Supporting Information

Halskau *et al.* 10.1073/pnas.0709881105

SI Materials and Methods

Differential Scanning Calorimetry (DSC). Hen-egg-white lysozyme (HEWL) and bovine *apo*- and *holo*-α-lactalbumin (*holo*-BLA and *apo*-BLA) were obtained from Sigma. The protein solutions for the calorimetric experiments were prepared by exhaustive dialysis against the buffer: 50 mM sodium acetate buffer, pH 4.5, for HEW Lyz; 10 mM phosphate, 5 mM citric acid, 1 mM EDTA, pH 5.6, for *apo*-BLA; and 10 mM phosphate, 5 mM citric acid, pH 5.6, 1 mM $CaCl_2$ for *holo*-BLA. Note that the buffer used for *apo*-BLA always contained 1 mM EDTA to ensure that the protein was in the *apo*, calcium-depleted form. DSC experiments were carried out with a VP-DSC calorimeter from MicroCal as we have previously described (see ref. 1 and references therein). Denaturation of both proteins was found to be highly reversible. Several experiments with different protein concentrations were carried for each protein and absolute heat capacities were calculated from the protein concentration dependence of the apparent heat capacities, as described in ref. 1.

Variable-Barrier Analysis of DSC Data. To eliminate as much as possible the enthalpy fluctuations not associated with folding-unfolding, the original development of the variable-barrier analysis (2) uses as heat capacity reference the baseline for native proteins, as described by Freire (3). Note, nevertheless, that at low temperature, the experimental absolute heat capacity values for both proteins (Fig. 2 *Upper*) were in good agreement with the predictions from Freire's equation, in particular, with the slope of the temperature dependence. Thus, we used as native baselines for the analyses of holo-BLA and HEWL profiles the baselines provided by the low-temperature experimental data. In the case of *apo*-BLA we used a native baseline fixed to the lowest temperature point and with the slope given by Freire's prediction. The general fitting approach is the same as we have previously used in which we obtained a grid of the goodness of the fit as a function of the barrier height (β) (2, 4). Fittings were based on the minimax criterion and, therefore, the maximum deviation (σ_m) is used as goodness of fit. For the three proteins under study here (*apo*- and *holo*-BLA and HEWL), the plots of σ_m versus β show clearly the optimal barrier height value. For BLA, fits of similar quality could be obtained by down-shifting the native baseline. This exercise, however, leads to even lower values for the barrier height.

Electrostatic Calculations. Electrostatic calculations are based on the Tanford–Kirkwood (TK) model as we have previously described in detail (5, 6), but without applying Gurd's correction. The structures used for the calculations were 1DPX.pdb for HEWL and 1F6S.pdb/1F6R.pdb for *holo*/*apo*-BLA. The electrostatic free-energy profiles of Fig. 4 *Lower* were obtained by projecting the TK electrostatic free energy for each conformation onto the number of native residues by using the expression:

$$G(n_U) = -RT\cdot\ln \sum_{n_U} \exp(-G_i/RT),$$

where the G_i is the electrostatic free energy of a given microstate with n_U residues in unfolded stretches and the sum is over all microstates with a given n_U value. In Fig. 4 *Lower* we give this free energy corrected for the degeneration associated with each degree of unfolding, $G'(n_U) = G(n_U) - RT\cdot\ln g(n_U)$, where $g(n_U)$ is the number of microstates compatible with the value of n_U. The degeneration term is simply a statistical term that has practically the same value for BLA and HEWL (since these two proteins have almost the same size) and the corresponding correction is only done for the purpose of facilitating the visual interpretation of the free-energy profiles. Also, the electrostatic energies shown in Fig. 4 *Lower* are referenced to the native state, that is, they are actually $G'(n_U) - G'(0)$.

pH Titration at Individual Residues Monitored by NMR. *holo*-BLA (1.6 mM) (Sigma-Aldrich) was prepared in 10 mM phosphate/5 mM citric acid buffer, 0.1 M NaCl, pH 5.6, 9% D_2O. DSS was added to some of the samples for use as a chemical shift reference. The TOCSY spectra were recorded with 1,024 complex points, 512 increments in T1 and 64 scans at a temperature of 40°C. Both dimensions had a spectral width of 10.5 ppm. The delay for the water suppression was 90 μs and the TOCSY mixing time was 80 ms. A squared cosine window function was used in both F1 and F2. The free induction decays were Fourier transformed into 2 K in both dimensions. Forward linear prediction to 2 K was used in F1. The titration series were stopped at pH 3.5 because of too poor quality of the TOCSY below this pH. TOCSY spectra for all pH values were examined in Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco, 1996–2001). H^N-H^β cross-peaks for aspartic acid and H^N-H^β and H^N-H^γ for glutamic acid were assigned by comparing with the already assigned TOCSY spectrum at pH 5.6 (7). In cases where the H^N-H^β cross-peaks were difficult to assign because of overlapping cross-peaks, the H^α-H^β cross-peaks were used in the pH titration. The relevant chemical shifts were plotted against pH. Nonlinear regression analysis was performed to fit the data by using the program DataFit from Oakdale Engineering. In Fig. 3, $H\beta'/H\gamma'$ and $H\beta''/H\gamma''$ refer to the β or γ protons with lowest and highest chemical shift, respectively.

1. Guzman-Casado M, Parody-Morreale A, Robic S, Marqusee S, Sanchez-Ruiz JM (2003) Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of Thermus thermophilus ribonuclease H. *J Mol Biol* 329:731–743.
2. Muñoz V, Sanchez-Ruiz JM (2004) Exploring protein-folding ensembles: A variable-barrier model for the analysis of equilibrium unfolding experiments. *Proc Natl Acad Sci USA* 101:17646–17651.
3. Freire E (1995) Differential scanning calorimetry. *Protein Stability and Folding. Theory and Practice*, ed Shirley BA (Humana Press, Totowa, NJ), pp 191–218.
4. Godoy-Ruiz R, *et al.* (2008) Estimating free energy barrier heights for an ultrafast folding protein from calorimetric and kinetic data. *J Phys Chem B*, in press.
5. Ibarra-Molero B, Loladze VV, Makhatadze GI, Sanchez-Ruiz JM (1999) Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* 38:8138–8149.
6. Sundd M, Iverson N, Ibarra-Molero B, Sanchez-Ruiz JM, Robertson AD (2002) Electrostatic interactions in ubiquitin: stabilization of carboxylates by lysine amino groups. *Biochemistry* 41:7586–7596.
7. Bartik K, Redfield C, Dobson CM (1994) Measurement of the individual pKa values of acidic residues of hen and turkey lysozymes by two-dimensional 1H NMR. *Biophys J* 66:1180–1184.