JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Real-Time NMR Characterization of Structure and Dynamics in a Transiently Populated Protein Folding Intermediate

Enrico Rennella,[†] Thomas Cutuil,[†] Paul Schanda,[†] Isabel Ayala,[†] Vincent Forge,[‡] and Bernhard Brutscher*[,†]

[†]Institut de Biologie Structurale, Université Grenoble 1, CEA, CNRS, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

[‡]DSV-iRTSV, Laboratoire de Chimie et Biologie des Métaux, Université Grenoble 1, CEA, CNRS, CEA-Grenoble, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France

Supporting Information

ABSTRACT: Recent advances in NMR spectroscopy and the availability of high magnetic field strengths now offer the possibility to record real-time 3D NMR spectra of short-lived protein states, e.g., states that become transiently populated during protein folding. Here we present a strategy for obtaining sequential NMR assignments as well as atom-resolved information on structural and dynamic features within a folding intermediate of the amyloidogenic protein β2-microglobulin that has a half-lifetime of only 20 min.

In the past, structural biology has mainly focused on studies of ground-state, low-energy conformations of biomolecules, providing a static picture of the molecular actors in biological processes. However, it has long been recognized that proteins and nucleic acids often do not fold into a single low-energy conformation but also sample higher energy structures that are separated from the ground state by sizable free-energy barriers. Such excited states can play important functional roles, or may be responsible for protein misfolding and fibril formation.[1] Hence there is a clear need for experimental techniques allowing a characterization of these excited-state conformations at atomic resolution. While X-ray crystallography has provided tremendous insight into structural details of protein ground states, NMR spectroscopy is the technique of choice to characterize the protein folding energy landscape in terms of conformational dynamics and structural properties of excited states. Low-populated excited protein states that are separated from the ground state by a kinetic barrier leading to interconversion on the millisecond time scale can be accessed by relaxation−dispersion NMR measurements.[2] Protein states that are separated from the ground state by even higher energy barriers are best studied by real-time NMR methods[3] where the refolding of the protein from an initial state, e.g., a highly unstructured conformation, to the native state is monitored by recording NMR spectra during the refolding process. The dead time of such a kinetic experiment can be limited to less than 1 s by an injection device[4] that allows fast mixing of the protein solution with a refolding buffer inside the NMR magnet.

Multidimensional (nD) NMR techniques provide site-specific information on local structure and dynamics in macromolecules. Unfortunately, using conventional nD NMR methods, the rate of data acquisition decreases from ~1 s^{-1} for 1D to ~10^{-2} s^{-1} for 2D and ~10^{-4} s^{-1} for 3D. Long acquisition times are the main limitation for the study of protein folding reactions by real-time 2D or 3D NMR. A first solution to this problem was introduced in 1996 by the Dobson group which proposed encoding kinetic information into the line shape of the correlation peaks observed in a single 2D spectrum acquired during protein folding.[5] While this yields reliable information on the folding kinetics, it does not provide any further information on the folding intermediates. An alternative, more powerful strategy is to reduce the NMR data acquisition time either by limiting the number of experimental repetitions (sparse sampling techniques) or by shortening the interscan delay needed for spin polarization recovery (fast pulsing techniques), or a combination of both.[6] Especially attractive in the context of real-time NMR are the fast-pulsing techniques BEST-HSQC (BH), BEST-TROSY (BT), and SOFAST-HMQC developed in our laboratory.[7,8] These techniques allow for significantly reduced acquisition times and thus increased dimensionality of the NMR experiments that can be recorded during protein folding. At the same time, they enhance the overall experimental sensitivity that is also a crucial issue for NMR studies of protein folding intermediates. Very recently, Balbach and co-workers[9] achieved NMR assignments of a folding intermediate of RNase T1 from a BH-HNCA spectrum recorded in 5 h during protein refolding. In this Communication, we present a general strategy for recording pure 3D correlation spectra of protein folding intermediates, or more generally short-lived transient states of macromolecules, with half-lifetimes of <1 h. The method is demonstrated for the amyloidogenic protein β2-microglobulin (B2M), the light chain of the human class I major histocompatibility complex. We show that sensitivity-optimized fast-pulsing real-time NMR experiments[8] performed at high magnetic field strength provide site-specific information on local structure and dynamics of a B2M folding intermediate.

