NMR Spectroscopy

DOI: 10.1002/anie.201301851

Protein Folding Studied by Dissolution Dynamic Nuclear Polarization**

*Hsueh-Ying Chen, Mukundan Ragavan, and Christian Hilty**

Protein folding occurs on a timescale that is not directly observable with most traditional nuclear magnetic resonance (NMR) spectroscopy methods.[1,2] The insights into this complex process that can potentially be gained from the high site resolution of NMR has led to developments such as stopped-flow NMR,[3] incorporation of specific isotope labels,[4] and pulse sequences tailored for rapid data acquisition.[5] Hyperpolarization, the generation of a non-equilibrium spin state, shows significant promise to enhance the sensitivity and, by removing the need for signal averaging, dramatically decrease signal acquisition time.[6,7] For the study of protein folding, chemically induced dynamic nuclear polarization (CIDNP) has been used to hyperpolarize tryptophan residues that undergo a cyclic reaction with a photosensitizer.[8] A different technique, dynamic nuclear polarization (DNP),[9] hyperpolarizes nuclei throughout a molecule via an admixed stable free radical. Combined with solid state NMR, DNP provides unique information on protein structure.[10] Using dissolution DNP,[11] NMR in the liquid state would be sensitive to structural changes across the entire macromolecule during the folding process. Liquid-state NMR signals of a DNP hyperpolarized, denatured protein, the ribosomal protein L23,[12,13] have recently been observed.[14]

Here, the re-folding of this hyperpolarized polypeptide is monitored following a pH jump (Figure 1). A frozen, hyperpolarized sample is brought to ambient temperature by dissolution with denaturing buffer, and injected into the NMR spectrometer, where it mixes with a buffer of higher pH. Folding is then monitored over a period of seconds by acquisition of a series of NMR spectra using small flip-angle excitation (Figure 2a). The signal intensity in these spectra gradually decreases because of spin relaxation and the effect of excitation pulses. Relative signal changes, which can occur from folding as well as non-uniform spin relaxation, become visible in Figure 2b. Only the differences to the first and to the last scan in the time series are shown, after rescaling of each spectrum to equal intensity integrated over carbonyl and aliphatic regions. Relative intensity increases are observed in the chemical-shift region corresponding to methyl groups in folded protein (near 12 ppm), as well as in off-center regions of the carbonyl chemical-shift range (170–185 ppm). The latter "shoulders", which are characteristic of the folded protein, are readily visible in re-normalized traces of the corresponding region (arrows in Figure 2c). Specifically, intensity from α-helix and β-sheet secondary structure is broadly centered around 177.9 ppm and 174.5 ppm, respectively.[15]

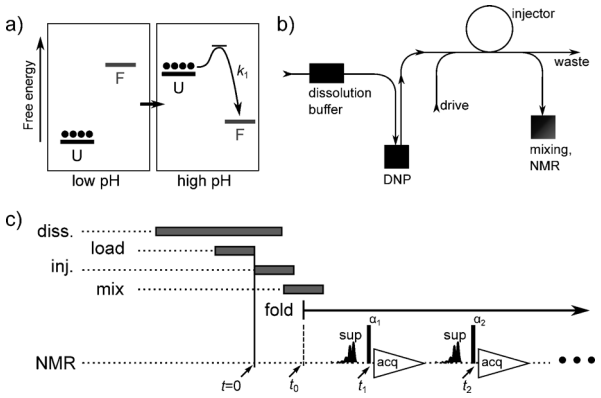

Figure 1. a) Representation of a pH jump experiment for re-folding of a protein (two-state folder shown). b) Experimental setup for DNP NMR. The sample is hyperpolarized ("DNP"), dissolved under denaturing conditions, and injected into an NMR instrument, where mixing with a refolding buffer occurs. c) Experimental timeline, indicating dissolution ("diss."), loading into the injector, injection into the NMR instrument, and mixing with refolding buffer. Acquisition of ^{13}C spectra ("acq") is preceded by solvent suppression ("sup") and excitation with small flip angle radio frequency pulses $\alpha_1 ... \alpha_n$.

