:** Fractional populations of the various states computed from their free energy levels.

assumes that the data between the sloping baselines reflect the populations of *U* and *N* only (the two-state assumption) and that the free energy gap between *U* and *N* (ΔG_U) varies linearly with denaturant. Both the fluorescence and the CD data are well fit by the two-state equation and when normalized yield identical melting curves (Fig. 1B). Such results are normally taken as convincing evidence that a protein melts in a two-state manner.

The apparent *U*/*N* ratio obtained through the optically-detected melting transition using the two-state assumption can be plotted as ΔG_U vs. denaturant and linearly extrapolated to zero denaturant concentration (Pace, 1986). The results yield a free energy for global unfolding at zero denaturant of 10.0 ± 0.4 kcal/mol and a slope (*m*) for the linearly extrapolated curve of 3.6 ± 0.1 kcal/mol/M (pDr 7, 30 °C; ±1 standard error).

Intermediates by hydrogen exchange

Figure 2A summarizes native state HX results measured for cyt *c* at low denaturant (pDr 7 and 30 °C (Bai et al., 1995b)) and extrapolates the results forward through the melting transition region where *U* crosses *N*. It appears that the highest energy (blue) unfolding represents the acquisition of the globally unfolded state (Bai et al., 1994, 1995b). The free energy found at zero denaturant for the blue unfolding is 12.7 kcal/mol (proline corrected) and the *m* value is 4.6 kcal/mol/M, in striking disagreement with the values found by optically detected melting data. The uncertainty in the HX values is about 0.4 kcal/mol in unfolding free energy and 0.1 kcal/mol/M in the *m* value.

The native state HX results summarized in Figure 2A further indicate that the cyt *c* molecule reversibly unfolds to a number of discrete intermediate forms as well as the fully unfolded state. It is often stated that denaturant destabilizes intermediates and makes them less apparent. Figure 2A illustrates the point. When denaturant is increased, partially unfolded intermediates are destabilized relative to *U*. However, the intermediates are promoted relative to *N*. *U* is promoted even more sharply (the blue unfolding) because it exposes more denaturant-sensitive surface.

Figure 2B shows that the intermediate states are infinitesimally populated at low denaturant. However, they become significantly populated within the melting transition zone where they account for up to about 20% of the total protein. This is contrary to the two-state assumption and the apparent two-state results of the classical melting analysis.

In a kinetic sense, the intermediates appear to form, one from another, by the reversible cooperative unfolding and refolding of the units of structure shown in Figure 2A (Bai et al., 1995b; Xu et al., 1998). Available information indicates that the red segment unfolds as a cooperative unit, the yellow intermediate involves the concerted unfolding of the red plus yellow segments, the green unfolding represents these three together, and the blue unfolding represents acquisition of the *U* state. However, the analysis described here depends not at all on whether the intermediates are on or off the kinetic folding pathway and only weakly on their detailed identity.

Multistate analysis

We want to consider how the population of partially folded intermediates can affect melting behavior. The denaturant-dependent populations measured by HX (Fig. 2B) and extrapolated into the melting region can be used to predict multistate melting curves that would be measured by various optical and HX probes. The normalized melting curves were predicted according to Equation 2:

$$S = \left(W_N N + \sum W_i I_i + W_U U\right) \Big/ \left(N + \sum I_i + U\right) \qquad (2)$$

where N, I_i, and U are the concentration of each species (as a function of denaturant) taken from Figure 2B, the W values represent the normalized contribution of each species to the measurement, and S is the normalized signal. In the following, the equation for each probe signal is written to make S move from 0 to 1 as denaturant increases.

The fluorescence signal, due to the single tryptophan residue of cyt *c* (Trp59), is assigned solely to U ($W_U = 1$; $W_N = W_i = 0$). In this approximation, the intermediates are assumed to represent relatively collapsed states with Trp59 much closer to the heme than the characteristic resonance energy transfer distance of 32 Å so that their fluorescence is almost completely quenched. The normalized fluorescence signal at any denaturant concentration is then given by Equation 3:

$$S_{\mathrm{Fl}} = U \Big/ \left(N + \sum I_i + U\right). \qquad (3)$$

The CD_{222} signal at any denaturant concentration is given by Equation 4. The contribution of each intermediate to helical CD at 222 nm was approximated according to the number of helical residues in each unit (Bushnell et al., 1990), namely 25 in the blue unit, 7 in green, 5 in yellow, and 4 in red for a total of 41 helical residues. In this approximation, $W_N = 1$, $W_U = 0$ (blue open), and the W_i values for the red open, yellow open, and green open intermediates are 0.90, 0.78, and 0.61, respectively.

