JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Resolution of Site-Specific Conformational Heterogeneity in Proline-Rich Molecular Recognition by Src Homology 3 Domains

Rachel E. Horness, Edward J. Basom, John P. Mayer, and Megan C. Thielges*

Department of Chemistry, Indiana University, 800 East Kirkwood, Bloomington, Indiana 47405, United States

Supporting Information

ABSTRACT: Conformational heterogeneity and dynamics are increasingly evoked in models of protein molecular recognition but are challenging to experimentally characterize. Here we combine the inherent temporal resolution of infrared (IR) spectroscopy with the spatial resolution afforded by selective incorporation of carbon-deuterium (C−D) bonds, which provide frequency-resolved absorptions within a protein IR spectrum, to characterize the molecular recognition of the Src homology 3 (SH3) domain of the yeast protein Sho1 with its cognate proline-rich (PR) sequence of Pbs2. The IR absorptions of C−D bonds introduced at residues along a peptide of the Pbs2 PR sequence report on the changes in the local environments upon binding to the SH3 domain. Interestingly, upon forming the complex the IR spectra of the peptides labeled with C−D bonds at either of the two conserved prolines of the PXXP consensus recognition sequence show more absorptions than there are C−D bonds, providing evidence for the population of multiple states. In contrast, the NMR spectra of the peptides labeled with ^{13}C at the same residues show only single resonances, indicating rapid interconversion on the NMR time scale. Thus, the data suggest that the SH3 domain recognizes its cognate peptide with a component of induced fit molecular recognition involving the adoption of multiples states, which have previously gone undetected due to interconversion between the populated states that is too fast to resolve using conventional methods.

Molecular recognition underlies all protein function and there is growing speculation that conformational heterogeneity (the population of different states) and dynamics (the interconversion between the populated states) contribute.[1,2] The conformational selection[3] and induced fit[4] models of molecular recognition, for example, explicitly evoke conformational heterogeneity and/or dynamics. The dynamics involved can cover a wide range of time scales. The contributions of states that interconvert on the ms and longer time scale have long been appreciated, at least in part due to our ability to detect them using conventional methods such as NMR spectroscopy. However, there is growing evidence that faster dynamics may contribute,[1,2,5−7] for example, being thought to dominate changes in the conformational binding entropy,[5,8] and to contribute to allostery,[8,9] the affinity maturation of antibodies,[10] and selectivity of enzyme catalysis.[7] In fact, any induced fit component of molecular recognition relies on dynamics that are fast relative to the lifetime of an initial encounter complex (ns).[2,6] Nonetheless, our understanding of the exact role of fast dynamics remains incomplete, at least in part due to the lack of experimental techniques capable of their characterization with sufficient temporal resolution to probe all populated states, as well as sufficient spatial resolution to probe specific parts of a protein.

In principle, infrared (IR) spectroscopy is capable of rigorously characterizing the dynamics of any molecule because its inherently fast (ps) time scale ensures detection of even rapidly interconverting states. Unfortunately, spectral congestion makes the characterization of any individual protein vibration challenging. However, this issue can be alleviated by incorporation of carbon−deuterium (C−D) bonds, which provide completely nonperturbative probes with IR frequencies in a spectral region free of any intrinsic protein absorptions.[11] The characterization of the incorporated C−D probes provides information about the presence and nature of populated environments from the number/relative intensity, line widths, and frequencies of observed absorption bands. Moreover, how the spectra change upon complex formation provides information about the mechanism of molecular recognition. For example, a conformational selection mechanism involves the sampling of multiple conformations in the free state, and from the ensemble the binding partner recognizes and selectively stabilizes the bound conformation.[3] This mechanism thus predicts the observation of multiple absorptions in the free state, of which one then dominates upon complex formation. In contrast, an induced-fit mechanism involves conformational changes that occur as a result of binding,[4] predicting that the number, line width, and/or frequency of IR absorptions will differ in the free and bound states. Obviously, these mechanisms are not mutually exclusive,[12] but importantly, they predict different spectral changes, while the lock-and-key model predicts little change.

