Supplementary Material for

High energy intermediates in protein unfolding characterized by thiol labeling under native-like conditions

Pooja Malhotra and Jayant B. Udgaonkar*

National Centre for Biological Sciences

Tata Institute of Fundamental Research

Bangalore 560065, India.

*Corresponding author: Professor Jayant B. Udgaonkar

Fax: 91-80-23636662

Email: jayant@ncbs.res.in

SI Methods

Fluorescence-Monitored Equilibrium and Kinetic Experiments. For equilibrium unfolding studies the protein was incubated in different concentrations of GdnHCl for 6 h prior to measurement. To monitor unfolding kinetics, the protein dissolved in native buffer was diluted into unfolding buffer containing a final GdnHCl concentration in the range of 2-8 M. The dead time of stopped flow mixing was 5 ms. Unfolding in 2-4 M GdnHCl was also performed by manual mixing with a dead time of 10 s. For refolding, the protein was incubated in unfolding buffer (3.5 M GdnHCl) for 6 h prior to the experiment. The refolding of the protein was initiated by diluting unfolded protein into native buffer to final concentrations of denaturant ranging between 0.2 and 1.5 M GdnHCl. Additionally, manual mixing experiments were carried out in the range of 0.6 to 1.5 M GdnHCl. For all fluorescence experiments, dithiothreitol was added in 100-fold excess of the protein concentration, to prevent protein dimerization.

Calculations of Solvent Accessible Surface area (SASA) and Depth of Residues. SASA was calculated for each of the residues shown in Figure 1A from the PDB file 1IV7.pdb, using the Lee and Richards algorithm (psa version 1.1 software).[1,2] The probe radius was set to 1.4Å. Reference values for the residues in the unfolded state were calculated for an Ala-X-Alatripeptide in the extended conformation, where X denotes any of the four residues under study. The depth of each residue was calculated using the online server DEPTH.[3,4]

SI Text

The folding kinetics at pH 8 show marginal differences from the kinetics at pH 7. The folding and unfolding mechanisms of MNEI appear to be the same, but there are minor differences in the sensitivity of different kinetic phases to different probes. Unlike at pH 7, where it is silent to intrinsic tryptophan fluorescence change, and where it can be detected only by

monitoring folding using ANS fluorescence, the very slow phase at pH 8 is accompanied by an intrinsic tryptophan fluorescence change. Also, unlike at pH 7, the observable kinetic phases do not account for the entire change in fluorescence that occurs during folding. About 15 % of the fluorescence change occurs in a burst phase (Figure 4C), which could be detected only by ANS fluorescence measurement at pH 7.[5]

Analysis of Rates and Free Energies to Estimate Barrier Heights. An intermediate can be detected by SX as long as its transition to another state is slower than the labeling reaction itself. The measured values of k_{op}, k_{cl}, and free energies, reported in Table 1, can be used to calculate each of the energy barriers shown in the free energy profile in Figure 7, in 0 M GdnHCl, providing a quantitative estimation of how long the protein resides in a particular state. Given the extremely low populations of the intermediates detected, it is important to show that each of the intermediate states is populated long enough for the labeling to take place from that particular state. Although it has been assumed that the intermediates occur sequentially along the unfolding pathway, it should be noted here that the present data has not been used to build a kinetic scheme. The opening rates measured by SX, for each of the thiols probed, is the rate at which the N state, in which the cysteine is buried, transforms into the intermediate (I) state, in which the cysteine is exposed to labeling. Likewise, the measured closing rates are the rates at which I, in which the cysteine is exposed, transitions to the N state.

