J|A|C|S COMMUNICATIONS

Published on Web 05/06/2010

A Method To Determine Residue-Specific Unfolded-State pK_a Values from Analysis of Stability Changes in Single Mutant Cycles

Jana K. Shen*

Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 92037

Received March 8, 2010; E-mail: jana.k.shen@ou.edu

Protein stability is defined by the equilibrium free energy difference between the folded and unfolded states. Thus, interactions in the unfolded state can affect protein stability and, if they persist in the transition state, also modulate folding kinetics (see a recent review[1]). Structural information about the unfolded state of a stable protein, however, is difficult to obtain because the unfolded state is rarely populated under folding conditions and it undergoes fast conformational exchange. Nevertheless, extrapolation of the experimental data obtained under denaturing conditions to native conditions has revealed the existence of native and non-native residual structures in the unfolded state (see recent reviews[1,2]). In the following discussion we focus on the unfolded state under native conditions.

As an alternative to direct structural determination, measurements of pH-dependent stabilities and native-state pK_a values offer evidence regarding nonrandom electrostatic interactions in the unfolded state. In this approach, the measured pH-dependent stability profile is compared with the one calculated using the Wyman−Tanford linkage equation,[3,4]

$$\partial \Delta G/\partial \mathrm{pH} = 2.303RT\Delta Q^{\mathrm{F}\rightarrow\mathrm{U}} = 2.303RT(Q^{\mathrm{U}} - Q^{\mathrm{F}}) \quad (1)$$

which relates the unfolding free energy (ΔG) to the difference of the net charge between the folded (F) and unfolded (U) states. If the unfolded state behaves in a "random-coil" manner, then charged side chains are fully exposed to solvent and not subject to net electrostatic effects. In this case, pK_a's for the unfolded state are the same as model compound pK_a's. Thus, a discrepancy between the measured and calculated pH-dependent stability profiles infers that some unfolded-state pK_a's are shifted from the model values, which offers proof for the presence of energetically significant electrostatic interactions in the unfolded state.

Based on this approach a number of proteins have been suggested to have pK_a shifts in the unfolded states.[5−11] To attribute the deviation between the measured and calculated pH-dependent stability profiles to residue-specific pK_a shifts, several *ad hoc* approaches have been attempted, for example, assuming a uniform pK_a shift[5] or using a structural model for representing the unfolded state,[11,12] or making an educated guess about a particular residue.[11] Recently, Raleigh and co-workers hypothesized that Asp8 has a depressed pK_a in the denatured state of the NTL9 protein, based on the observation that the deviation between the measured and calculated pH-dependent stability profiles is largely abolished when Asp8 is mutated to Asn.[10] However, the magnitude of the pK_a shift for Asp8 remains unknown. Here we derive an analytic expression based on the Wyman−Tanford theory[3,4] to allow determination of site-specific unfolded-state pK_a's from the native-state pK_a's and equilibrium stabilities of the wild type and charge-neutralizing mutants at two pH conditions. To illustrate the approach we determine the unfolded-state pK_a's for NTL9.

We begin by integrating the Wyman−Tanford linkage equation (eq 1) and applying the Henderson−Hasselbach equation to obtain the unfolding free energy as a function of pH.

$$\Delta\Delta G = \Delta G(\mathrm{pH}) - \Delta G(\mathrm{pH}^{\mathrm{ref}}) = RT\sum_i \ln \frac{(1 + 10^{(\mathrm{p}K_a^{\mathrm{F}}(i) - \mathrm{pH})})(1 + 10^{(\mathrm{p}K_a^{\mathrm{U}}(i) - \mathrm{pH}^{\mathrm{ref}})})}{(1 + 10^{(\mathrm{p}K_a^{\mathrm{U}}(i) - \mathrm{pH})})(1 + 10^{(\mathrm{p}K_a^{\mathrm{F}}(i) - \mathrm{pH}^{\mathrm{ref}})})} \quad (2)$$

Here pH$^{\mathrm{ref}}$ is a reference pH and the summation runs over all residues titrable in the range of pH to pH$^{\mathrm{ref}}$. p$K_a^{\mathrm{F}}(i)$ and p$K_a^{\mathrm{U}}(i)$ are the respective folded- and unfolded-state pK_a's. We neglect Hill coefficients in eq 2 because a small deviation from 1 has little effect on the resulting free energy change.[13] Equation 2 gives a means to break down the pH-dependent stability change into residue-based electrostatic contributions. Since pH$^{\mathrm{ref}}$ is arbitrary, we can choose it to be much higher than the folded and unfolded-state pK_a's. Thus, $1 + 10^{(\mathrm{p}K_a^{\mathrm{U}}(i) - \mathrm{pH}^{\mathrm{ref}})} \approx 1$ and $1 + 10^{(\mathrm{p}K_a^{\mathrm{F}}(i) - \mathrm{pH}^{\mathrm{ref}})} \approx 1$. Substituting them into eq 2 we obtain the electrostatic contribution from residue *i*.

$$\Delta\Delta G_i^{\mathrm{ele}} = RT \ln \frac{1 + 10^{(\mathrm{p}K_a^{\mathrm{N}}(i) - \mathrm{pH})}}{1 + 10^{(\mathrm{p}K_a^{\mathrm{D}}(i) - \mathrm{pH})}} \quad (3)$$

If the pH is much lower than the native- and unfolded-state pK_a's, e.g., residue *i* is fully protonated, the above equation can be further reduced to

$$\Delta\Delta G_{i,\mathrm{max}}^{\mathrm{ele}} = 2.303RT(\mathrm{p}K_a^{\mathrm{F}}(i) - \mathrm{p}K_a^{\mathrm{U}}(i)) \quad (4)$$

This is the maximum electrostatic contribution from residue *i*, e.g., the free energy change associated with protonation.

