Communication

Dynamic NMR Line-Shape Analysis Demonstrates that the Villin Headpiece Subdomain Folds on the Microsecond Time Scale

Minghui Wang, Yuefeng Tang, Satoshi Sato, Liliya Vugmeyster, C. James McKnight, and Daniel P. Raleigh

J. Am. Chem. Soc., **2003**, 125 (20), 6032-6033• DOI: 10.1021/ja028752b • Publication Date (Web): 30 April 2003

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information

J|A|C|S
COMMUNICATIONS

Published on Web 04/30/2003

Dynamic NMR Line-Shape Analysis Demonstrates that the Villin Headpiece Subdomain Folds on the Microsecond Time Scale

Minghui Wang,[†] Yuefeng Tang,[†] Satoshi Sato,[†,||] Liliya Vugmeyster,[†,⊥] C. James McKnight,[*,‡] and Daniel P. Raleigh[*,†,§]

Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, Department of Physiology and Biophysics, Boston University School of Medicine, Boston, Massachusetts 02118, and Graduate Programs in Biophysics and in Biochemistry and Structural Biology, State University of New York at Stony Brook, Stony Brook, New York 11794

Received September 30, 2002; E-mail: draleigh@notes.cc.sunysb.edu; cjmck@bu.edu

The mechanism by which an initially unfolded polypeptide chain folds into its final native structure is still not fully understood despite the obvious importance of the protein folding problem. A recent popular approach is to simulate the folding of small proteins either by all atom molecular dynamics calculations or by models which make use of a simplified representation of the peptide chain.[1,2] The small, independently folded, 36-residue villin headpiece subdomain (HP36) has been the subject of at least six computational studies, but there are no reported experimental determinations of its folding rate.[2,3] Recently, distributed computing methods have been applied to the protein folding problem, and again HP36 has been an important test system.[3] In this Communication, we report the use of dynamic NMR line-shape analysis to measure the folding rate of the villin headpiece subdomain. The protein folds on the time scale of 10 μs. These measurements provide an important benchmark to compare with molecular dynamics simulations and with theoretical efforts. The results should also prove useful in refinements of diffusion-collision models of folding.

The chicken villin headpiece subdomain (HP36) is the C-terminus of the actin-bundling protein villin.[4] The numbering system used here corresponds to that used in other publications. The first residue is a Met. The second residue is designated Leu42, and the C-terminus is Phe76. Our construct has an amidated C-terminus. This does not affect the structure, although it slightly decreases the melting temperature. HP36 is made up of three helices (residues Asp44-Lys48, Arg55-Phe58, and Leu63-Lys70) which are packed together to form the hydrophobic core (Figure 1).[5] The structure of the subdomain is identical to the same region in the intact headpiece.[5,6] The subdomain folds cooperatively and reversibly. The T_m of our construct is 67 °C at pD 5.0. The free energy of unfolding, $\Delta G°$, is 2.44 kcal mol^{-1} at 25 °C.

Dynamic NMR line-shape analysis is an attractive method for measuring rapid folding rates, and HP36 is an excellent candidate for such investigations.[7] The one-dimensional proton NMR spectrum exhibits several well-resolved resonances. The resonance due to the C4 proton of Phe47 is strongly upfield shifted to 5.51 ppm because of its close proximity to the phenyl rings of Phe51 and Phe58. A resonance from Phe51 is also well resolved at 6.32 ppm. One of the Val50 methyl resonances is very well resolved and is found at −0.11 ppm in the native state. Examination of the NMR spectra recorded on a 500 MHz instrument as a function of temperature shows that the peaks shift and broaden throughout the unfolding transition region (Figure 2), indicating that the native and denatured state are in intermediate to fast exchange.

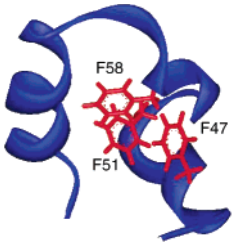

Figure 1. Ribbon diagram of the chicken villin headpiece subdomain, HP36, created with WebLab ViewerLite (MSI). The three Phe residues which contribute to the hydrophobic core are shown.

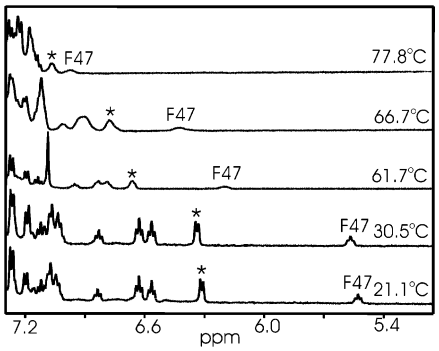

Figure 2. Temperature-dependent ^{1}H 500 MHz NMR spectra of HP36. The aromatic region with the C4 ^{1}H resonance of F47 is labeled. The resonance due to Phe51 is labeled with a star. Spectra were recorded in D_2O at pD 5.0 (uncorrected meter reading).