Binding of Src homology 3 (SH3) domains to proline-rich (PR) sequence motifs has emerged as a paradigm of molecular recognition.[13−15] SH3 domains are among the most prevalent in human and other eukaryotic proteomes, and their binding of different PR sequence motifs (PXXP) underlies diverse cellular processes, including signaling, cytoskeletal remodeling, and development.[14,15] Interestingly, previous studies of SH3-mediated PR sequence recognition have implicated a role for protein dynamics.[16−19] In many cases, for instance, cognate peptide recognition by the Src, Lck, and Sem-5 SH3 domains,[16−18] NMR relaxation studies find conformational

Received: November 16, 2015
Published: January 19, 2016

DOI: 10.1021/jacs.5b11999
J. Am. Chem. Soc. 2016, 138, 1130−1133

restriction of ns−ps motions in loop region(s) of the SH3 domains upon PR peptide binding. Interestingly, NMR chemical shift perturbation analysis[19] has provided evidence for the existence of multiple dynamically engaged states in the complex of the Abp1p SH3 domain and its cognate PR peptides; however, the presence of such heterogeneity remains uncertain as the individual states could not be observed, presumably due to their rapid interconversion on the NMR time scale.

Toward investigating the molecular recognition of the PR sequence of Pbs2 by the SH3 domain of Sho1 ($SH3^{Sho1}$), we prepared peptides consisting of the Pbs2 PR recognition sequence (Ac-VNKPLPPLPVA-NH_2) individually labeled with (methyl-d_3)Leu at L5′ or (C_α-d_1,C_δ-d_2)Pro (d_3Pro) at P6′, P7′, or P9′ (d_3L5′, d_3P6′, d_3P7′, and d_3P9′, respectively, where the prime denotes the peptide) (Figure 1) and characterized them by

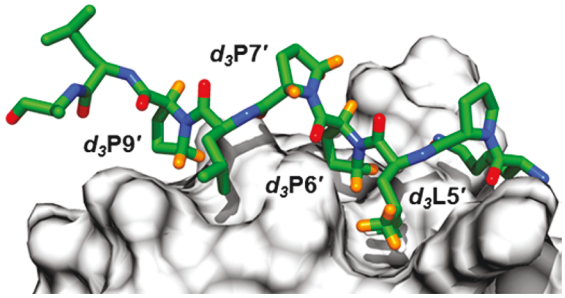

Figure 1. Structural model of $SH3^{Sho1}$-PR peptide binding interface (pdb ID 2VKN). Deuterium labels are highlighted in orange. Image generated in UCSF Chimera.

FT IR spectroscopy both in the free state and when bound to $SH3^{Sho1}$. The spectra were fit to a sum of Gaussian functions to extract the frequencies, line widths, and amplitudes of the C−D absorption bands (Table 1, Supporting Information). To aid spectral interpretation, molecular dynamics (MD) simulations were performed for the peptide-$SH3^{Sho1}$ complex (Supporting Information).

The FT IR spectrum of d_3L5′ in the free state shows two overlapping bands of high intensity at high frequency (2213 and 2226 cm^{-1}) and a band of moderate intensity at low frequency (2069 cm^{-1}), which, based on previous literature,[20] we assign to the two asymmetric stretches and the symmetric stretch vibrations of the d_3-methyl group, respectively (Figure 2A). A very weak band at intermediate frequency (2125 cm^{-1}) is also apparent, which has been attributed to a combination or