B2M forms amyloid fibrils in the joints and connective tissues of patients undergoing long-term dialysis (dialysis-related amyloidosis).[10] B2M has been studied in the past by a variety of biochemical and biophysical methods,[11−13] highlighting the presence of a long-lived folding intermediate (I-state) characterized by a non-native trans-peptide bond

Received: March 16, 2012
Published: May 3, 2012

dx.doi.org/10.1021/ja302598j | J. Am. Chem. Soc. 2012, 134, 8066−8069

between residues His[31] and Pro[32].[11] This I-state is thought to be involved in the onset of amyloid fibril formation, although the exact relationship between this folding intermediate and fibrillogenesis has not yet been established. Here we have used the B2M mutant W60G as a model system to structurally and dynamically characterize the I-state at atomic resolution. Although the refolding kinetics of W60G-B2M are slightly faster than for the wild-type species, the W60G mutant was preferred because of a higher I-state population after the refolding burst phase.[12] At 15 °C, the I-state of W60G-B2M has a half-lifetime of ∼20 min (Figure 1a). While this makes

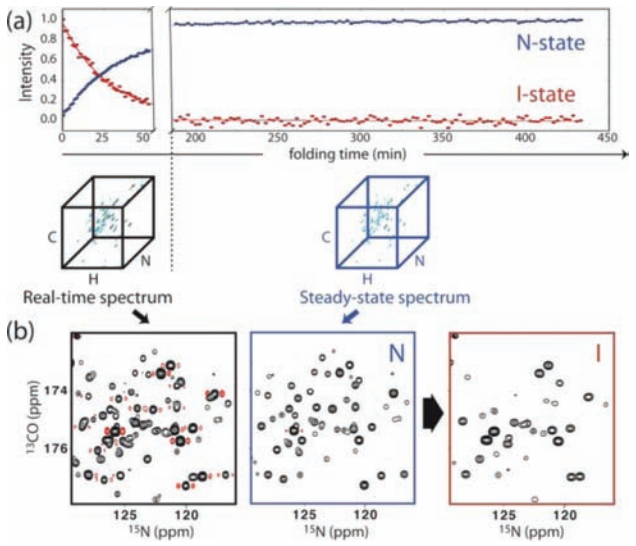

Figure 1. Strategy used to reconstruct pure I-state 3D spectra. (a) Kinetic profiles of the N- and I-states are shown during the folding of W60G-B2M, along with the 3D cubes of the acquired HNCO spectra. The red and blue dots are intensities of resolved I- and N-state peaks measured in a series of 2D SOFAST-HMQC spectra recorded during W60G-B2M folding. (b) Projections of the real-time (left), steady-state (center), and reconstructed (right) 3D HNCO spectra of B2M. The experimental time for the real-time 3D spectra recorded at 800 MHz for a 0.8 mM ^{13}C,^{15}N-labeled W60G-B2M sample was 40 min.

W60G-B2M one of the slowly folding single-domain proteins, this refolding time is still 6 times shorter than what has been reported for the RNaseT1 system of Balbach and co-workers.[9] Therefore, an NMR investigation of this W60G-B2M folding intermediate presents a serious challenge because each *n*D experiment has to be recorded within the short lifetime of the I-state.