Analysis of the 500 MHz line shape of the Phe47 C4 proton allows the exchange rate, $k_{ex} = k_f + k_u$, to be determined. Knowledge of the relative population of the native and unfolded states in turn allows the folding rate, k_f, and unfolding rate, k_u, to be determined. The line shape could be fit using the relative populations as a variable; however, this increases the number of adjustable parameters and decreases the precision of the measurement (see Supporting Information). Instead, we independently determined the population of the two states by conducting CD monitored thermal unfolding experiments under identical conditions. The 500 MHz line shape of Phe47 at 66.7 °C is shown in Figure 3. The simulations indicate a folding rate of 0.37×10^5 s^{-1}. To investigate the uncertainty in k_f, we systematically varied k_{ex} and calculated the residuals between the theoretical and experimental line shape. Changes of 30−40% in k_{ex} are required to generate clear effects (Figure 3 and Supporting Information).

[†] Department of Chemistry, State University of New York at Stony Brook.
[‡] Boston University School of Medicine.
[§] Graduate Programs in Biophysics and in Biochemistry and Structural Biology, State University of New York at Stony Brook.
[||] Present address: Centre for Protein Engineering MRC, Hills Road, Cambridge, CB2 2QH, U.K.
[⊥] Present address: Institut de chimie moléculaire et biologique, Ecole Polytechnique Fédérale de Lausanne, BCH 1015 Lausanne, Switzerland.

10.1021/ja028752b CCC: $25.00

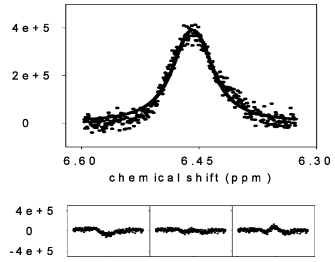

Figure 3. Representative example of the line-shape analysis. Top: Comparison between experimental line shape (●) and calculated line shape (−) for the Phe47 resonance at 500 MHz and 66.7 °C. Bottom: Plots of the residuals of calculated versus experimental line shape. The center panel corresponds to the simulation on the top. The left-hand panel corresponds to a k_{ex} 30% less than the best fit, and the right-hand panel corresponds to a k_{ex} 30% greater than the best fit.

Table 1. Kinetic Parameters for the Folding of HP36 Derived from Variable Temperature Line-Shape Analysis at 500 and 700 MHz[a]

(A) 500 MHz

T (°C)	p_D	k_{ex} (×10^5 s^{-1}) F47	k_f (×10^5 s^{-1}) F47	k_u (×10^5 s^{-1}) F47
56.3	0.28	1.79	1.28	0.51
61.7	0.40	1.75	1.05	0.70
66.7	0.53	0.79	0.37	0.42
72.6	0.68	1.26	0.40	0.86
77.8	0.80	2.59	0.52	2.07

(B) 700 MHz

		k_{ex} (×10^5 s^{-1})			k_f (×10^5 s^{-1})			k_u (×10^5 s^{-1})		
T (°C)	p_D	F47	V50	F51	F47	V50	F51	F47	V50	F51
57.2	0.30	3.09		1.13	2.16		0.79	0.93		0.34
62.3	0.41	2.53	1.62	1.65	1.49	0.96	0.97	1.04	0.66	0.68
68.1	0.57	2.39	1.88	1.61	1.03	0.81	0.69	1.36	1.07	0.92
73.7	0.72	3.64	1.47	1.53	1.02	0.41	0.43	2.62	1.06	1.10
76.8	0.79	2.78	1.41	1.19	0.58	0.30	0.25	2.20	1.11	0.94

[a] p_D denotes population denatured.

It would be desirable to independently estimate the folding rate using other resonances. This provides an addition test of the precision of k_f and also helps to test if folding is really two-state. The large exchange rate and smaller chemical shift differences mean that none of the other resonances are well suited for line-shape analysis at 500 MHz. The Val50 methyl resonance can be analyzed at 700 MHz. The line shape is sensitive to variations in k_{ex} of 30−40%. The folding rates determined from this resonance at 700 MHz are within a factor of 2 of those estimated from the Phe47 resonance at either field. The Phe51 resonance (indicated by a star in Figure 2) can also be analyzed at 700 MHz. The calculated exchange rates vary from 1.65×10^5 s^{-1} at 62.3 °C to 1.19×10^5 s^{-1} at 76.8 °C and are all within 20% of the values determined from the Val50 resonance.