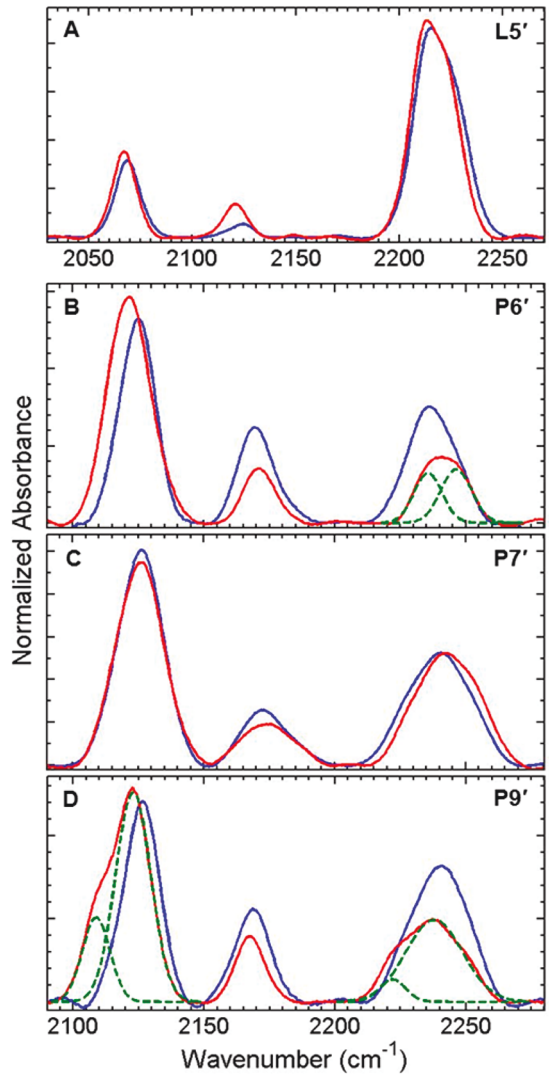

Figure 2. FT IR spectra of free (blue) and bound (red) PR peptide individually labeled with d_3Leu (A) or d_3Pro (B−D). Gaussian fits are shown as green dashed lines. Spectra are normalized to the total integrated area of the symmetric and asymmetric bands.

Table 1. FT IR Gaussian Fit Parameters for Individually Labeled PR Peptides

		$C_\delta D_2$ symmetric stretch		$C_\delta D_2$ asymmetric stretch		$C_\alpha D$ stretch	
		ν (cm^{-1})	fwhm[a] (cm^{-1})	ν (cm^{-1})	fwhm (cm^{-1})	ν (cm^{-1})	fwhm (cm^{-1})
d_3L5′	free	2068.8 ± 0.1	14.2 ± 0.2	2212.8 ± 1.0	16.7 ± 1.8		
				2226.5 ± 1.7	18.6 ± 1.8		
	bound	2067.2 ± 0.1	13.7 ± 0.6	2211.5 ± 0.7	16.9 ± 1.3		
				2224.5 ± 1.1	16.6 ± 2.1		
d_3P6′	free	2124.7 ± 0.9	15.6 ± 1.7	2236.5 ± 1.0	20.3 ± 2.8	2169.3 ± 0.6	14.8 ± 1.1
	bound	2121.2 ± 0.6	18.5 ± 1.2	2247.0 ± 0.9	13.6 ± 2.4	2171.3 ± 0.4	14.9 ± 1.5
				2235.9 ± 1.1	13.6 ± 2.5		
d_3P7′	free	2125.8 ± 0.3	19.9 ± 1.6	2239.9 ± 2.6	28.2 ± 1.4	2173.1 ± 2.1	20.4 ± 2.4
	bound	2125.5 ± 0.4	20.8 ± 2.4	2242.1 ± 2.2	28.3 ± 1.7	2174.2 ± 1.1	23.2 ± 1.9
d_3P9′	free	2126.1 ± 0.1	15.8 ± 0.4	2240.6 ± 1.9	24.8 ± 2.8	2168.5 ± 0.3	15.7 ± 0.6
	bound	2122.6 ± 0.2	15.2 ± 1.0	2238.8 ± 1.2	24.7 ± 1.3	2167.3 ± 0.6	12.5 ± 0.7
		2108.8 ± 0.4	11.8 ± 1.0	2221.9 ± 0.8	10.5 ± 1.8		

[a]full width at half the maximum peak amplitude.

DOI: 10.1021/jacs.5b11999
J. Am. Chem. Soc. 2016, 138, 1130−1133

overtone bending mode and was not analyzed further.[21] Upon binding to SH3^{Sho1}, the symmetric stretch shifts by 1.6 cm^{-1} to lower frequency, and the asymmetric stretches appear to red-shift by a similar amount, but for these latter modes the differences are within error. These spectral changes are consistent with a slightly less polar environment for the L5′ side chain upon binding SH3^{Sho1}. The bound spectrum is however similar to those of d_3Leu previously characterized in an unfolded protein,[22] suggesting the environment of the side chain may be more polar than expected from the crystal structure, which shows the side chain packed within a hydrophobic pocket (Figure 1).[23] This is consistent with trajectories from the MD simulations, in which the L5′ side chain adopts configurations more solvent-exposed than found in the crystal structure (Figure S18).