The strategy used for the reconstruction of pure I-state NMR spectra is illustrated in Figure 1. A first 3D spectrum is recorded during the refolding process (*real-time spectrum*). The experimental time of this 3D data acquisition has to match the lifetime of the I-state; in other words, the acquisition should stop once the intensity of the I-state signals reaches the noise level. To remove the contribution of the N-state resonances to the *real-time spectrum*, a second 3D spectrum is acquired under steady-state conditions (*steady-state spectrum*), i.e., immediately after the end of the folding process (Figure 1a). This data set is typically acquired over a much longer time period in order to increase the signal-to-noise ratio (SNR). During data processing, an apodization function, corresponding to the folding kinetics of the N-state, is applied to the steady-state, pure N-state spectrum. Finally, a pure I-state spectrum, devoid of any N-state contributions, is obtained as the difference between the real-time and the apodized steady-state spectra. Kinetics-induced line broadening in the I-state spectra can be limited by choosing short scan times and long maximal evolution times in the last incremented dimension (see Supporting Information for details).

The strategy outlined above was first applied to HNCO, the most sensitive 3D triple-resonance experiment. At high magnetic field strengths (≥800 MHz), BEST-TROSY generally outperforms BEST-HSQC in terms of sensitivity and spectral resolution.[8] The ^{15}N−^{13}CO projections of the real-time, N-state, and reconstructed I-state 3D BT-HNCO spectra recorded at 800 MHz ^{1}H frequency are shown in Figure 1b. The quality of these H−N−CO data was sufficient to detect a similar number of correlation peaks (71) as observed in the I-state ^{1}H−^{15}N SOFAST-HMQC spectrum (Figure S1b). The next step was to record a BT-HNCA spectrum required for sequential NMR assignment. This experiment was performed at a magnetic field strength of 1000 MHz in order to further boost the experimental sensitivity. An average SNR increase of ∼50% was observed for BT-HNCA when comparing steady-state data recorded at 800 and 1000 MHz (Figure S2). The H−N−CA data (Figure 2) allowed unambiguous sequential

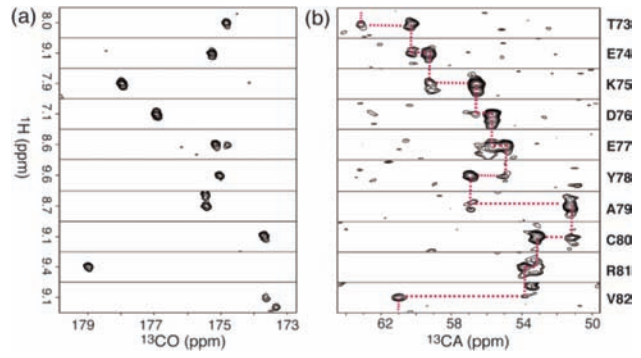

Figure 2. Example strip plots extracted from the reconstructed I-state spectra of (a) 3D BT-HNCO and (b) 3D BT-HNCA recorded on a 0.8 mM ^{13}C,^{15}N-labeled W60G-B2M sample at 800 MHz (HNCO) and 1 GHz (HNCA). The assignment walk for residues Thr[73]-Val[82] is highlighted by dotted (red) lines.

assignment of 63 H−N moieties out of the 95 non-proline B2M residues. This corresponds to the majority of correlation peaks detected in the ^{1}H−^{15}N correlation spectrum (Figure S1). Interestingly, the peptide regions encompassing Met[1]-Lys[6], Val[27]-Glu[36], Ser[53]-Lys[58], and Phe[62]-Leu[65] could not be assigned, because they did not give rise to detectable NMR signals in the 3D spectra, most likely due to extensive line broadening (see below).

The chemical shift dispersion observed in the ^{1}H−^{15}N correlation spectrum of the I-state (Figure S1b) provides a first indication that the overall structure of this transient state is "native-like". This is further supported by the ^{13}CO and ^{13}CA secondary chemical shift (SCS) data, calculated as the difference between the measured chemical shifts and tabulated random coil values corrected for next neighbor effects.[14] These SCS data are sensitive reporters of local structure (backbone dihedral angles). In particular, a series of consecutive negative SCS values are indicative for β-strand structures. Comparison of measured values for the N- and I-states (Figure 3a) shows that the location and population of β-strands in B2M is very