Equation 4 reveals that the unfolded-state pK_a can be calculated if $\Delta\Delta G_{i,\mathrm{max}}^{\mathrm{ele}}$ and the folded-state pK_a are known. However, the former is difficult to obtain because it is impossible to experimentally turn off/on the charge on a single side chain. An experiment that gives the closest result is the substitution of an ionizable residue by a charge-neutral one without introducing significant structural perturbation. For example, Asp→Asn and Glu→Gln are the commonly used mutations that serve this purpose. However, the stability change due to a charge-neutralizing mutation also contains a nonelectrostatic contribution.

For a moment, let us assume that the mutation does not affect pK_a's of other titrable residues in the native or unfolded state. In this case, eq 2 shows that the electrostatic contribution can be obtained as the difference between the pH-dependent stability changes of wild type (WT) and mutant (MT).

$$\Delta\Delta G^{\mathrm{WT}} - \Delta\Delta G^{\mathrm{MT}} \approx \Delta\Delta G_i^{\mathrm{ele}} \quad (5)$$

Based on this realization we propose to estimate the unfolded-state pK_a by performing stability measurements for WT and MT at two pH values. A thermodynamic cycle can be then constructed (Scheme 1),

$$\Delta\Delta\Delta G = \Delta\Delta G^{\mathrm{WT}} - \Delta\Delta G^{\mathrm{MT}} = \Delta\Delta G^{\mathrm{pH}_1} - \Delta\Delta G^{\mathrm{pH}_2} \quad (6)$$

 10.1021/ja101761m © 2010 American Chemical Society

Scheme 1. A Single-Mutant Cycle Analysis for Determination of Unfolded-State pK_a's

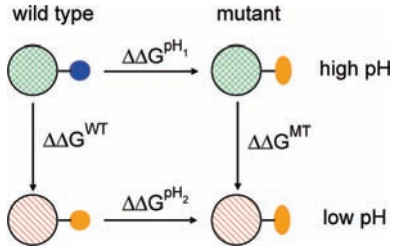

Table 1. Determination of Unfolded-State pK_a's of NTL9 Using Stability Data of WT and Mutants as well as the Native-State pK_a's [a]

	$\Delta\Delta G^{MT}$	$\Delta\Delta\Delta G$	pK_a^F	$pK_a^{U,0}$	pK_a^U	pK_a^{Frag}
D8N	−2.50	0.00	2.99	2.99	3.05	3.84
E17Q	−2.20	−0.30	3.57	3.83	3.87	4.11
D23N	−1.80	−0.70	3.05	3.77	3.87	4.11
E38Q	−2.00	−0.50	4.04	4.43	4.45	4.63
E48Q	−2.40	−0.10	4.21	4.29	4.35	4.31
E54Q	−2.50	−0.00	4.21	4.21	4.26	4.32

[a] $\Delta\Delta G^{MT}$ and $\Delta\Delta G^{WT}$ (−2.50 kcal/mol) are the measured stability changes in going from pH 6 to 3 (estimated from the urea denaturation data presented in ref 10). $\Delta\Delta\Delta G$ is defined in eq 6. pK_a^F denotes the measured native-state pK_a's.[10] $pK_a^{U,0}$ is calculated using eqs 3 and 5, while pK_a^U includes the correction term in eq 7. pK_a^{Frag} denotes the measured pK_a's using the fragment peptides derived from the sequence of the intact NTL9.[13] pK_a^{Frag} values deviate from the model values for Asp (4.0) and Glu (4.4) due to local electrostatic effects.[10]

which, combined with eq 3, suggests that the unfolded-state pK_a can be determined through either the pH-dependent (vertical) or the mutation (horizontal) arms.

Now we remove the assumption that mutation does not perturb other pK_a's in the native state while keeping the condition that mutation does not perturb unfolded-state pK_a's. The latter is a good approximation because the electrostatic coupling between titratable residues in the unfolded state is negligible. From eqs 2 and 3 it is evident that mutation-induced electrostatic perturbation in the native state can be accounted for by adding the following correction term into eq 5.

$$\Delta = \sum_{j\neq i} RT \ln \frac{1 + 10^{(pK_a^{F,WT}(j)-pH)}}{1 + 10^{(pK_a^{F,MT}(j)-pH)}} \qquad (7)$$

If mutational perturbations in pK_a's are small, eq 7 shows that the correction for the unfolded-state pK_a is also small. Equations 3, 5, and 7 form the basis of the proposed approach.