The IR spectrum of each free d_3Pro-labeled PR peptide shows three bands, each of which is well fit by a single Gaussian (Table 1, Figure 2B–D). Based on previous literature,[24] we assign the bands at lowest (∼2125 cm^{-1}) and highest (∼2240 cm^{-1}) frequency to the symmetric and asymmetric stretches, respectively, of the $C_\delta D_2$ methylene groups, while we assign the band at intermediate frequency (∼2170 cm^{-1}) to the $C_\alpha D$ vibration. In the free state, the spectra of the labeled variants are generally similar. However, the $C_\delta D_2$ frequencies for d_3P6′ are lower in frequency (∼1 and 4 cm^{-1} for symmetric and asymmetric stretches, respectively) than those for d_3P7′ and d_3P9′ labeled PR peptides. In addition, all the absorptions of the d_3P7′ peptide are broader than those of the other two labeled peptides, and the $C_\alpha D$ band is shifted by ∼4 cm^{-1} to higher frequency. Thus, small position-dependent differences exist among the residue environments in the free peptide.

Unlike the IR spectra of the free d_3Pro-labeled PR peptide variants, those of the SH3^{Sho1}-bound peptides show large and position-dependent differences. For d_3P7′, no significant changes occur upon binding to SH3^{Sho1} (Figure 2C). In contrast, the $C_\delta D_2$ absorptions of both d_3P6′ and d_3P9′ PR peptides show substantial changes upon SH3^{Sho1} binding. For d_3P9′, both the symmetric and asymmetric stretch $C_\delta D_2$ absorptions require two Gaussians for fitting (Figures 2D and S8). The dominant band is similar to the single absorption observed in the free state, with the only significant change being a 3.5 cm^{-1} red-shift of the symmetric stretch, and the second, minor band is both highly red-shifted (∼18 cm^{-1}) and narrower. For d_3P6′, as with d_3P9′, fitting the asymmetric stretch absorption in the bound state requires two Gaussians. In this case, the Gaussians are of similar integrated absorbance, and one again has the same frequency as the single absorption in the free state, although it is narrower, and the second is narrower and dramatically shifted to higher energy by ∼11 cm^{-1}. The symmetric stretch absorption can be well fit by a single Gaussian band shifted to lower frequency by 3.5 cm^{-1} but is broader in line width (Figure 2B).

To verify that the observed changes are associated with specific binding, we synthesized the d_3P9′ peptide where P6′ was mutated to Ala (d_3P9′/P6′A). The altered peptide binds ∼150-fold weaker to SH3^{Sho1} (Figure S4), and under the conditions employed, 71% of d_3P9′/P6′A is expected to be bound. The IR spectra of the mixture of SH3^{Sho1} and d_3P9′/P6′A was well fit by a superposition of the bound and unbound spectra of d_3P9′, with relative contributions of 63% and 37%, respectively, demonstrating that the complexity of the bound form results from specific binding to the SH3 domain (Supporting Information).

It is unlikely that the additional bands observed for the $C_\delta D_2$ stretch absorptions of d_3P9′ and d_3P6′ originate from a Fermi resonance. This conclusion is based on the relative intensities and symmetry restrictions (see Supporting Information for further discussion). They are also unlikely to arise from contributions of the fully unbound state, as none of the absorptions, with the single exception of the $C_\delta D_2$ asymmetric stretch of P9′, match those of the free peptide. We therefore attribute the multiple bands observed for d_3P9′ and d_3P6′ to the population of multiple conformations in the bound state. To further investigate this conformational heterogeneity, we prepared and characterized by NMR spectroscopy the PR peptide selectively labeled with $^{13}C_5$Pro at P9′ or P6′ (Supporting Information). The 800 MHz $^{13}C-^1H$ HSQC NMR spectra of the free PR peptides show $^{13}C_\delta H_2$ resonances with similar chemical shifts along both the carbon and proton dimensions, consistent with the similar C–D absorptions measured at the equivalent sites with FT IR spectroscopy (Figure 3). The resonances for both residues show chemical shift

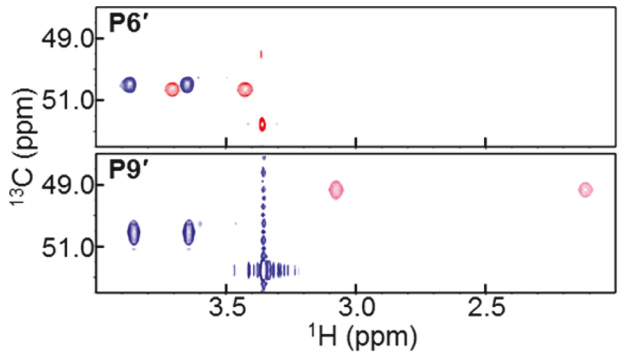

Figure 3. $^{13}C-^1H$ HSQC spectra of $^{13}C_5$P6′ (top) and $^{13}C_5$P9′ (bottom) of the free (blue) and SH3^{Sho1}-bound (red) PR peptide.

perturbation upon binding SH3^{Sho1} but, interestingly, neither shows evidence for the appearance of multiple resonances in the bound state in either the ^1H or ^{13}C dimension. While the absence of multiple resonances may result from interconversion of the two IR-observed conformations or from binding and dissociation that are fast on the NMR time scale, it is clear that the more conventional technique failed to provide any evidence of conformational heterogeneity.