Quantitatively, the kinetic rate constant of folding can be extracted from analysis of the signal intensities in the DNP NMR spectra. As L23 has been reported to follow a two-state folding process,[13] the populations of the folded and unfolded species at each time point are obtained by decomposing the individual scans into a linear combination of reference spectra of folded and acid-denatured L23 (Figure 3a and Supporting Information, Figure S7). The reconstructed spectra are in good agreement with the DNP NMR spectra, except for deviations near 180 ppm and 183 ppm (the latter outside of the fitted region), which likely pertain to side-chain carbonyl chemical shifts. Side-chain carbonyl chemical shifts exhibit a pH dependence in function of protonation state, and furthermore relative intensity differ-

[*] H.-Y. Chen,[+] Prof. Dr. C. Hilty
Department of Chemistry
Texas A&M University, College Station, TX 77843 (USA)
E-mail: chilty@chem.tamu.edu
M. Ragavan[+]
Department of Biochemistry and Biophysics
Texas A&M University, College Station, TX 77843 (USA)
[+] These authors contributed equally to this work.

[**] Financial support from the National Science Foundation (Grants CHE-0846402 and CHE-0840464), the Welch Foundation (Grant A-1658), and Texas A&M University is acknowledged. We thank Prof. Dr. Mikael Oliveberg, Stockholm University for providing the plasmid containing the gene for L23, and Prof. Dr. François Gabbaï for use of the fluorimeter.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201301851.

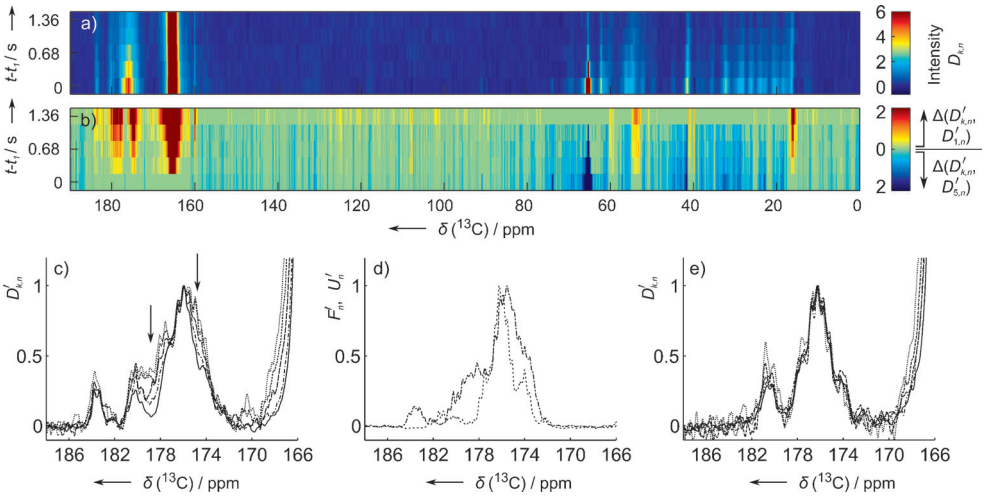

Figure 2. a) Intensity plot of a time series of 5 spectra from a single hyperpolarized sample of [u-^{13}C, 50%-^{2}H]-L23, after pH jump to initiate protein folding (40 μM L23, 100 mM ethylene glycol, 45 mM urea, 10% (v/v) methanol (MeOH), 5 mM KH_2PO_4 and 45 mM 4-morpholineethanesulfonic acid (MES) buffer, pH 5.8, $T = 301$ K, $B = 9.4$ T); t_1 is the time of the first scan. b) Composite plot showing in cyan–blue the difference compared to the last spectrum, $\Delta(D'_{k,n}, D'_{5,n}) = \max(0, D'_{k,n} - D'_{5,n})$, and in yellow–red the difference compared to the first spectrum, $\Delta(D'_{k,n}, D'_{1,n}) = \max(0, D'_{k,n} - D'_{1,n})$ in the time series. Each spectrum $D'_{k,n}$ is obtained by rescaling $D_{k,n}$ to the same total integration value in the range 185.7–170 ppm and 63–7 ppm prior to taking the difference. c) Time evolution of carbonyl resonances (increasingly shorter dashes for later times; spectra are re-scaled to unit maximum intensity near 176 ppm). d) Reference spectra measured without hyperpolarization of unfolded L23 (-----; 40 μM L23, 100 mM ethylene glycol, 45 mM urea, 10% (v/v) MeOH, 5 mM KH_2PO_4, pH 3.1, 32000 scans) and folded L23 (–•–•; using sample from (a), 32000 scans), cryoprobe at $B = 11.7$ T. e) Series of spectra as in (c), but without the pH jump.