dx.doi.org/10.1021/ja302598j | J. Am. Chem. Soc. 2012, 134, 8066−8069

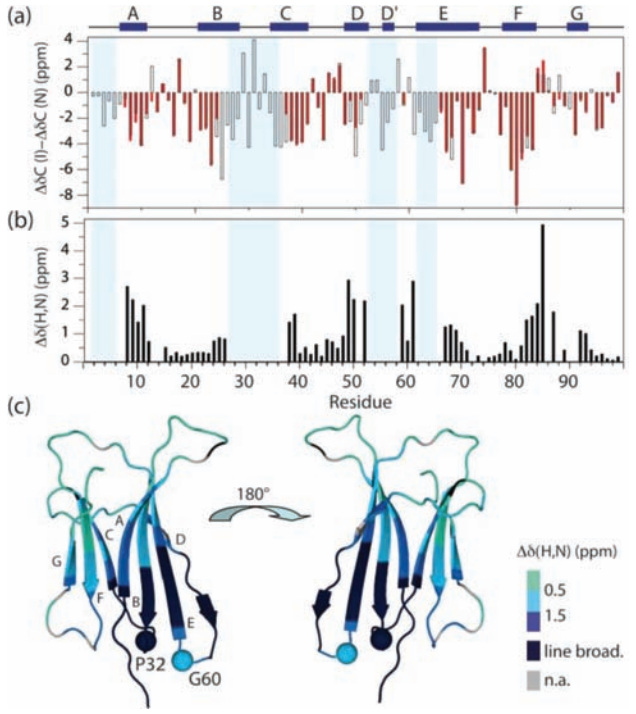

Figure 3. (a) Secondary ^{13}C chemical shifts ($\Delta\delta CO + \Delta\delta CA$) measured for the N-state (open bars) and I-state (filled red bars) of W60G-B2M. The ^{1}H, ^{15}N chemical shift difference $\Delta\delta HN = [(10\Delta\delta H)^2 + \Delta\delta N^2]^{0.5}$ between I and N is plotted in (b) as a function of sequence and in (c) on the cartoon structure of B2M. NMR-invisible I-state regions are indicated by cyan bars in (a) and (b), and dark blue bars in (c).

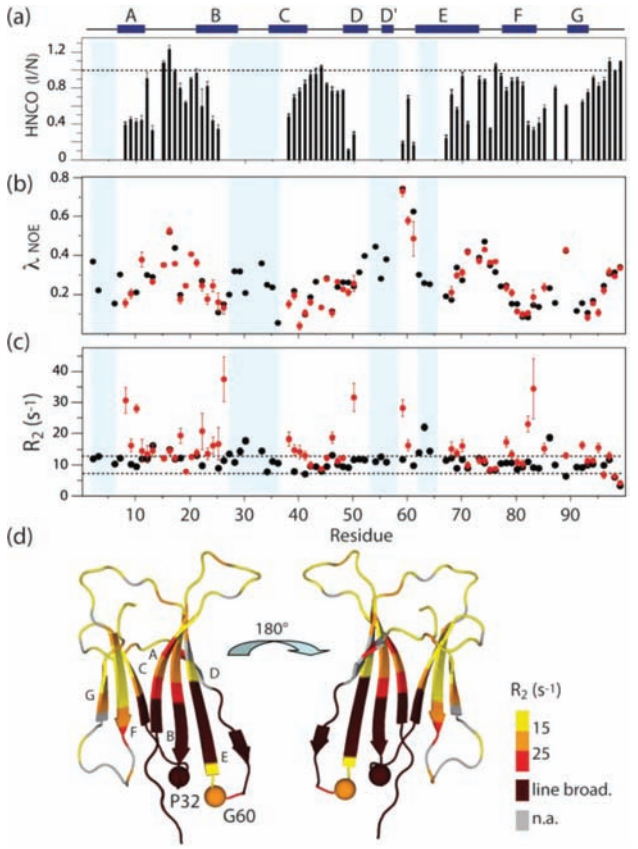

Figure 4. (a) Ratios of peak intensities measured in I-state and N-state 3D HNCO spectra. (b) λ_{NOE} values obtained from HET-SOFAST experiment. (c) $R_2(N_x + 2N_xH_z)$ relaxation rates measured with R_2-BEST-TROSY sequence (Figure S4). N- and I-state values are plotted in black and red, respectively. Both experiments were performed at 800 MHz on a 0.6 mM ^{13}C,^{15}N-labeled W60G-B2M sample. In (d) the measured R_2 values are color-coded on the cartoon structure of B2M. NMR-invisible I-state regions are indicated by cyan bars in (a)−(c), and dark red in (d).