Thiol labeling of Cys13 identifies I_1, in which the C13 side-chain is exposed to solvent. Hence, the k_{op} measured for Cys13 (>2 s^{-1}) is the rate at which the N to I_1 barrier is crossed. The k_{cl}, measured for Cys13, is therefore the rate at which I_1 reverses back to the N state. Since the equilibrium constant K_{op} is 0.01 for the N to I_1 transition (Figure 6D), and k_{op} has a lower limit of 2 s^{-1}, hence k_{cl} would have a lower limit of 2000 s^{-1}. In addition to closing back to the N state,

I_1can also transition to I_2. The difference in the heights of the barrier from N to I_2 and the barrier from I_1 to I_2 is equivalent to the free energy difference between I_1 and the N state; i.e. the ΔG_{op} measured for I_1 (2.7 kcal mol^{-1}). Since the height of the barrier from N to I_2 corresponds to the k_{op} measured for Cys42, the rate of crossing the I_1 to I_2 barrier can be calculated, as 2 s^{-1}.The barrier for I_1 directly transitioning to the U state would necessarily be higher than this. Since the labeling of an exposed cysteine takes place at a rate of 200 s^{-1}, at the highest PDS concentration used (10 mM), the conversion from I_1 to I_2 is much slower than the labeling process itself. Hence, the labeling of the C13 side-chain, in I_1, is limited only by the opening and closing rates between N and I_1.

Thiol labeling of C42 identifies I_2, which is formed from the N state, at a rate equal to the k_{op} of Cys42 (0.01 s^{-1}). Even though the N to I_1 rate (> 2s^{-1}) and the I_1 to I_2 rate (2 s^{-1}) arefast , the even faster backward rate, from I_1 to N (> 2000 s^{-1}), results in the net rate from N to I_2 being much lower. Once formed, I_2 can either close back to the N state, at a rate equivalent to the k_{cl} (1790 s^{-1}), measured for Cys42, or it can cross the major unfolding barrier and transition to the I_3 state. The calculation for the I_2 to I_3 rate is similar to that described earlier for the rate of I_1 transitioning to I_2. The k_{op} of I_3 (1.58×10^{-5}) and the ΔG_{op} of I_2 (6 kcal mol^{-1}) yield a value of 0.25 s^{-1} for the I_2 to I_3 transition. The rate of labeling of an exposed cysteine is 5700 s^{-1}, at the highest DTNB concentration used (60 mM). These calculations therefore suggest that once I_2 is formed, the exposed C42 side-chain gets labeled much faster than the rate at which it closes back to the N state or transitions to the I_3 state.

The apparent rate of formation of the intermediate I_3 from the N state is equivalent to the k_{op} measured for Cys74 (1.58×10^{-5}) and Cys63 (1.7×10^{-4}). This rate is also much lower that the individual rates of N to I_1, $I_1$to I_2 and I_2 to I_3 (Figure 7) due to the very high closing rates. The

labeling of the C74 and C63 residues in the intermediate I_3 depends upon the rates at which I_3 crosses the major barrier back to the native side and the rate at which it converts into the unfolded state. The former rate is the k_{cl} measured for the two proteins (1100 s^{-1} for Cys74 and 183 s^{-1} for Cys63). (See the main text for why the rates measured for k_{cl} from I_3 are different for the Cys74 and Cys63 proteins). Figures 8A-C show how the relative barriers between N, I_3 and U are modulated in different GdnHCl concentrations and provide an estimate of the minimum lifetime of I_3 for labeling to occur in it.

SI References

(1) Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.55*, 379–IN4.

(2) Richards, F. M. (1977) Areas, Volumes, Packing, and Protein Structure. *Annu. Rev. Biophys. Bioeng.6*, 151–176.

(3) Chakravarty, S., and Varadarajan, R. (1999) Residue depth: a novel parameter for the analysis of protein structure and stability. *Structure7*, 723–32.

(4) Tan, K. P., Varadarajan, R., and Madhusudhan, M. S. (2011) DEPTH: a web server to compute depth and predict small-molecule binding cavities in proteins. *Nucleic Acids Res.39*, W242–8.

(5) Patra, A. K., and Udgaonkar, J. B. (2007) Characterization of the folding and unfolding reactions of single-chain monellin: evidence for multiple intermediates and competing pathways. *Biochemistry46*, 11727–43.

Supplementary figures

Figure S1. Determination of the bimolecular rate constant (k_b) of labeling. The dependence of k_b on GdnHCl concentration was measured for (A) Cys74 (B) Cys63 (C) Cys42 and (D) Cys13. Measurements were with the protein (filled symbols), at GdnHCl concentrations higher than that corresponding to the midpoint of unfolding, and with the corresponding pentapeptide (open symbols), at low denaturant concentrations. The solid lines are linear regression fits through the data. The error bars represent the spread in the measurements from at least two different experiments.