$$S_{\mathrm{CD}} = 1 - \left(N + \sum W_i I_i\right) \Big/ \left(N + \sum I_i + U\right). \qquad (4)$$

These equations show that the midpoint of the measured melting curve occurs where U is close to the sum of all non-U species, i.e., where $U \approx N + \sum I_i$ and not $U = N$ as in two-state unfolding. The HX results in Figure 2B show that near the midpoint the cyt *c* intermediates ($\sum I_i$) account for a significant fraction, about 20%, of the total protein. Accordingly, the usual two-state calculation for U/N through the melting zone will be in error. The apparent midpoint can be shifted. The melting curve appears significantly broader than it would be in two-state unfolding, decreasing the m value. The effect on the m value dominates so that the linear extrapolation to zero denaturant finds an artificially low free energy gap between the N and U states.

Figure 3A shows that the melting curves predicted for fluorescence (dashed) and CD_{222} (solid) fall close to each other, indicating that the prediction is not very sensitive to the detailed assumptions. Also the multistate HX results, measured well below the melting transition, closely predict the measured melting data. Uncertainty in the predicted apparent midpoint is ~0.05 M GmCl, which translates into $\Delta\Delta G^\circ$ ~0.2 kcal/mol. The predictive analysis used has no free parameters although some assumptions are necessary concerning the contribution of each intermediate to the signal in question. The melting results can be fit even more precisely by the six parameter two-state equation (Fig. 1B), but this equation will fit well to a wide range of melting curves, even when they are not two-state, and will discriminate between two-state and multistate melting only in extreme cases. This fact was stressed by Santoro and Bolen (1988) but has often been overlooked in subsequent work.

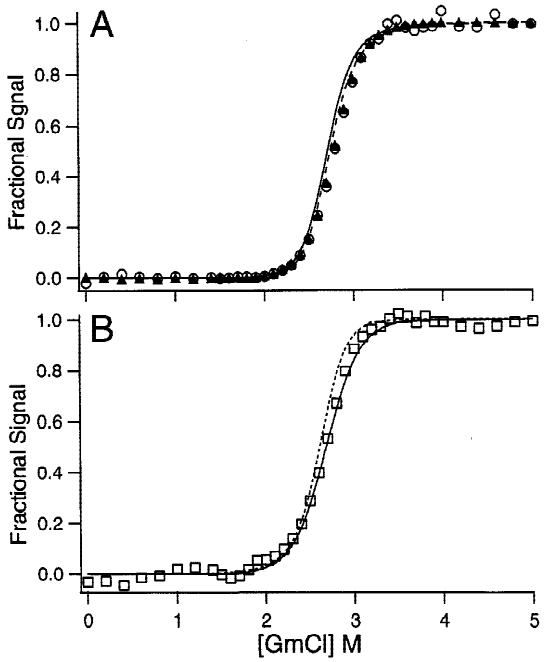

Fig. 3. Comparison between normalized optically measured melting data and the predicted multistate melting curves. **A:** Fluorescence (○) and CD_{222} (▲) data compared with predicted fluorescence (dashed) and CD_{222} (solid) melting curves, from Equations 3 and 4 together with the HX results in Figure 2. **B:** CD_{695} data with the two-state fit (solid) and the predicted curve (dotted) from Equation 5 and the HX results in Figure 2.

The 695 nm probe

Cyt *c* provides an unusual optical probe, an absorbance band at 695 nm, due to the Met80-S to heme iron ligation. Melting followed by absorbance at 695 nm has been seen to slightly precede the global melting observed by fluorescence and CD_{222}, suggesting that cyt *c* may not unfold in a two-state manner (Myer, 1984).

This behavior is predicted by the population of intermediate states shown in Figure 2. Previous work shows that loss of the A_{695} band provides a direct measurement of the unfolding of the red loop, which contains the Met80 residue, and that the red loop is unfolded also in all of the higher energy forms of cyt *c* (Xu et al., 1998). In this case, $W_N = 1$, where the ligation is intact, and $W_i = W_U = 0$, leading to Equation 5:

$$S_{695} = 1 - N \Big/ \left(N + \sum I_i + U\right). \qquad (5)$$

Figure 3B compares melting data observed by CD at 695 nm with the two-state fit (solid) and the multistate prediction from Equation 5 (dotted).

