J|A|C|S COMMUNICATIONS

Published on Web 12/22/2009

ϕ-Value Analysis for Ultrafast Folding Proteins by NMR Relaxation Dispersion

Jae-Hyun Cho,[†] Nichole O'Connell,[†] Daniel P. Raleigh,*[,‡] and Arthur G. Palmer III*[,†]

Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, New York 10032, and Department of Chemistry and Graduate Program in Biochemistry and Structural Biology, State University of New York at Stony Brook, Stony Brook, New York 11794

Received October 23, 2009; E-mail: draleigh@notes.cc.sunysb.edu; agp6@columbia.edu

Proteins that fold rapidly, on the (sub-) microsecond time scale, offer the exciting prospect of direct comparison between experimental data and molecular dynamics simulations.[1−6] The standard method for assessing the role of amino acid side chains in the transition state for folding is a protein engineering approach commonly referred to as ϕ-value analysis (Figure S1, Supporting Information, SI).[7−9] Application of ϕ-value analysis to ultrafast folding proteins is stymied by several technical difficulties: (*i*) folding rates are too fast for conventional stopped-flow methods, (*ii*) suitable spectroscopic probes often are not available in natural amino acid sequences,[10] and (*iii*) mutational effects on the denatured state ensemble obscure the interpretation of ϕ-values.[11] The third concern is particularly critical because residual structure in the denatured state may be important for ultrafast folding.[1,12] In this Communication, we demonstrate that these difficulties are obviated by the use of NMR spin relaxation dispersion methods to determine ϕ-values and probe effects of mutations on the denatured (D), or intermediate (I), state ensembles of ultrafast folding proteins. The described approach provides information that is difficult, if not impossible, to obtain by other means.

NMR spin relaxation dispersion measurements have been used to perform ϕ-value analysis for proteins in which at least some spins in the molecule fall outside of the fast-exchange limit (on the chemical shift time scale).[13,14] In this case, a complete analysis of relaxation dispersion data yields the populations of the native (N) and D (or I) states, p_N and p_D, respectively; the folding and unfolding kinetic rate constants, k_f and k_u, respectively; and the chemical shift differences between N and D (or I) states ($\Delta\omega = \Omega_N - \Omega_D$).[15] In addition, comparison of $\Delta\omega$ for wild-type and mutant proteins provides a powerful probe of potential mutational effects on the D (or I) state. Kay and co-workers have pioneered this approach with their studies of SH3 domains.[13,16]

Ultrafast folding proteins exhibit chemical exchange line broadening in the fast-exchange limit ($k_{ex} = k_f + k_u \geq \Delta\omega \approx 10^3-10^4$ s^{-1}). For fast-limit two-site chemical exchange, the transverse relaxation rate constant is $R_2 = R_2^0 + R_{ex}$, in which R_2^0 is the population-average relaxation rate constant for N and D (or I) states in the absence of chemical exchange processes, $R_{ex} = \phi_{ex}/k_{ex}$ and $\phi_{ex} = p_N p_D \Delta\omega^2$. Note that ϕ_{ex} is not the ϕ-value. In this regime, NMR spectroscopy would appear to have limited application for ϕ-value analysis, because the product $p_N p_D$ cannot be determined independently of $\Delta\omega^2$ and only the magnitude of $\Delta\omega$ can be obtained.[15] Nonetheless, NMR-based ϕ-value analysis is even more facile for protein folding in the fast-exchange limit.

The effects of mutation on chemical exchange line broadening are given by

$$k_B T \ln(k_{ex}^{mt}/k_{ex}^{wt}) = \Delta\Delta G^{\dagger} + k_B T \ln(p_N^{mt}/p_N^{wt}) \approx \Delta\Delta G^{\dagger} - k_B T(p_D^{wt} - p_D^{mt}) \approx \Delta\Delta G^{\dagger} \quad (1)$$

$$k_B T \ln(\phi_{ex}^{mt}/\phi_{ex}^{wt}) = -\Delta\Delta G^{\circ} + 2k_B T \ln(p_N^{mt}/p_N^{wt}) \approx -\Delta\Delta G^{\circ} - 2k_B T(p_D^{wt} - p_D^{mt}) \approx -\Delta\Delta G^{\circ} \quad (2)$$

in which k_B is the Boltzmann constant, and $\Delta\Delta G^{\circ}$ and $\Delta\Delta G^{\dagger}$ are the changes in stability and activation free energies, respectively (defined in Figure S1 in Supporting Information, SI). Equation 2 is obtained assuming that the values of $\Delta\omega$ are not affected by mutation. Using eqs 1, and 2,