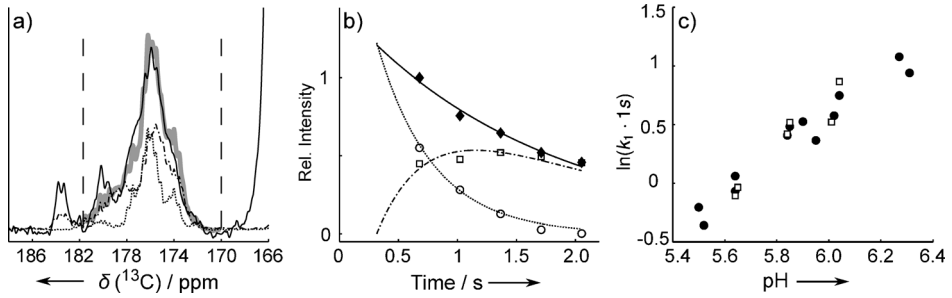

Figure 3. a) Decomposition of a representative spectrum (black ——; second scan of 5-scan time series) into a linear combination of unfolded (-----) and folded (–•–•) reference spectra. Reference spectra are drawn to scale according to their fractional contribution. The spectrum reconstituted by linear combination is shown as a thick gray line. Vertical dashed lines indicate fitting region. b) Time course of signal intentities. Shown are total signal (◆), and fractions of signal from folded (□) and unfolded (○) forms obtained from linear decompositions. Fit results are $r = 0.59\ s^{-1}$, $k_1 = 1.52\ s^{-1}$, and $t_0 = 0.32$ s. c) pH dependence of folding rate constants for L23 from DNP NMR measurements (□) and from fluorescence measurements (●) for validation.

ences can arise because the spin-lattice relaxation time (T_1) is longer in side chains than in the backbone. Figure 3b shows the time evolution of the signal fractions obtained from the decompositions. The first data point already contains a contribution from folded protein, as folding starts during the dead time of the experiment. Nevertheless, from a fit with kinetic model equations, folding and spin relaxation rates can be determined. The folding rate constants from several measurements with different final pH are shown in Figure 3c and Supporting Information, Figure S8b, and are in excellent agreement with stopped flow fluorescence measurements. The model used herein is based on a two-state folding mechanism, and assumes equal T_1 for folded and unfolded forms (see the Experimental Section). The latter is valid because under present conditions, with $r_F/r_U = 0.76$, the expected error in the final result is smaller than the observed spread of values from multiple measurements (Supporting Information, Figure S9). In the more general case, a model with two known relaxation time constants could also be used. Observed signal intensities are in principle convolved with an input function due to sample mixing. Calculations using model equations for injection with time constants of 100–200 ms[6,16] however show insignificant deviations in the determined rate constant (Supporting Information, Figure S10 and Table S1). To test the influence of side-chain peak intensity on the decomposition, a reference spectrum of urea denatured protein at pH 5.8 was used (Supporting Information, Figures S8b and S11). Results were indistinguishable within experimental error, suggesting that the effect of side chain carbonyl chemical-shift variations is not significant.

In the absence of known reference spectra, the observed signals can alternatively be decomposed into contributions from α-helix, β-sheet, and random-coil structure. With this procedure, an increase in secondary structure content was observed as expected (Supporting Information, Figure S12), but compared to Figure 3 the random coil contribution is overestimated. This may be due to the specific chemical shifts in L23, which are better captured by the actual reference spectra, as well as due to overlap of the α-helix, β-sheet, and random-coil chemical-shift ranges.

Along with global rate constants, site-specific information can be obtained if individual chemical shifts are resolved.