similar in these two states. The only remarkable difference is observed for strand D, where the SCS data indicate a less pronounced population of β-strand conformation. In addition to local secondary structural changes, ^{1}H and ^{15}N chemical shifts are also sensitive to hydrogen bonding and tertiary contacts, e.g., to aromatic side chains that induce ring-current shifts. Significant changes ($\Delta\delta(H,N) > 1.5$ ppm) in ^{1}H, ^{15}N chemical shifts between the N- and I-states are found for residues Gln[8], Val[9], Ser[11], Leu[39], Val[49], Glu[50], Asp[59], Ser[61], Asn[83]-Val[85], and Leu[87] (Figure 3b), i.e., generally neighboring unassigned (i.e., undetected) regions of the protein in the I-state. A plot of the observed chemical shift deviations and the NMR-invisible protein regions on the 3D cartoon structure of B2M (Figure 3c) points out that structural differences between the I- and N-states exist in the apical region close to Pro[32], with a particularly strong effect in the second half of the β-sandwich (A, B, D, E), while strands F and G are less affected.

The missing NMR signals for several consecutive regions that are structurally close to Pro[32] indicate the presence of conformational dynamics in this part of the protein. Toward these regions, the NMR signal decreases gradually, as nicely seen from the plot of intensity ratios of I- versus N-state peaks in HNCO spectra (Figure 4a). To obtain additional, more quantitative information on the local conformational dynamics in the I-state of B2M, we performed two simple spin relaxation measurements during protein refolding. Such measurements require the recording of a small number of 2D $^{1}H-^{15}N$ correlation spectra (pseudo-3D experiment) using different parameter sets, e.g., varying a relaxation delay. The total acquisition time for such a pseudo-3D experiment is generally only a few minutes, much shorter than the protein folding time. Therefore, data acquisition is repeated several times during real-time folding (see Figure S5 for details).

First, a HET-SOFAST experiment[15] was performed that exploits the difference in amide ^{1}H T_1 relaxation in the presence/absence of aliphatic ^{1}H saturation to probe the structural compactness (local ^{1}H density) and sub-ns time scale dynamics. The intensity ratios ($\lambda_{NOE} = I_{sat}/I_{ref}$) measured for individual amide protons in the N- and I-states of W60G-B2M are compared in the graph of Figure 4b. Low λ_{NOE} values are indicative of rigid, compact local structure, while λ_{NOE} values close to 1 are expected for highly flexible loop regions. From these data we can conclude that, within the experimental error, the two states behave very similarly in terms of local structural compactness and fast time scale dynamics, even in the regions with significant differences in ^{1}H, ^{15}N chemical shifts (Figure 3b,c). Of course, no conclusion can be drawn for the protein regions that are NMR invisible.

In a second experiment, we measured the transverse relaxation rates of the ^{15}N coherence $N_x + 2N_xH_z$ (TROSY line) by inserting a variable spin−echo delay into the $^{1}H-^{15}N$ BEST-TROSY experiment (Figure S4a). This R_2-BEST-TROSY experiment yields information on both fast motions (ps−ns), comprising the molecular tumbling and local bond-vector fluctuations, and conformational exchange dynamics