ΔG_U and ΔG_{HX}

The unfolding behavior determined from the HX rate of hydrogens that exchange only through the globally unfolded U state ($W_U = 1$; $W_N = W_i = 0$) measures the value in Equation 6:

$$S_{HX} = U \Big/ \left(N + \sum I_i + U \right) \approx U/N. \quad (6)$$

Equation 6 for HX looks like Equation 3 for fluorescence but there are crucial differences. Native state HX data are taken under native conditions, far below the melting region, where the denominator in Equation 6 is dominated by N ($N \approx 1 \gg (\sum I_i + U) \approx 0$). Thus, the global unfolding is measured directly, independently of partially unfolded intermediates. Similarly, the separate intermediate states and their free energy levels are each individually measured.

Figure 4 compares the global unfolding behavior inferred from two-state melting analysis and from native state HX, with both measured under the same solution conditions but in different ranges of denaturant. The melting results extrapolate to a global unfolding free energy at zero denaturant of 10.0 kcal/mol compared to 13 kcal/mol for the HX analysis. A part of the disagreement in free energy, 0.3 kcal/mol, is accounted for by the proline isomerization effect (Bai et al., 1994) (see Materials and methods). Most of the discrepancy, however, stems from the fact that intermediates reach significant population in the melting region, invalidating the assumption of two-state melting.

A difference between HX and melting data, like that in Figure 4, was seen before at higher temperature (50 °C) (Bai et al., 1994). The apparent HX upcurve, above the linearly extrapolated melting data, was previously suggested to show the behavior expected from the theoretical denaturant melting formulations of Tanford (Aune & Tanford, 1967) and Schellman (1987, 1990). The present results show instead that the discrepancy results from the significant population of intermediates in the melting zone and the depressed linear extrapolation that is found when two-state behavior is incorrectly assumed.

Discussion

One is interested in the properties and roles of protein folding intermediates and the ways in which they can be measured or can escape detection. Here, we have considered the relationship of folding intermediates to classical melting experiments.

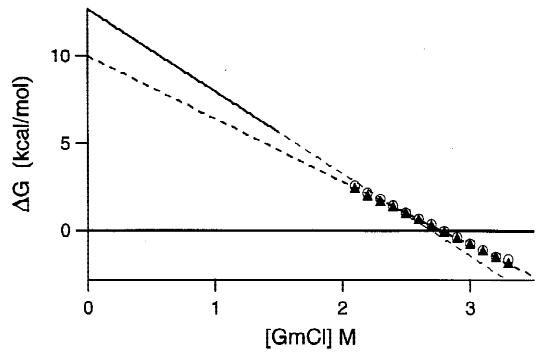

Fig. 4. Free energy results for cyt *c* by HX and by melting analysis. Optical melting data measured through the melting transition were fit by the Santoro–Bolen equation and linearly extrapolated (dashed line) back to zero denaturant. HX results measured at low denaturant (solid) and extrapolated forward (dashed) through the transition region are for several globally exchanging amides in the native protein.

Equilibrium melting

The present exercise deals with a particular case in which the classical and HX analyses of equilibrium melting disagree. The results obtained show that the discrepancy is due to a significant divergence from the two-state behavior generally assumed in classical melting analysis. In the presence of significantly populated intermediates, the midpoint of an observed melting curve does not occur where $U = N$, two-state analysis of the melting curve does not yield the correct U/N ratio, the slope of the curve of ΔG_U vs. denaturant is too low, and the linear extrapolation to native conditions yields an artificially low free energy.

The intermediates that cause these anomalies are not detected by the fit of melting data to the usual two-state equation. In the cyt *c* case, the presence of intermediates is not detected even by a comparison of CD and fluorescence melting curves. It seems probable that the distributed CD and fluorescence probes common to most proteins will often fail to provide distinct melting curves that would signal the presence of significant intermediates. The remaining equilibrium melting test involves a comparison of van't Hoff and calorimetric enthalpies. Calorimetric melting is not in itself sensitive to the presence of intermediates, and the comparison can be limited by curved baselines, high melting temperature, and thermal irreversibility or aggregation.