Figure S2. Comparison of the bimolecular rate constants (k_b) of labeling of Cys42 by PDS (black) and DTNB (cyan), at different GdnHCl concentrations. Measurements were with the protein (filled symbols), at GdnHCl concentrations higher than the C_m, and with the corresponding pentapeptide (open symbols), at low denaturant concentrations. The solid lines are linear regression fits through the data. The error bars represent the spread in the measurements from at least two different experiments.

Figure S3. Fluorescence monitored unfolding kinetic rates, at pH 8.0, 25^{o}C, for Cys74 (○), Cys63 (Δ), Cys42 (□) and Cys13 (▽). The solid lines through the data are fits to equation2. The error bars represent the spread in the measurements from at least two different experiments.

Figure S4. Fluorescence monitored refolding kinetics of Cys74 and Cys63 at pH 8.0, 25 ^{o}C. The observed rates of refolding (upper panel) and relative amplitudes (lower panel) of the very fast (▲), fast (●), slow (▼) and very slow (■) phases have been measured in different GdnHCl concentrations for Cys74 (left panels) and Cys63 (right panels). The error bars represent the spread in the measurements from at least two different experiments. The lines have been drawn by inspection only.

Figure S5. Thiol labeling of Cys13 at high denaturant concentrations. The free energy of opening, ΔG_{op}, was measured between 0 and 2 M GdnHCl (□), by allowing the protein to equilibrate at each GdnHCl concentration and then initiating the labeling reaction by the addition of PDS, without any change in denaturant concentration. ΔG_{op} was measured in the range of 2 to 7 M GdnHCl (□) by diluting the native protein to a high GdnHCl concentration, in the presence of 10-fold excess of PDS. k_{obs} was obtained by fitting the resulting kinetic trace to a single-exponential equation, and showed a linear dependence on PDS concentration, till 6 M GdnHCl, confirming that labeling takes place in the SX2 limit at high denaturant concentrations as well. ΔG_{op} was therefore calculated using equation 6. The free energy associated with global unfolding $\Delta G_u(\Delta)$, was obtained from fluorescence-monitored equilibrium unfolding experiments. The error bars represent the spread in the measurements from at least two different experiments.