dx.doi.org/10.1021/ja302598j | J. Am. Chem. Soc. 2012, 134, 8066−8069

occurring on the μs−ms time scale. Relaxation rates were obtained by fitting the cross-peak intensities measured for three relaxation times T (1, 20, and 60 ms) to the function $I(T) = A \exp(-R_2 T)$ (Figure S4b). The measured rates are shown in Figure 4c. A first interesting conclusion from these data is that, for protein regions that do not show significant chemical shift variations between I- and N-states, the measured relaxation rates are also very similar (~10 s^{-1}). This observation indicates similar molecular tumbling correlation times for the I- and N-states, excluding the hypothesis of a monomer−oligomer exchange process as a possible explanation for the missing I-state correlation peaks. Most interestingly, however, transverse relaxation in the I-state is increased in regions close to the NMR-invisible residues, further supporting the idea that the absence of NMR signal is due to conformational exchange processes on the μs−ms time scale in the apical side of the protein close to Pro[32] (Figure 4d). Based on the observed profiles of chemical shift changes and ^{15}N transverse relaxation rates, it is tempting to speculate that this conformational exchange process involves a collective opening motion in the apical half of B2M that is destabilized by the non-native *trans*-prolyl bond. The transient breaking of local structure makes hydrophobic side chains accessible at the surface of the protein, thus favoring B2M oligomerization as the onset of amyloid fibril formation. The conclusions drawn here from our real-time NMR data are in agreement with recent results obtained from a conventional steady-state NMR investigation of Δ6N-B2M,[13] a mutant lacking the first six N-terminal amino acids. NMR relaxation data of Δ6N-B2M also revealed increased μs−ms (but not ps−ns) dynamics in the BC and DE loop regions, similar to our findings on the transient I-state. Our real-time NMR data thus further support the idea that Δ6N-B2M can be considered an equilibrium analogue to the I-state in terms of structure and conformational dynamics.

In summary, we have demonstrated the ability of sensitivity-enhanced, fast real-time 3D NMR spectroscopy combined with high-field NMR instruments to provide insight into the structural and dynamic properties of a protein folding intermediate of B2M at atomic resolution. In particular, we have shown that, once sequential NMR assignment has been obtained, quantitative information on local dynamics at various time scales can be obtained from real-time spin relaxation measurements. This approach can be readily applied to other biologically relevant systems and provides a powerful strategy for NMR-based characterization of protein states with half-lifetime of a few tens of minutes. As demonstrated here, fast real-time 3D NMR represents a unique method for the characterization of "long-lived" transiently populated protein states. It may also prove useful for in-cell NMR studies[16] where the protein of interest has often a similar short lifetime, and where structural changes occurring after some cellular stimulus may be monitored by real-time NMR methods.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional NMR data and experimental details on the algorithm used for reconstructing I-state spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

bernhard.brutscher@ibs.fr

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank A. Corazza and G. Esposito (Udine) for stimulating discussion on this project. This work was financially supported by grants from ANR (BLAN-08-0210, PDOC-10-011-ProtDyn-ByNMR) and EU (FP7-I3-BIO-NMR 261862). We also thank the TGIR-RMN THC FR3050 for access to the 1 GHz NMR spectrometer in Lyon.