In short, melting data are sensitive to the presence of intermediates only when they achieve substantial population in the melting region and even then will often fail to detect them.

Equilibrium HX

Huyghues-Despointes et al. (1999) list 17 proteins for which available HX rate data and extrapolated melting data yield stabilization free energies that are in reasonable agreement. The present results deal with a case in which these determinations differ because cryptic intermediates contribute to the measurement. Accordingly, one possible explanation for the agreement found by Huyghues-Despointes et al. is that intermediates do not in general exist. Another possibility is that intermediates exist but often do not achieve significant population within the transition region. This might happen for several reasons. It is especially noteworthy that even the lowest lying intermediates detected by native state HX tend to have free energies that are remarkably high, generally 6 kcal/mol and more (Bai et al., 1995b; Chamberlain et al., 1996; Fuentes & Wand, 1998a, 1998b; Hollien & Marqusee, 1999; Llinas et al., 1999). The proteins listed by Huyghues-Despointes et al. that appear to be two-state folders all have global unfolding free energy less than 8 kcal/mol. One can picture Figure 2A with the global unfolding curve (blue) moved down to this level. In this case, discrete intermediates, although they may be present and functionally important, for example in kinetic folding processes (Creighton et al., 1996), will not be resolved in the native state HX analysis, will not be significantly populated in the melting transition zone, and will not affect the classical melting analysis.

Kinetic folding

Analogous detection problems occur in kinetic studies. Proteins that appear to fold in a two-state kinetic manner (Jackson, 1998)

are often asserted to fold without intermediates. It is noteworthy, however, that most kinetic intermediates are unlikely to be seen even if they are important in kinetic folding. An intermediate will accumulate and become measurable in the usual kinetic folding experiment only when it occupies a free energy well that is deeper than all prior wells and is blocked by a barrier that is higher (trough to peak) than all prior barriers. Further, it appears that the intrinsic rate-limiting barrier in two-state folding may occur as an initial step in kinetic folding (Sosnick et al., 1996). In this case, pathway intermediates will be invisible. They will kinetically accumulate and become visible only when the protein happens to encounter a subsequent nonintrinsic barrier, which has been pictured as a time-consuming error correction process ("misfold-reorganization" barrier; Sosnick et al., 1994).

Conclusions

The firmly entrenched view that protein molecules tend to be monolithically cooperative two-state structures is based largely on observations that will fail to demonstrate intermediates even when they are present. These considerations lead to an important perspective. It seems remarkably difficult in general to distinguish multistate protein folding from apparent two-state folding. Neither equilibrium unfolding nor kinetic refolding provide a dependable test. In favorable cases, the native state HX experiment can separately identify and characterize intermediates but this approach also is limited, especially by the requirement for sufficient dynamic range in the free energy dimension, and also by other constraints (Bai & Englander, 1996; Englander, 2000; Milne et al., 1999).

The present work extends previous indications of this situation to the case of equilibrium melting. An understanding of the multistate nature of protein molecules and its role in equilibrium and kinetic folding is important for all aspects of protein science. The problem stands as a fundamental challenge for future work.

Materials and methods

Commercial equine cytochrome *c* was used (Sigma Chemical Co., St. Louis, Missouri). Melting experiments used an Aviv model 202 CD spectrometer with fluorescence and autotitrator attachments. To match HX conditions (Bai et al., 1995b), all experiments were in D_2O at pDr 7.0 and 30 °C.

In the HX results shown (Fig. 2) and in the calculations that lead from Figure 2B to Figures 3 and 4, a small correction was applied to account for the effect of proline *cis-trans* isomerization to make the populations measured by HX applicable to the populations that will occur in the equilibrium melting experiment. In kinetic HX experiments, proline residues do not have time to relax to isomeric equilibrium in the transiently populated unfolded forms before HX goes to completion and so they maintain their native isomers. In melting experiments, the prolines do relax to isomeric equilibrium and therefore reduce the free energy of the unfolded forms according to Equation 7 (Sharp & Englander, 1994):

$$\Delta\Delta G_U = -RT \ln(1 + K_{iU})/(1 + K_{iN}) \quad (7)$$

where K_i is the equilibrium constant for *cis-trans* isomerization of each proline in the unfolded (U) and native (N) forms (Bai et al., 1994). K_{iN} was taken as zero since the prolines in native cyt *c* are all *trans*. In the unfolded state, the isomerization constant depends on the prior residue (Reimer et al., 1998). The corrections are 0.15 for the two prolines in the red loop, 0.05 for one in the yellow loop, and 0.09 for one in the green loop, which add up to relax the unfolding free energy measured by HX by 0.15 kcal/mol for the red-open intermediate, 0.20 for the yellow-open intermediate, and 0.29 for the green intermediate and for global unfolding.