To further support the presence of heterogeneity in the bound complex, as well as to elucidate the nature of the two states populated, we performed MD simulations using the AMBER software package (Supporting Information). Analysis of the MD trajectories reveals that the sampled structures of the complex generally cluster into two bound states that place the P6′ and P9′ side chains in two distinct environments (Figure 4). One state is similar to that found for the crystal structure, while in the second the peptide is shifted and rotated relative to the SH3^{Sho1} surface (Figure S14). In the former state, the side chain of P9′ closely packs between the aromatic rings of two Tyr side chains (Figures 4A,C and S17), an environment thought to promote C–H···π interactions that are expected to polarize the $C_\delta D_2$ bonds,[25] and the population of the state that optimizes this interaction may be the origin of the dramatically red-shifted absorptions observed in the complex. In the second state, the $C_\delta D_2$ bonds of d_3P9′ are more solvent-exposed, suggesting their association with the absorptions more similar to the free state. In contrast, the P6′ side chain shows significant contact with the surface of SH3^{Sho1} in both of the two states (Figures 4B,D and S15). One state places the P6′ side chain in proximity to the positively charged side chain of Arg39, which might be the origin of the shifted

DOI: 10.1021/jacs.5b11999
J. Am. Chem. Soc. 2016, 138, 1130–1133

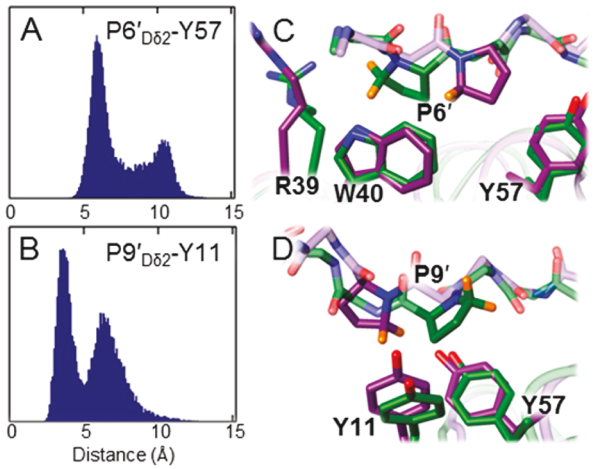

Figure 4. Histograms from analysis of MD trajectories showing relative occurrence of distances between proline $C_\delta D$ for (A) P6′ and center of aromatic ring of Y57 of $SH3^{Sho1}$ and for (B) P9′ and center of aromatic ring of Y11 of $SH3^{Sho1}$. Structural models showing environment of (C) P6′ and (D) P9′ in the two bound states based on MD snapshots representative of the center of the histogram distributions shown in A and B.

absorptions. P7′ is also predicted to populate two distinct conformations, one more hydrophobic and one more solvent exposed. While only single absorptions were observed for the symmetric and asymmetric stretches of d_3P7′, its increased line width relative to the other peptides is consistent with the population of these two states. Thus, the calculations strongly support the suggestion that the $SH3^{Sho1}$-PR peptide complex adopts two states.

Together, the IR data and simulations suggest heterogeneity in the $SH3^{Sho1}$-PR peptide complex. The binding-induced spectral changes observed for the PR peptides imply different bound and free states, which suggests that recognition is mediated, at least in part, by an induced fit mechanism. While an induced fit mechanism has been evoked for SH3-mediated molecular recognition,[26,27] this is the first direct observation of rapidly interconverting states that might facilitate it. Moreover, the association of the conformational heterogeneity with the conserved PXXP motif suggests that it is a ubiquitous feature of SH3 molecular recognition. More generally, the data demonstrate that the inherent fast time scale of IR spectroscopy and the spatial selectivity afforded by incorporation of C–D bonds permits detection and initial characterization of states that interconvert too fast to be resolved by the more conventional NMR techniques, and so otherwise would likely remain unnoticed. Application of improved tools for direct characterization of fast dynamics will make possible the elucidation of their involvement in the molecular recognition of these and other proteins.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b11999.