Folding rates could be estimated over the range of 56−78 °C (Table 1). Outside of this range, the two states are not sufficiently populated, or the resonances of interest overlap with other peaks. Many studies of protein folding are conducted near 25 °C, and it would be useful to estimate the folding rate near 25 °C. The folding rate at 25 °C was obtained by fitting spectra collected in the transition region of a guanidine denaturation experiment. At 1.93 M guanidine, the folding rate is 1.5×10^4 s^{-1}. There is a large uncertainty in the rate determination from extrapolation to 0 M guanidine, but it is on the order of 1×10^4 s^{-1}. At 33 °C, the estimated folding rate is 2.2×10^4 s^{-1}.

The measured folding rate of HP36 is consistent with recent molecular dynamics simulations. The folding rate is much less sensitive to temperature than the unfolding rate: k_f varies by 5−10× over the range of 33−77.8 °C, while k_u varies by more than 200×. This could arise if entropic effects dominate the barrier to folding, or it could reflect the shallow temperature dependence of $\Delta G_{folding}^{\ddagger}$ expected for a system where $\Delta C_{p,folding}^{\ddagger}$ is small.[1g] The small size of HP36 naturally leads to small values of ΔC_p° and $\Delta C_{p,folding}^{\ddagger}$.

Acknowledgment. We thank Francis Picart and Martine Ziliox for their assistance with some of the NMR measurements. This research was supported by NIH grant GM 54233 to D.P.R., and GM 62886 to C.J.M.

Supporting Information Available: Table and figure of simulations; Arrhenius plot for HP36 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Mayor, U.; Johnson, C. M.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13518−13522. (b) Jager, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373−393. (c) Myers, J. K.; Oas, T. G. *Nat. Struct. Biol.* **2001**, *8*, 552−558. (d) Daggett, V. *Acc. Chem. Res.* **2002**, *35*, 422−429. (e) Zhou, Y.; Karplus, M. *Nature* **1999**, *401*, 400−403. (f) Burton, R. E.; Myers, J. K.; Oas, T. G. *Biochemistry* **1998**, *37*, 5337−5343. (g) Chan, H. S.; Dill, K. A. *Proteins* **1998**, *30*, 2−33. (h) Klimov, D. K.; Thirumalai, D. *J. Chem. Phys.* **1998**, *109*, 4119−4125. (i) Gutin, A.; Sali, A.; Abkevich, V.; Karplus, M.; Shakhnovich, E. I. *J. Chem. Phys.* **1998**, *108*, 6466−6483. (j) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 12952−12953. (k) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258−11259.

(2) (a) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740−744. (b) Duan, Y.; Wang, L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9897−9902. (c) Sullivan, D. C.; Kuntz, I. D. *J. Phys. Chem. B* **2002**, *106*, 3255−3262. (d) Shen, M. Y.; Freed, K. F. *Proteins* **2002**, *49*, 439−445. (e) Islam, S. A.; Karplus, M.; Weaver, D. L. *J. Mol. Biol.* **2002**, *318*, 199−215. (f) Srinivas, G.; Bagchi, B. *Curr. Sci.* **2002**, *82*, 179−185. (g) Fernandez, A.; Shen, M. Y.; Colubri, A.; Sosnick, T. R.; Berry, R. S.; Freed, K. F. *Biochemistry* **2003**, *42*, 664−671.

(3) (a) Zagrovic, B.; Snow, C. D.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 927−37. (b) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153−164.

(4) McKnight, C. J.; Doering, D. S.; Matsudaira, P. T.; Kim, P. S. *J. Mol. Biol.* **1996**, *260*, 126−134.

(5) McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. *Nat. Struct. Biol.* **1997**, *4*, 180−184.

(6) Vardar, D.; Buckley, D. A.; Frank, B. S.; McKnight, C. J. *J. Mol. Biol.* **1999**, *294*, 1299−1310.

(7) (a) Huang, G. S.; Oas, T. G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6878−6882. (b) Kuhlman, B.; Boice, J. A.; Fairman, R.; Raleigh, D. P. *Biochemistry* **1998**, *37*, 1025−1032. (c) Burton, R. E.; Huang, G. S.; Daugherty, M. A.; Calderone, T. L.; Oas, T. G. *Nat. Struct. Biol.* **1997**, *4*, 305−310. (d) Gillespie, B.; Plaxco, K. *Protein Sci.* **2002**, *11*, 208A.

JA028752B