$$\ln(k_{ex}^{mt}/k_{ex}^{wt})/\ln(\phi_{ex}^{mt}/\phi_{ex}^{wt}) = -(\phi - \zeta)/(1 - 2\zeta) \approx -\phi(1 + 2\zeta) + \zeta \approx -\phi \quad (3)$$

and $\xi = (k_B T/\Delta\Delta G^{\circ}) \ln(p_N^{mt}/p_N^{wt}) \approx (k_B T/\Delta\Delta G^{\circ})(p_D^{wt}/p_D^{mt})$. As shown in the SI, $0 \leq \xi \leq \max(p_D^{wt}, p_D^{mt})$; thus, for $0 \leq \phi \leq 1$, the maximum absolute error in ϕ is equal to $\max(p_D^{wt}, p_D^{mt})$ and this error vanishes when $\phi = 0.5$. Proteins for which relaxation dispersion measurements have been reported frequently have $p_D^{wt} < 0.05$ ($\Delta G^{\circ} \geq 7$ kJ/mol). By extension,

$$k_B T \ln(R_{ex}^{mt}/R_{ex}^{wt}) = -\Delta\Delta G^{\dagger} - \Delta\Delta G^{\circ} + 3k_B T \ln(p_N^{mt}/p_N^{wt}) \approx -\Delta\Delta G^{\dagger} - \Delta\Delta G^{\circ} + 3k_B T(p_D^{wt} - p_D^{mt}) \approx -\Delta\Delta G^{\circ} - \Delta\Delta G^{\dagger} \quad (4)$$

Equations 1−4 are the principal results for effects of mutations on fast-limit chemical exchange broadening. R_{ex} can be determined using Hahn spin−echo, Carr−Purcell−Meiboom−Gill (CPMG), or $R_{1\rho}$ relaxation dispersion experiments; k_{ex} and ϕ_{ex} can be determined using CPMG or $R_{1\rho}$ experiments; and the assumption that chemical exchange is in the fast-exchange limit can be confirmed from the static magnetic field dependence of R_{ex}.[15]

Nuclear spins in the wild-type (mutant) protein that are affected by the same kinetic process exhibit the same exchange rate constant k_{ex}^{wt} (k_{ex}^{mt}). Nuclear spins that further satisfy the critical assumption that $\Delta\omega$ is unaffected by mutation can be identified because a graph of ϕ_{ex}^{mt} vs ϕ_{ex}^{wt} (or R_{ex}^{mt} vs R_{ex}^{wt}) will follow a straight line through the origin. Nuclear spins whose environment in the D (or I) state is affected by mutation can be identified because population-average chemical shifts in the native state ($\sim\Omega_N$ for $p_N^{wt} \approx 1$) are observed directly in NMR spectra.

The proposed method is demonstrated for the villin headpiece domain HP67 (Figure S2). Relaxation data for backbone ^{15}N spins for wild-type and H41Y mutant have been reported previously.[3,17] Significant line broadening is observed predominantly for the ^{15}N spins of amino acid residues in the N-terminal subdomain of HP67 owing to equilibrium (un)foldng between N and I states.[3] A plot of R_{ex}^{mt} vs R_{ex}^{wt}, determined from Hahn-echo experiments, exhibits

[†] Columbia University.
[‡] State University of New York at Stony Brook.

10.1021/ja909052h

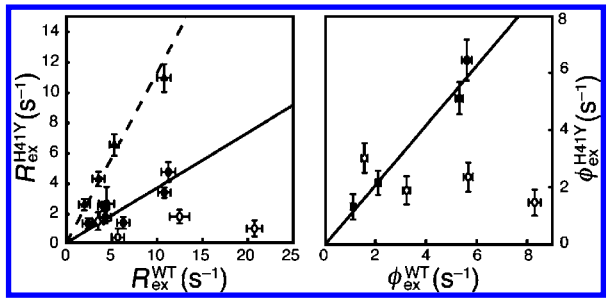

Figure 1. Comparison of (a) R_{ex}^{H41Y} vs R_{ex}^{wt} and (b) ϕ_{ex}^{H41Y} vs ϕ_{ex}^{wt}. (a) The residues in Group A (D19, L21, A25, E27, D28, G32, and E39) and Group B (L18, V20, V22, and T24) are shown in closed circles and triangles, respectively. The fitted line for each group is shown in solid (Group A) and dashed lines (Group B), respectively. Open circles (T15, F16, D34, and L42) represent the residues for which $\Delta\omega$ is changed significantly upon mutation and are not included in the analysis. (b) Four residues (D19, L21, E27, and D28) from Group A are shown as closed circles with a fitted line. Open circles (T15, F16, E39, and L42) represent the residues for which $\Delta\omega$ is changed upon mutation (see SI).

