J|A|C|S
COMMUNICATIONS

Published on Web 06/01/2005

Measuring pK_a Values in Protein Folding Transition State Ensembles by NMR Spectroscopy

Martin Tollinger,*[,†,‡,§] Lewis E. Kay,[‡,¶] and Julie D. Forman-Kay[†,‡]

Structural Biology and Biochemistry, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada, and Protein Engineering Network Centres of Excellence and Departments of Biochemistry, Medical Genetics, and Chemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Received March 26, 2005; E-mail: martin.tollinger@univie.ac.at

Folding kinetics for a number of proteins have been shown to follow simple monoexponential behavior, indicating that folding occurs without the observable accumulation of folding intermediates.[1] The kinetic folding and unfolding data for such proteins are generally analyzed and interpreted within the framework of transition state theory, where the interconversion between the folded and the unfolded states is limited by the formation of a transition state, representing the ensemble of conformations of highest free energy along the folding pathway.[1] The transition state is, therefore, only transiently populated. Together with structural analyses of the end states of the folding reaction, characterization of the structure of the transition state ensemble (TSE) is of crucial importance for obtaining a comprehensive picture of the folding pathway and for understanding the mechanisms by which proteins fold. Much insight into protein folding has been gained in recent years; however, knowledge of the electrostatic contacts in the transition state of a given protein or in its unfolded state is still very limited. While studies have shown that the strength of electrostatic interactions in the TSE are protein specific, most analyses assume uniform pK_a values for titrating residues (of a given type).[2]

Our recent determination of the pK_a values of the Asp, Glu, and His side chains in both folded and unfolded states of the marginally stable N-terminal SH3 domain of the *Drosophila* protein drk (drkN SH3) enabled calculation of the pH dependence of its stability and demonstrated that simple models of electrostatic interactions in the unfolded state do not adequately account for experimental behavior.[3] Rather, for the drkN SH3 domain, local electrostatic interactions that are present in the unfolded state give rise to nonuniform pK_a values. It can, therefore, be expected that folding transition states, which generally display higher degrees of structure and are considerably more compact than unfolded states, will show even more nonuniformity of electrostatic interactions.

Here, protein folding kinetic data have been obtained for the drkN SH3 domain in order to investigate electrostatic properties of a conserved residue, Asp8, in the folding TSE. The interconversion between folded and unfolded states of this domain has previously been shown to be two-state[4] and is slow on the NMR chemical shift time scale (folding and unfolding rate constants, k_f and k_u, at pH 6.0, 15 °C are 0.99 ± 0.05 and 0.47 ± 0.04 s^{-1}, respectively). Thus, separate sets of resonances for each state are observed in NMR spectra. This enables the use of a longitudinal magnetization exchange experiment[5] to simultaneously obtain folding and unfolding rate constants *under equilibrium conditions without perturbation* of the folding/unfolding equilibrium. In this experiment, magnetization transfer from the folded (F_{exch}) to the unfolded (U_{exch}) state gives rise to a cross-peak (Figure 1A (blue)), whose intensity buildup is directly related to the unfolding rate (blue curves in Figure 1B,C), with a corresponding correlation and buildup (red) for the folding reaction. This experiment is particularly useful where $|\Delta G_{F\rightarrow U}| < \sim 1$ kcal/mol. In cases where $|\Delta G_{F\rightarrow U}| < \sim 3$ kcal/mol, folding kinetics under equilibrium conditions can be studied via line shape analysis and spin relaxation dispersion techniques.[6]

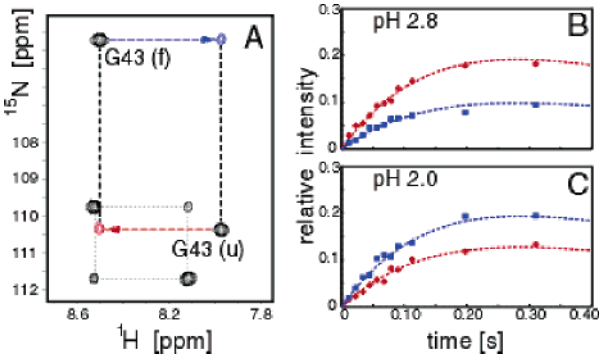

Figure 1. (A) In the longitudinal exchange NMR experiment, a pair of auto peaks corresponding to the folded and unfolded states as well as a pair of exchange cross-peaks, which result from the transfer of magnetization due to folding (red arrow) and unfolding (blue arrow) transitions, are observed for each residue. Only the high-field (^{15}N) portion of a 2D correlation map is shown. (B and C) Relative intensities of folding (red) and unfolding (blue) cross-peaks for residue Gly43 as a function of time (normalized to the intensity of the corresponding auto peak) for two pH values. Folding (k_f) and unfolding (k_u) rate constants are determined simultaneously by fitting the buildup of cross-peaks together with decay of auto peaks (not shown) to analytical expressions for two-site exchange including longitudinal relaxation.[5] All data were recorded on a 1.0 mM sample of wild-type ^{15}N-labeled drkN SH3 domain in 50 mM sodium phosphate, 92% H_2O/8% D_2O, 15 °C at 500 MHz.