We now apply the method to determine the unfolded-state pK_a's of NTL9 for which pH-dependent stability and native-state pK_a data are available for WT and mutants of all acidic residues.[10] The results are summarized in Table 1. Remarkably, the pK_a of Asp8 in the unfolded state is shifted down to ∼3, identical to that in the native state. This result explains why mutation D8N largely abolishes the discrepancy between measured pH-dependent stabilities and calculation using model pK_a's.[10] We note that errors in the calculated pK_a^U values can be estimated from eq 4. Given a maximum error in stability measurements, 0.32 kcal/mol at pH 3, the maximum error for estimated pK_a^U values is 0.23 pH units. Interestingly, the unfolded-state pK_a's obtained here are in excellent agreement with the results from pH-coupled molecular dynamics simulations of the unfolded state of WT NTL9 (J.K.S., unpublished data). From Table 1 it can be seen that the corrections for the unfolded-state pK_a's due to mutational perturbation are small. However, this is not a general case. To further validate the unfolded-state pK_a's obtained here, we sum up the residue-based electrostatic contributions using eq 2. The resulting pH-dependent stability profile

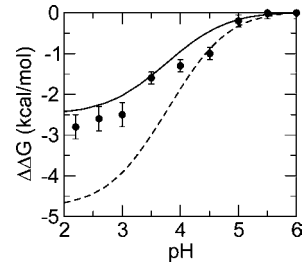

Figure 1. Comparison of calculated and measured pH-dependent stability profiles for NTL9. Experimental data (adapted from ref 10) are shown in solid circles with estimated error bars. Calculation based on the unfolded-state pK_a's obtained in this work is shown in solid curve. Calculation based on model compound pK_a's is shown in dashed curve.

is in good agreement with the measurements although the discrepancy from the calculation using model pK_a's is somewhat overestimated (Figure 1). The latter is most likely the result of uncertainties in the stability measurements and warrants further investigation.

In summary, an unfolded-state pK_a can be calculated from eqs 3, 5 and 7 using its native-state pK_a and the stabilities of WT and the corresponding charge-neutralizing mutant measured at two pH values. The high pH condition is chosen such that the residue of interest is fully unprotonated in both folded and unfolded states. The low pH condition is chosen such that the native state remains significantly populated and the interested residue is at least partially protonated. Presence of the native-state population is also a condition for the measurement of the native-state pK_a value.

Finally, it can be proven that the present approach is a generalization of the one used by Fersht and co-workers, in which both folded- and unfolded-state pK_a values are obtained from plotting the difference between $\Delta Q^{F\rightarrow U}$ of WT and that of mutant as a function of pH.[6] The major difference is that our approach accounts for the mutational perturbation of native-state electrostatics. Furthermore, our approach requires only four stability measurements instead of the entire pH profiles as required by the method of Fersht and co-workers. Knowledge of unfolded-state pK_a's allows quantitative estimation of the unfolded-state electrostatic effects on protein stability. It also offers valuable benchmarks for the improvement of protein force fields and validation of microscopic electrostatics from pH-coupled protein folding simulations.[14]

Acknowledgment. I thank Dr. Daniel P. Raleigh for critical reading of the manuscript and for pointing out the early work by Fersht and co-workers.[6] Financial support from the ACS Petroleum Research Fund and University of Oklahoma is greatly appreciated.

References

(1) Cho, J.-H.; Sato, S.; Horng, J.-C.; Anil, B.; Raleigh, D. P. *Arch. Biochem. Biophys.* **2008**, *469*, 20–28.
(2) Bartlett, A. I.; Radford, S. E. *Nat. Struct. Mol. Biol.* **2009**, *16*, 582–588.
(3) Wyman, J., Jr. *Adv. Protein Chem.* **1964**, *19*, 223–286.
(4) Tanford, C. *Adv. Protein Chem.* **1970**, *25*, 1–93.
(5) Oliveberg, M.; Arcus, V. L.; Fersht, A. R. *Biochemistry* **1995**, *34*, 9424–9433.
(6) Tan, Y.-J.; Oliveberg, M.; Davis, B.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254*, 980–992.
(7) Whitten, S. T.; García-Moreno E., B. *Biochemistry* **2000**, *39*, 14292–14304.
(8) Pace, C. N.; Alston, R. W.; Shaw, K. L. *Protein Sci.* **2000**, *9*, 1395–1398.
(9) Marti, D. N.; Bosshard, H. R. *Biochemsitry* **2004**, *43*, 12436–12447.
(10) Cho, J.-H.; Raleigh, D. P. *J. Mol. Biol.* **2005**, *353*, 174–185.
(11) Lindman, S.; Linse, S.; Mulder, F. A. A.; André, I. *Biophys. J.* **2007**, *92*, 257–266.
(12) Fitzkee, N. C.; García-Moreno E., B. *Protein Sci.* **2008**, *17*, 216–227.
(13) Kuhlman, B.; Luisi, D. L.; Young, P.; Raleigh, D. P. *Biochemistry* **1999**, *38*, 4896–4903.
(14) Khandogin, J.; Chen, J.; Brooks, C. L., III. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18546–18550.

JA101761M