two groups of residues, denoted A and B, based on significantly different slopes of 0.379 ± 0.028 and 1.117 ± 0.087, respectively, yielding $\Delta\Delta G^\circ - \Delta\Delta G^\dagger = -2.36 \pm 0.18$ and 0.27 ± 0.19 kJ/mol, respectively, from eq 4 (Figure 1a and SI). In principle, the difference between the slopes for Group A and B could represent a difference in either $\Delta\Delta G^\circ$ or $\Delta\Delta G^\dagger$, or both. However, the mutational effect on the stability is global; therefore, the difference in slopes must arise from multiple transition states (see SI). This conclusion is confirmed by $R_{1\rho}$ relaxation dispersion experiments described below. Data for residues T15, F16, and D34 deviate from the fitted lines and must have large differences in $\Delta\omega$ between wild-type and mutant HP67. Comparison of NMR spectra indicates that the mutation affects Ω_N for T15, D34, and L42.[3] In contrast, F16 shows little change in Ω_N, indicating that Ω_D is altered by mutation.

The ϕ-value can be obtained from eq 4 if the stability difference ($\Delta\Delta G^\circ$) is known independently. If conventional methods (CD or fluorescence) are not applicable or more rigorous analysis is preferred, NMR relaxation dispersion experiments enable the ϕ-value to be calculated from eqs 1 and 4 or eqs 1−3. The thermodynamic stability of the N-terminal subdomain of HP67 is difficult to measure because of lack of a unique spectroscopic probe to monitor protein folding independently of the C-terminal subdomain.[17] Thus, $R_{1\rho}$ relaxation dispersion measurements were performed at two static magnetic fields and fit globally for residues in Groups A and B for both wild-type and mutant HP67.[3] Dispersion curves for Group A were described by a rate constant, k_{ex}, of 5700 $\pm$ 100 and 14 000 $\pm$ 1000 s^{-1} in the wild-type and mutant proteins, respectively. The same analysis for Group B yielded $k_{ex} = (4.2 \pm 0.5) \times 10^4$ and $(4.6 \pm 0.8) \times 10^4$ s^{-1} in the wild-type and mutant, respectively. Thus, the kinetic process of Group B is clearly distinct from that of Group A, as indicated by Figure 1. Using eqs 1 and 4 yields $\Delta\Delta G^\dagger = 2.19 \pm 0.18$ and 0.22 ± 0.51 kJ/mol and $\Delta\Delta G^\circ = 0.17 \pm 0.25$ and 0.49 ± 0.51 kJ/mol for Groups A and B, respectively. Thus, the kinetic process contributing to exchange broadening of the spins in group B is not affected by the mutation. More detailed analysis suggests that this process arises from repacking of the hydrophobic core of HP67.[3,18]

A graph of ϕ_{ex}^{mt} vs ϕ_{ex}^{wt} for Group A is shown in Figure 1b. The solid line through the origin is fitted to data for residues D19, L21, E27, and D28 and has a slope of 1.05 ± 0.08. Using eq 2 gives $\Delta\Delta G^\circ = 0.12 \pm 0.18$ kJ/mol, consistent with the value obtained from eqs 1 and 4. In this case, $|\phi| \gg 1$ for residues in Group A. Noncanonical ϕ-values indicate that non-native interactions are formed in the transition state and/or that energetics of the D (or I) state are affected by mutation. Analysis of data recorded at pH 6 yields $\phi \approx 1$ for the H41Y mutation (Figure S3). The large change in ϕ-value may reflect differences in the free energy of the intermediate state at pH 7 and 6 (see SI). More detailed analysis suggests that the chemical exchange process affecting spins in Group A arises from a kinetic transition from the N to I state, in which the N-terminal subdomain of HP67 is (largely) unfolded.[18] The analysis of H41Y HP67 demonstrates the importance of multisite detection for accurate ϕ-analysis. Thus, NMR-based methods can significantly reduce the effort of ϕ-analysis by obviating the need to repeatedly alter the locations of spectroscopic probes.[19] The same formalism can be applied to investigations of changes in solution conditions, such as pH and isotope effects, as illustrated for HP67 by Figures S3 and S4.