Figure S1

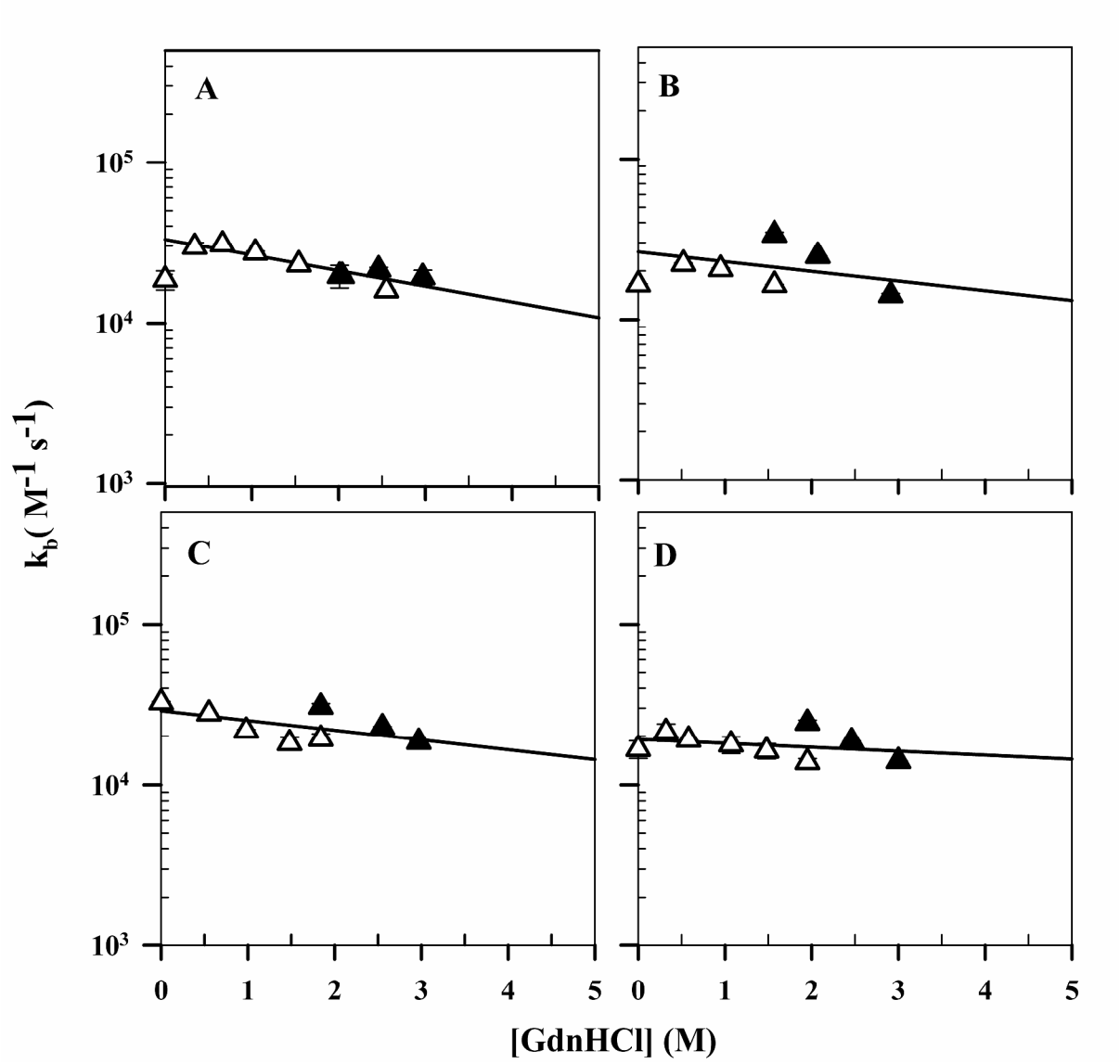

Figure S2

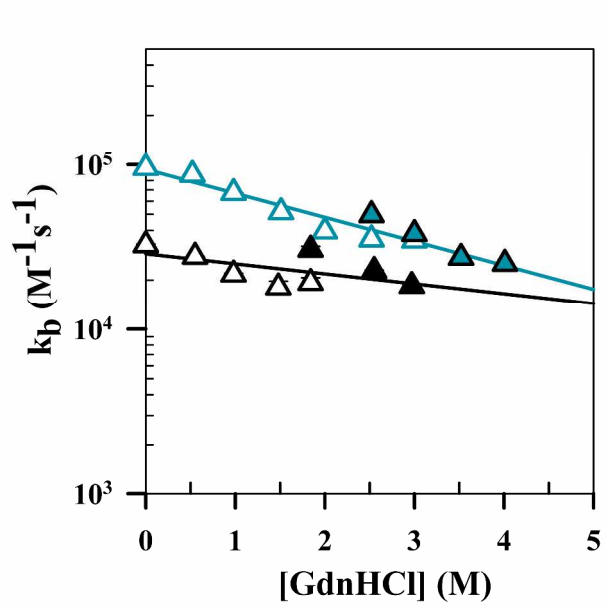