To investigate electrostatic interactions in the TSE, folding and unfolding kinetics of the drkN SH3 domain have been measured as a function of pH using this equilibrium NMR approach. The free energy difference between states A and B, $\Delta G_{A\rightarrow B} = G_B - G_A$, is given by[3]

$$\Delta G_{A\rightarrow B} = \Delta G^0_{A\rightarrow B} - RT \sum \ln \frac{([H^+] + K_a^B(i))}{([H^+] + K_a^A(i))} \qquad (1)$$

where $\Delta G^0_{A\rightarrow B}$ is the contribution to free energy from all terms unrelated to pH, K_a^j is the acid dissociation constant for state j = {U,F,‡}, the TSE is denoted by ‡ and the summation is over all titrating groups i. In addition, from transition state theory, it follows that[1]

$$-RT \ln k_f = -RT \ln\left(\frac{\kappa k_b T}{h}\right) + \Delta G_{U\rightarrow ‡} \qquad (2)$$

† Hospital for Sick Children.
‡ Department of Biochemistry, University of Toronto.
¶ Departments of Medical Genetics and Chemistry, University of Toronto.
§ Current address: Department of Biomolecular Structural Chemistry, University of Vienna, A-1030 Vienna, Austria.

10.1021/ja051942c CCC: $30.25

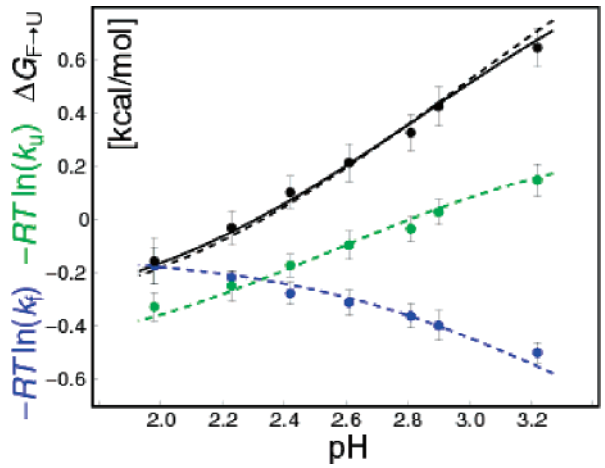

Figure 2. The pH stability profile of the wild-type drkN SH3 domain between pH 2.0 and 3.2, $\Delta G_{F\rightarrow U}$, calculated from folding and unfolding rate constants using $\Delta G_{F\rightarrow U} = -RT\ln(k_u/k_f)$ (black circles), 15 °C. The solid black line is the pH dependence of $\Delta G_{F\rightarrow U}$ calculated from experimental, residue-specific pK_a values for both states as described.[3] The black dashed line corresponds to the contribution of Asp8 *only* to the pH stability profile. Experimentally determined values of $-RT\ln k_f$ (blue circles) and $-RT\ln k_u$ (green circles) are shown along with profiles obtained from the best fit of the experimental data to eqs 1 and 2 to obtain p$K_a(\ddagger)$ (dashed blue, green lines).

where k_b and h are the Boltzmann and Planck constants, respectively, and κ is a transmission coefficient, related to the probability that the reaction proceeds to product from the TSE. The corresponding equation for $-RT\ln k_u$ is obtained by replacing U with F in eq 2. The free energy of unfolding, $\Delta G_{F\rightarrow U}$, along with free energy differences between the TSE and states U and F, $\Delta G_{U\rightarrow\ddagger}$, and $\Delta G_{F\rightarrow\ddagger}$, respectively, can be obtained directly from eq 1 by replacing A and B with the appropriate values for j above.

Values of $\Delta G_{F\rightarrow U}$ were calculated from folding and unfolding rate constants measured as a function of pH using the relation $\Delta G_{F\rightarrow U} = -RT\ln(k_u/k_f)$ (black circles in Figure 2). The black solid line shows the pH profile of $\Delta G_{F\rightarrow U}$ obtained by including contributions from all titrating groups in both F and U states from individual experimentally determined pK_a values as described,[3] with $\Delta G^0_{F\rightarrow U}$ fitted to minimize the difference between experiment (circles) and calculation (solid line). In contrast, the black dashed line corresponds to $\Delta G_{F\rightarrow U}$ calculated assuming that ionization of Asp8 *only* contributes to the pH stability profile over the pH range from 2.0 to 3.2. The fractional difference between free energy values computed by including all Asp, Glu, and His side chains or only Asp8 is less than 5.5% between pH 2.0−3.2, indicating that the pH dependence of stability can be well accounted for by including contributions from Asp8 only in this pH range. This allows a determination of the pK_a of Asp8 in the transition state ensemble, p$K_a(\ddagger)$, by fitting the pH profiles of $-RT\ln k_f$ (blue circles) and $-RT\ln k_u$ (green circles) simultaneously using previously determined values for the pK_a of Asp8 in states F (2.18 ± 0.11) and U (3.75 ± 0.03) and eqs 1 and 2. A value of 2.92 ± 0.09 is obtained for p$K_a(\ddagger)$.