Acknowledgments

This work was supported by grants from the National Institutes of Health and the Mathers Charitable Foundation.

References

Aune K, Tanford C. 1967. Thermodynamics of the denaturation of lysozyme by GdmCl. *Biochemistry 8*:4586–4590.

Bai Y, Milne JS, Mayne L, Englander SW. 1994. Protein stability parameters measured by hydrogen exchange. *Proteins Struct Funct Genet 20*:4–14.

Bai Y, Englander JJ, Mayne L, Milne JS, Englander SW. 1995a. Thermodynamic parameters from hydrogen exchange measurements. *Methods Enzymol 259*:344–356.

Bai Y, Sosnick TR, Mayne L, Englander SW. 1995b. Protein folding intermediates: Native-state hydrogen exchange. *Science 269*:192–197.

Bai Y, Englander SW. 1996. Future directions in folding: The multi-state nature of protein structure. *Proteins Struct Funct Genet 24*:145–151.

Bushnell GW, Louie GV, Brayer GD. 1990. High-resolution three dimensional structure of horse heart cytochrome *c*. *J Mol Biol 213*:585–595.

Chamberlain AK, Handel TM, Marqusee S. 1996. Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nature Struct Biol 3*:782–787.

Creighton TE, Darby NJ, Kemmink J. 1996. The roles of partly folded intermediates in protein folding. *FASEB J 10*:110–118.

Englander SW. 2000. Protein folding intermediates and pathways studied by hydrogen exchange. *Annu Rev Biophys Biomol Struct 29*:213–238.

Fuentes EJ, Wand AJ. 1998a. Local dynamics and stability of apocytochrome b562 examined by hydrogen exchange. *Biochemistry 37*:3687–3698.

Fuentes EJ, Wand AJ. 1998b. Local stability and dynamics of apocytochrome b562 examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry 37*:9877–9883.

Hollien J, Marqusee S. 1999. A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. *Biochemistry 38*:3831–3836.

Huyghues-Despointes BMP, Scholtz JM, Pace CN. 1999. Protein conformational stabilities can be determined from hydrogen exchange rates. *Nat Struct Biol 6*:910–912.

Jackson SE. 1998. How do small single-domain proteins fold? *Fold Des 3*:R81–91.

Llinas M, Gillespie B, Dahlquist Frederick W, Marqusee S. 1999. The energetics of T4 lysozyme reveal a hierarchy of conformations. *Nat Struct Biol 6*:1072–1078.

Milne JS, Xu Y, Mayne LC, Englander SW. 1999. Experimental study of the protein folding landscape: Unfolding reactions in cytochrome *c*. *J Mol Biol 290*:811–822.

Myer YP. 1984. Ferricytochrome *c*, refolding and the methionine 80-sulfur-iron linkage. *J Biol Chem 259*:6127–6133.

Pace CN. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol 131*:266–280.

Reimer U, Scherer G, Drewello M, Kruber S, Schutkowski M, Fischer G. 1998. Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J Mol Biol 279*:449–460.

Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry 27*:8063–8068.

Schellman JA. 1987. Selective binding and solvent denaturation. *Biopolymers 26*:549–559.

Schellman JA. 1990. A simple model for solvation in mixed solvents. Applications to the stabilization and destabilization of macromolecular structures. *Biophys Chem 37*:121–140.

Sharp KA, Englander SW. 1994. How much is a stabilizing bond worth? *Trends Biochem Sci 19*:526–529.

Sosnick TR, Mayne L, Hiller R, Englander SW. 1994. The barriers in protein folding. *Nat Struct Biol 1*:149–156.

Sosnick TR, Mayne L, Englander SW. 1996. Molecular collapse: The rate-limiting step in two-state cytochrome *c* folding. *Proteins Struct Funct Genet 24*:413–426.

Xu Y, Mayne L, Englander SW. 1998. Evidence for an unfolding and refolding pathway in cytochrome *c*. *Nat Struct Biol 5*:774–778.