Figure S3

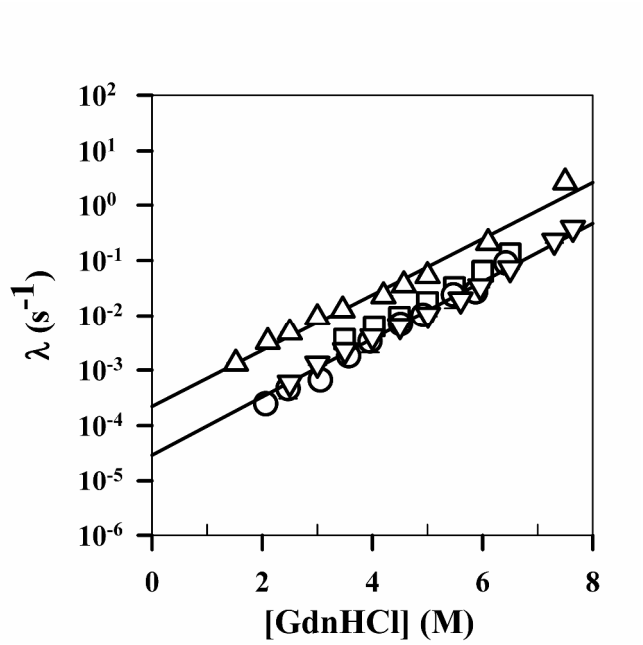

Figure S4

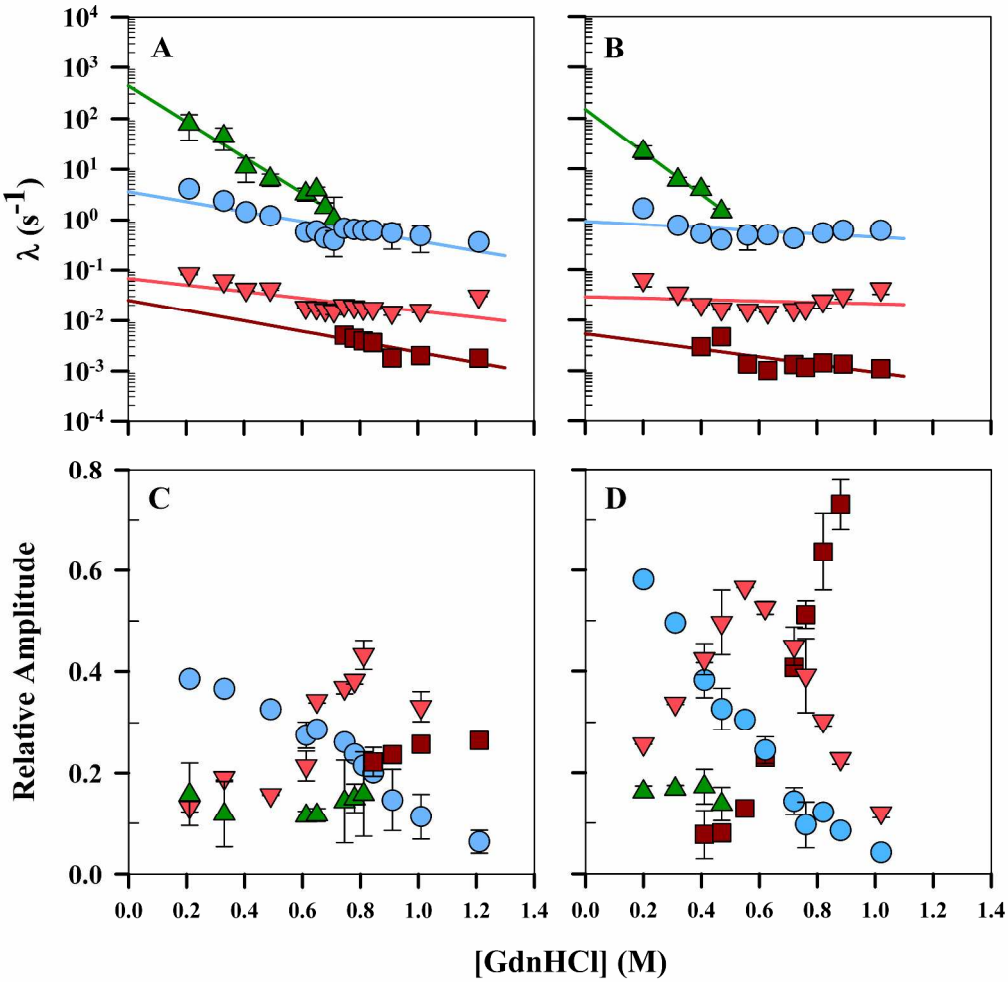

Figure S5