The side chain acidic group of Asp8 makes a nonlocal electrostatic interaction in the folded state, a hydrogen bond with the backbone amide of Lys21, that is conserved in SH3 domains,[7] and we have previously shown that the backbone amide chemical shift of Lys21 changes with an apparent pK_a that coincides with that of the side chain of Asp8.[3] In addition, a local interaction between Asp8 and the imidazole moiety of His7 is also present that contributes considerably less to the pH stability profile.[3] To investigate how each interaction affects p$K_a(\ddagger)$, we have repeated the analysis described above on a His7Ala mutant of the drkN SH3 domain for which pK_a values in the F and U states have also been measured, 2.66 ± 0.06 and 3.98 ± 0.03, respectively. A value of p$K_a(\ddagger)$ = 3.25 ± 0.17 is obtained for the His7Ala mutant (see Supporting Information) that reflects the contribution from the Asp8−Lys21 interaction exclusively, assuming that the interactions in the TSE are those that are found in the folded protein.

For pH values much greater than those of pK_a(U), the electrostatic interactions involving Asp8 stabilize the TSE and F states by 1.1(1.0) and 2.1(1.7) kcal/mol relative to the U state for wild-type (His7Ala) drkN SH3 domain. An interpretation consistent with these energetics is one where the electrostatics involving Asp8 are present approximately to 55% in the TSE for both wild-type and mutant proteins. This assumes, of course, that the interactions are of similar magnitude in both the TSE and F states, and an alternative and perhaps more likely scenario is that electrostatic contacts of reduced magnitude are present more than 55% of the time in the TSE. This result complements ϕ_f value analysis that suggests that the region encompassing residue Asp8 is only weakly structured in the TSE, with $\phi_f = 0.18$ for the His7Ala drkN SH3 domain and $\phi_f < 0.2$ for α-spectrin and Src SH3 domain positions analogous to Asp8 of drkN SH3.[8] It is noteworthy that single mutant ϕ_f values represent averages over *all* energetic components involving a particular side chain, whereas here, we have obtained energetic information on a *specific* electrostatic interaction in the TSE.

In conclusion, our kinetic data on the drkN SH3 domain folding have allowed an estimate of the pK_a value in the TSE for an aspartic acid residue that is conserved in SH3 domains. The approach, which records kinetic data from magnetization transfer rates on a sample under equilibrium, nondenaturing conditions, can facilitate the study of other properties of the TSE, utilizing a variety of perturbations other than pH.

Acknowledgment. The research was supported by grants from the Canadian Institutes of Health Research (J.D.F.-K. and L.E.K.). M.T. is a recipient of an E. Schrödinger Return Fellowship of the Austrian Science Fund (R18). L.E.K. holds a Canada Research Chair in Biochemistry.

Supporting Information Available: Description of sample preparation, NMR experiments, and analysis of data. Figure of $-RT\ln k_f$, $-RT\ln k_u$, and $\Delta G_{F\rightarrow U}$ versus pH for the His7Ala mutant of the drkN SH3 domain. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Fersht, A. R. *Structure and Mechanism in Protein Science*; W. H. Freeman and Company: New York, 1998.
(2) (a) Jamin, M.; Geierstanger, B.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6127−6131. (b) Oliveberg, M.; Fersht, A. R. *Biochemistry* **1996**, *35*, 6795−6805. (c) Luisi, D. L.; Raleigh, D. P. *J. Mol. Biol.* **2000**, *299*, 1091−1100. (d) Sato, S.; Raleigh, D. P. *J. Mol. Biol.* **2002**, *318*, 571−582. (e) Horng, J.-C.; Cho, J.-H.; Raleigh, D. P. *J. Mol. Biol.* **2005**, *345*, 163−173.
(3) Tollinger, M.; Crowhurst, K. A.; Kay, L. E.; Forman-Kay, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4545−4550.
(4) (a) Mok, Y. K.; Elisseeva, E. L.; Davidson, A. R.; Forman-Kay, J. D. *J. Mol. Biol.* **2001**, *307*, 913−928. (b) Tollinger, M.; Skrynnikov, N. R.; Mulder, F. A. A.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 11341−11352.
(5) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. *J. Biomol. NMR* **1994**, *4*, 727−734.
(6) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *331*, 204−238.
(7) Larson, S. M.; Davidson, A. R. *Protein Sci.* **2000**, *9*, 2170−2180.
(8) (a) Martinez, J. C.; Serrano, L. *Nat. Struct. Biol.* **1999**, *6*, 1010−1016. (b) Riddle, D. S.; Grantcharova, V. P.; Santiage, J. V.; Alm, E.; Ruczinski, I.; Baker, D. *Nat. Struct. Biol.* **1999**, *6*, 1016−1024.

JA051942C