In conclusion, we have shown that NMR spin relaxation data for ultrafast folding proteins with chemical exchange rates in the fast-exchange limit can be used to (*i*) assess mutational effects on the D (or I) state, (*ii*) identify nuclear spins in chemical sites suitable for accurate ϕ-value analysis, and (*iii*) derive ϕ-values. The chemical exchange line broadening rate constant, R_{ex}, can be obtained from efficient Hahn-echo experiments, without recourse to full CPMG or $R_{1\rho}$ relaxation dispersion methods, and provides a particularly straightforward approach, provided that the change in stability of the mutant protein can be determined independently. A wide range of ultrafast folding protein domains give well-resolved NMR spectra; thus, the approach is expected to be broadly applicable. The method was demonstrated using the H41Y mutant of HP67, which is a particularly challenging test case because the substitution of H41 disrupts residual non-native interactions in an intermediate state.

Acknowledgment. This work was supported by NIH Grant GM59273 (A.G.P.) and NSF Grant MCB0919860 (D.P.R.).

Supporting Information Available: Theory, data analysis, schematic representation of the thermodynamic states and ϕ-values, and structure of HP67 with the position of the analyzed residues. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Mayor, U.; Guydosh, N. R.; Johnson, C. M.; Grossmann, J. G.; Sato, S.; Jas, G. S.; Freund, S. M. V.; Alonso, D. O. V.; Daggett, V.; Fersht, A. R. *Nature* **2003**, *421*, 863–867.
(2) Vugmeyster, L.; Kroenke, C. D.; Picart, F.; Palmer, A. G.; Raleigh, D. P. *J. Am. Chem. Soc.* **2000**, *122*, 5387–5388.
(3) Grey, M. J.; Tang, Y.; Alexov, E.; McKnight, C. J.; Raleigh, D. P.; Palmer, A. G. *J. Mol. Biol.* **2006**, *355*, 1078–1094.
(4) Kubelka, J.; Eaton, W. A.; Hofrichter, J. *J. Mol. Biol.* **2003**, *329*, 625–630.
(5) Meisner, W. K.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15639–15644.
(6) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102–106.
(7) Fersht, A. R.; Matouschek, A.; Serrano, L. *J. Mol. Biol.* **1992**, *224*, 771–782.
(8) Goldenberg, D. P.; Frieden, R. W.; Haack, J. A.; Morrison, T. B. *Nature* **1989**, *338*, 127–132.
(9) Matthews, C. R. *Methods Enzymol.* **1987**, *154*, 498–511.
(10) Ferguson, N.; Schartau, P. J.; Sharpe, T. D.; Sato, S.; Fersht, A. R. *J. Mol. Biol.* **2004**, *344*, 295–301.
(11) Cho, J. H.; Raleigh, D. P. *J. Am. Chem. Soc.* **2006**, *128*, 16492–16493.
(12) Ferguson, N.; Sharpe, T. D.; Schartau, P. J.; Sato, S.; Allen, M. D.; Johnson, C. M.; Rutherford, T. J.; Fersht, A. R. *J. Mol. Biol.* **2005**, *353*, 427–446.
(13) Neudecker, P.; Lundstroem, P.; Kay, L. E. *Biophys. J.* **2009**, *96*, 2045–2054.
(14) Neudecker, P.; Zarrine-Afsar, A.; Choy, W.-Y.; Muhandiram, D. R.; Davidson, A. R.; Kay, L. E. *J. Mol. Biol.* **2006**, *363*, 958–976.
(15) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204.
(16) Korzhnev, D. M.; Kay, L. E. *Acc. Chem. Res.* **2008**, *41*, 442–451.
(17) Tang, Y.; Grey, M. J.; McKnight, C. J.; Palmer, A. G.; Raleigh, D. P. *J. Mol. Biol.* **2006**, *355*, 1066–1077.
(18) O'Connell, N. E.; Grey, M. J.; Tang, Y.; Kosuri, P.; Miloushev, V. Z.; Raleigh, D. P.; Palmer, A. G. *J. Biomol. NMR* **2009**, *45*, 85–98.
(19) Sato, S.; Religa, T. L.; Fersht, A. R. *J. Mol. Biol.* **2006**, *360*, 850–864.

JA909052H