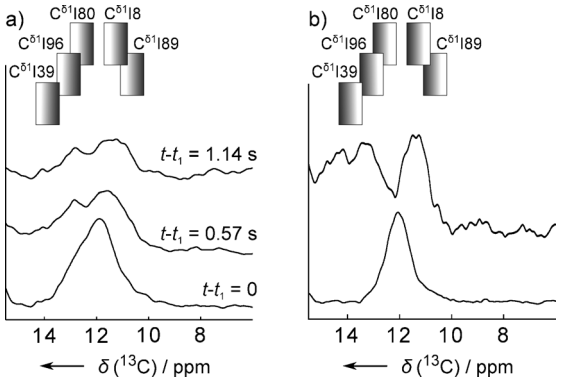

Figure 4. a) DNP NMR spectra showing the time evolution of Ile $C^{\delta1}$ chemical shifts after a pH jump (40 μM L23, 60 mM KH_2PO_4, pH 5.7, 10% (v/v) MeOH, 400 mM urea, 100 mM ethylene glycol, $T = 301$ K). b) Non-hyperpolarized reference spectra of L23 unfolded (bottom; 0.4 mM L23, 10% (v/v) MeOH/50 mM KH_2PO_4, pH 3.1) and folded (top; sample as in (a), 50000 scans on cryoprobe at 11.7 T). In (a) and (b), chemical shifts of folded L23 are indicated with ranges covering isotopomers CD_3, CD_2H, CDH_2, and CH_3.[21] Dark shading indicates the position of highest relative detection sensitivity; relaxation losses in the DNP experiment are lowest for CD_3, and equilibrium polarization in the non-hyperpolarized experiment is largest for CH_3 (not considering relative isotopomer concentration and line width differences).

Here, Ile $C^{\delta1}$ groups are sufficiently separated for this purpose (Figure 4). These methyl groups are often located in the hydrophobic core of a protein (in L23, Ile 8, 39, and 89 are part of the hydrophobic core) and experience a significant change in chemical environment upon folding. Spectra taken at three time points (Figure 4) reveal a large initial contribution from random-coil chemical shifts, followed at later time by resolution of individual peaks for the different methyl groups in the folded protein.

In proteins with a more complicated folding pathway, this experiment could be useful for the identification of on-pathway intermediate states, provided that their lifetime is at least on the order of 100 ms. Peaks of such intermediates can additionally be assigned by forward or backward correlation to the known chemical shifts in the folded or unfolded protein.[17] Additional site-specific information can potentially be recovered using techniques aimed at reducing signal overlap, such as selective isotope enrichment[18] or multidimensional spectroscopic techniques.[19]

The length of the time interval accessible to the DNP NMR method is governed by T_1 relaxation. Using deuteration, ^{13}C relaxation rates even in large polypeptides can be on the order of seconds.[20] In this time frame, the DNP NMR measurement can provide a substantial amount of information on tertiary structure formation. Apart from the observation of global folding kinetics, DNP NMR includes the capability to simultaneously resolve different sites on the macromolecule. These features may enable the observation of differential kinetics in proteins that do not follow a two-state folding mechanism, of contacts established during assembly of protein complexes, or even yield structural information on folding intermediates.

Experimental Section

Ribosomal protein L23 (*T. thermophilus*) was produced as reported.[12,14] Sample aliquots consisted of 10 μL of nominally 4 mM protein in 60% (v/v) ethylene glycol/water, 6 M urea, 15 mM tris-{8-carboxyl-2,2,6,6-tetrakis [2-(1-hydroxethyl)]-benzo(1,2-d:4,5-dS)bis-(1,3)dithiole-4-yl} methyl free radical sodium salt (Oxford Instruments, Abingdon, U.K.), and 1 mM diethylenetriamine pentaacetic acid gadolinium complex (Sigma–Aldrich, St. Louis, MO). Samples were hyperpolarized on ^{13}C for 4–6 h in a HyperSense DNP polarizer (Oxford Instruments) at $T = 1.4$ K, using microwaves of 60 mW power at 93.974 GHz. Samples were then dissolved in 10% (v/v) MeOH/5 mM KH_2PO_4 in water, pH 3.1, heated to a vapor pressure of 10 bar. Dissolved samples were injected into a 400 MHz NMR spectrometer (Bruker Biospin, Billerica, MA).[22] In the NMR spectrometer, samples were mixed at a ratio of 17:1 (v/v) with 25 μL of 810 mM MES or KH_2PO_4 buffer, pH 5.6–6.2 (KH_2PO_4 does not show ^{13}C NMR peaks). NMR spectra were acquired at $T = 301$ K in intervals of 340 ms using solvent suppression, small flip-angle excitation,[23] and (^1H, ^2H) decoupling. Reference spectra of folded and unfolded L23 were measured using signal averaging on a 500 MHz NMR spectrometer with cryoprobe (Bruker). Data was processed (20 Hz line broadening, polynomial baseline correction) with TopSpin 3.0 (Bruker) and analysis was performed using Matlab (MathWorks, Natick, MA). For decomposition, reference spectra and first scan of the DNP NMR experiments were normalized to the integral of signal intensities between 170 and 181.7 ppm. Unfolded ($\check{v}_k$) and folded ($\hat{\phi}_k$) signal fractions were determined by minimizing the root mean square difference between linear combinations of the reference spectra and individual scans from the DNP NMR experiment. The apparent relaxation rate (r) was found by fitting a single exponential to total signal intensity, then folding rate constant k_1 was determined by fitting $\phi(t) = ce^{-r(t-t_0)}\left(1 - e^{-k_1(t-t_0)}\right)$ and $v(t) = ce^{-(k_1+r)(t-t_0)}$. Fit parameters were the intensity scaling factor, c, and mixing start time, t_0.