■ REFERENCES

(1) Dobson, C. M. *Semin. Cell. Dev. Biol.* **2004**, *15*, 3.
(2) (a) Korzhnev, D. M.; Kay, L. E. *Acc. Chem. Res.* **2007**, *41*, 442. (b) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312. (c) Bouvignies, G.; Vallurupalli, P.; Hansen, D. F.; Correia, B. E.; Lange, O.; Bah, A.; Vernon, R. M.; Dahlquist, F. W.; Baker, D.; Kay, L. E. *Nature* **2011**, *477*, 111.
(3) Balbach, J.; Forge, V.; van Nuland, N. A. J.; Winder, S. L.; Hore, P. J.; Dobson, C. M. *Nat. Struct. Biol.* **1995**, *2*, 865−870.
(4) (a) Mok, K. H.; Nagashima, T.; Day, I. J.; Jones, J. A.; Jones, C. J. V.; Dobson, C. M.; Hore, P. J. *J. Am. Chem. Soc.* **2003**, *125*, 12484. (b) Schanda, P.; Forge, V.; Brutscher, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11257.
(5) Balbach, J.; Forge, V.; Lau, W. S.; van Nuland, N. A.; Brew, K.; Dobson, C. M. *Science* **1996**, *274*, 1161.
(6) (a) Freeman, R.; Kupce, E. *J. Biomol. NMR* **2003**, *27*, 101. (b) Felli, I. C.; Brutscher, B. *Chemphyschem* **2009**, *10*, 1356. (c) Kazimierczuk, K.; Stanek, J.; Zawadzka-Kazimierczuk, A.; Kozminski, W. *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *57*, 420.
(7) (a) Schanda, P.; Brutscher, B. *J. Am. Chem. Soc.* **2005**, *127*, 8014. (b) Schanda, P.; Van Melckebeke, H.; Brutscher, B. *J. Am. Chem. Soc.* **2006**, *128*, 9042. (c) Farjon, J.; Boisbouvier, J.; Schanda, P.; Pardi, A.; Simorre, J. P.; Brutscher, B. *J. Am. Chem. Soc.* **2009**, *131*, 8571.
(8) Favier, A.; Brutscher, B. *J. Biomol. NMR.* **2011**, *49*, 9.
(9) Haupt, C.; Patzschke, R.; Weininger, U.; Gröger, S.; Kovermann, M.; Balbach, J. *J. Am. Chem. Soc.* **2011**, *133*, 11154.
(10) Floege, J.; Ehlerding, G. *Nephron.* **1996**, *72*, 9.
(11) (a) Kameda, A.; Hoshino, M.; Higurashi, T.; Takahashi, S.; Naiki, H.; Goto, Y. *J. Mol. Biol.* **2005**, *348*, 383. (b) Jahn, T. R.; Parker, M. J.; Homans, S. W.; Radford, S. E. *Nat. Struct. Mol. Biol.* **2006**, *13*, 195. (c) Sakata, M.; Chatani, E.; Kameda, A.; Sakurai, K.; Naiki, H.; Goto, Y. *J. Mol. Biol.* **2008**, *382*, 1242. (d) Eichner, T.; Radford, S. E. *J. Mol. Biol.* **2009**, *386*, 1312. (e) Kameda, A.; Morita, E. H.; Sakurai, K.; Naiki, H.; Goto, Y. *Protein Sci.* **2009**, *18*, 1592. (f) Rennella, E.; Corazza, A.; Giorgetti, S.; Fogolari, F.; Viglino, P.; Porcari, R.; Verga, L.; Stoppini, M.; Bellotti, V.; Esposito, G. *J. Mol. Biol.* **2010**, *401*, 286−297. (g) Eichner, T.; Radford, S. E. *FEBS J.* **2011**, *278*, 3868.
(12) Corazza, A.; Rennella, E.; Schanda, P.; Mimmi, M. C.; Cutuil, T.; Raimondi, S.; Giorgetti, S.; Fogolari, F.; Viglino, P.; Frydman, L.; Gal, M.; Bellotti, V.; Brutscher, B.; Esposito, G. *J. Biol. Chem.* **2010**, *285*, 5827−5835.
(13) Eichner, T.; Kalverda, A. P.; Thompson, G. S.; Homans, S. W.; Radford, S. E. *Mol. Cell* **2011**, *41*, 161.
(14) Schwarzinger, S.; Kroon, G. J.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970.
(15) Schanda, P.; Forge, V.; Brutscher, B. *Magn. Reson. Chem.* **2006**, *44*, S177.
(16) (a) Reckel, S.; Hänsel, R.; Löhr, F.; Dötsch, V. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 91. (b) Inomata, K.; Ohno, A.; Tochio, H.; Isogai, S.; Tenno, T.; Nakase, I.; Takeuchi, T.; Futaki, S.; Ito, Y.; Hiroaki, H.; Shirakawa, M. *Nature* **2009**, *458*, 106. (c) Selenko, P.; Frueh, D. P.; Elsaesser, S. J.; Haas, W.; Gygi, S. P.; Wagner, G. *Nat. Struct. Mol. Biol.* **2008**, *15*, 321.