Experimental details of protein expression, peptide preparation, spectroscopic analysis, and MD simulations; complete fit parameters, plots of example fits, details of model evaluation; full NMR spectra, amide I′ IR data, plots of fluorescence binding titrations, peptide mass spectra, and MD structural models (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*thielges@indiana.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge Alan Davidson for $SH3^{Sho1}$ expression plasmid, Richard DiMarchi for materials for peptide synthesis, Hongwei Wu for help with NMR spectroscopy, and David Giedroc for helpful discussions. R.E.H. and E.J.B. were supported by the Graduate Training Program in Quantitative and Chemical Biology (T32 GM109825) and Indiana University.

■ REFERENCES

(1) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450*, 964–972.
(2) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789–796.
(3) Foote, J.; Milstein, C. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 10370–10374.
(4) Koshland, D., Jr *Proc. Natl. Acad. Sci. U. S. A.* **1958**, *44*, 98.
(5) Lee, A.; Kinnear, S.; Wand, A. *Nat. Struct. Biol.* **2000**, *7*, 72–77.
(6) Zhou, H.-X. *Biophys. J.* **2010**, *98*, L15–L17.
(7) Basom, E. J.; Spearman, J. W.; Thielges, M. C. *J. Phys. Chem. B* **2015**, *119*, 6620–6627.
(8) Stone, M. J. *Acc. Chem. Res.* **2001**, *34*, 379–388.
(9) Tzeng, S.; Kalodimos, C. G. *Curr. Opin. Struct. Biol.* **2011**, *21*, 62–67.
(10) Zimmermann, J.; Oakman, E. L.; Thorpe, I. F.; Shi, X.; Abbyad, P.; Brooks, C. L.; Boxer, S. G.; Romesberg, F. E. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13722–13727.
(11) Chin, J.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 2426–2427.
(12) Liu, S.-Q.; Ji, X.-L.; Tao, Y.; Tan, D.-Y.; Zhang, K.-Q.; Fu, Y.-X. Protein folding, binding and energy landscape: a synthesis. In *Protein Engineering*; Kaumaya, P., Ed.; InTech: Rijeka, Croatia, 2012; pp 207–252.
(13) Pawson, T.; Nash, P. *Genes Dev.* **2000**, *14*, 1027–1047.
(14) Pawson, T.; Scott, J. D. *Science* **1997**, *278*, 2075–2080.
(15) Mayer, B. J. *Journal of Cell Science* **2001**, *114*, 1253–1263.
(16) Bauer, F.; Sticht, H. *FEBS Lett.* **2007**, *581*, 1555–1560.
(17) Ferreon, J. C.; Hilser, V. *Protein Sci.* **2003**, *12*, 982–996.
(18) Wang, C.; Pawley, N. H.; Nicholson, L. K. *J. Mol. Biol.* **2001**, *313*, 873–887.
(19) Stollar, E. J.; Lin, H.; Davidson, A. R.; Forman-Kay, J. D. *PLoS One* **2012**, *7*, e51282.
(20) Cremeens, M. E.; Fujisaki, H.; Zhang, Y.; Zimmermann, J.; Sagle, L. B.; Matsuda, S.; Dawson, P. E.; Straub, J. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6028–6029.
(21) Zimmermann, J.; Gundogdu, K.; Cremeens, M. E.; Bandaria, J. N.; Hwang, G. T.; Thielges, M. C.; Cheatum, C. M.; Romesberg, F. E. *J. Phys. Chem. B* **2009**, *113*, 7991–7994.
(22) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 3384–3385.
(23) PDB ID: 2VKN.
(24) Adhikary, R.; Zimmermann, J.; Liu, J.; Forrest, R. P.; Janicki, T. D.; Dawson, P. E.; Corcelli, S. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2014**, *136*, 13474–13477.
(25) Zondlo, N. *Acc. Chem. Res.* **2013**, *46*, 1039–1049.
(26) Arold, S.; Franken, P.; Strub, M.-P.; Hoh, F.; Benichou, S.; Benarous, R.; Dumas, C. *Structure* **1997**, *5*, 1361–1372.
(27) Cordier, F.; Wang, C.; Grzesiek, S.; Nicholson, L. K. *J. Mol. Biol.* **2000**, *304*, 497–505.

DOI: 10.1021/jacs.5b11999
J. Am. Chem. Soc. 2016, 138, 1130–1133