Received: March 5, 2013
Revised: May 7, 2013
Published online: July 16, 2013

Keywords: hyperpolarization · NMR spectroscopy · protein folding

[1] K. W. Plaxco, C. M. Dobson, *Curr. Opin. Struct. Biol.* **1996**, *6*, 630–636.
[2] J. Balbach, C. Steegborn, T. Schindler, F. X. Schmid, *J. Mol. Biol.* **1999**, *285*, 829–842.
[3] C. Frieden, S. D. Hoeltzli, I. J. Ropson, *Protein Sci.* **1993**, *2*, 2007–2014.
[4] S. D. Hoeltzli, C. Frieden, *Biochemistry* **1996**, *35*, 16843–16851.
[5] P. Schanda, V. Forge, B. Brutscher, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11257–11262.
[6] S. Bowen, C. Hilty, *Angew. Chem.* **2008**, *120*, 5313–5315; *Angew. Chem. Int. Ed.* **2008**, *47*, 5235–5237.
[7] H. Zeng, Y. Lee, C. Hilty, *Anal. Chem.* **2010**, *82*, 8897–8902.
[8] C. M. Dobson, P. J. Hore, *Nat. Struct. Biol.* **1998**, *5*, 504–507.
[9] A. Abragam, M. Goldman, *Rep. Prog. Phys.* **1978**, *41*, 395–467.
[10] M. Rosay, J. C. Lansing, K. C. Haddad, W. W. Bachovchin, J. Herzfeld, R. J. Temkin, R. G. Griffin, *J. Am. Chem. Soc.* **2003**, *125*, 13626–13627.
[11] J. H. Ardenkjær-Larsen, B. Fridlund, A. Gram, G. Hansson, L. Hansson, M. H. Lerche, R. Servin, M. Thaning, K. Golman, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10158–10163.
[12] A. Öhman, A. Rak, M. Dontsova, M. B. Garber, T. Härd, *J. Biomol. NMR* **2003**, *26*, 131–137.

[13] K. L. Maxwell, D. Wildes, A. Zarrine-Afsar, M. A. De Los Rios, A. G. Brown, C. T. Friel, L. Hedberg, J.-C. Horng, D. Bona, E. J. Miller, A. Vallée-Blisle, E. R. Main, F. Bemporad, L. Qiu, K. Teilum, N.-D. Vu, A. M. Edwards, I. Ruczinski, F. M. Poulsen, B. B. Kragelund, S. W. Michnick, F. Chiti, Y. Bai, S. J. Hagen, L. Serrano, M. Oliveberg, D. P. Raleigh, P. Wittung-Stafshede, S. E. Radford, S. E. Jackson, T. R. Sosnick, S. Marqusee, A. R. Davidson, K. W. Plaxco, *Protein Sci.* **2005**, *14*, 602 – 616.
[14] M. Ragavan, H.-Y. Chen, G. Sekar, C. Hilty, *Anal. Chem.* **2011**, *83*, 6054 – 6059.
[15] H. Zhang, S. Neal, D. S. Wishart, *J. Biomol. NMR* **2003**, *25*, 173 – 195.
[16] H. Y. Chen, Y. Lee, S. Bowen, C. Hilty, *J. Magn. Reson.* **2011**, *208*, 204 – 209.
[17] S. Bowen, C. Hilty, *Anal. Chem.* **2009**, *81*, 4543 – 4547.
[18] V. Tugarinov, V. Kanelis, L. E. Kay, *Nat. Protoc.* **2006**, *1*, 749 – 754.
[19] L. Frydman, D. Blazina, *Nat. Phys.* **2007**, *3*, 415 – 419.
[20] A. Eletsky, O. Moreira, H. Kovacs, K. Pervushin, *J. Biomol. NMR* **2003**, *26*, 167 – 179.
[21] K. H. Gardner, L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 357 – 406.
[22] S. Bowen, C. Hilty, *Phys. Chem. Chem. Phys.* **2010**, *12*, 5766 – 5770.
[23] H. Zeng, S. Bowen, C. Hilty, *J. Magn. Reson.* **2009**, *199*, 159 – 165.

