THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

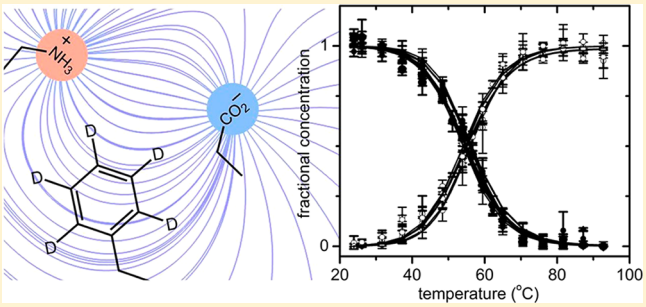
Experimental Characterization of Electrostatic and Conformational Heterogeneity in an SH3 Domain

Ramkrishna Adhikary, Jörg Zimmermann, Jian Liu, Philip E. Dawson, and Floyd E. Romesberg*

Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

Supporting Information

ABSTRACT: Electrostatic and conformational heterogeneity make central contributions to protein function, but their experimental characterization requires a combination of spatial and temporal resolution that is challenging to achieve. Src homology 3 (SH3) domains mediate protein−protein interactions, and NMR studies have demonstrated that most possess conformational heterogeneity, which could be critical for their function. Here, we use the IR absorptions of carbon−deuterium (C−D) bonds site-selectively incorporated throughout the N-terminal SH3 domain from the murine adapter protein Crk-II to characterize its different microenvironments with high spatial and temporal resolution. The C−D absorptions are only differentiated in the folded state of the protein where they show evidence of significant environmental heterogeneity. However, the spectra of the folded state are independent of temperature, and upon thermal denaturation the protein undergoes a single, global unfolding transition. While some evidence of conformational heterogeneity is found within the peptide backbone, the majority of the environmental heterogeneity appears to result from electrostatics.



INTRODUCTION

There is great interest in characterizing microenvironmental heterogeneity within proteins due to its potential roles in biological function. For example, conformational heterogeneity has important ramifications for protein evolution, as the different conformations may possess different physicochemical properties and might serve as starting points for the evolution of new function.[1−3] While conformational heterogeneity involving states that exchange relatively slowly may be characterized via X-ray crystallography[4] or NMR spectroscopy,[5−7] the heterogeneity involving more rapidly exchanging states remains more difficult to characterize, especially with site-specific resolution. Proteins also have internal electric fields that may establish unique electrostatic microenvironments (i.e., electrostatic heterogeneity) that may also contribute to biological function.[8,9] However, experimental techniques capable of characterizing the electrostatic environments at different positions within a protein in the absence of potentially perturbative probes are not generally available.

The temporal resolution of infrared (IR) spectroscopy is inherently high, and when combined with site-specific labeling techniques, IR spectroscopy also provides high structural resolution.[10] To take advantage of this temporal and structural resolution to study electrostatic and conformational heterogeneity, we have been developing a technique based on the site-specific incorporation of carbon−deuterium (C−D) bonds within proteins.[11−18] C−D bonds are nonperturbative and provide stretching absorptions within an otherwise unobstructed region of a protein's IR spectrum (~2000−2300 cm^{-1}). Because C−D bonds are inert and local mode in character, when not adjacent to a heteroatom where hyperconjugative effects can contribute,[11] the observed shifts in their absorption frequencies when different environments are probed (e.g., the same site within the folded versus unfolded protein or two different sites within the folded protein) largely reflect differences in local electrostatics.[19] Correspondingly, the observation of multiple C−D absorptions originating from a single C−D bond is direct evidence of multiple local environments (i.e., conformational heterogeneity).[15]

SH3 (Src homology 3) domains are found in many different proteins and function autonomously to mediate protein−protein interactions. They have emerged as a paradigm for biological molecular recognition;[20−24] and NMR studies have demonstrated that their native states are commonly conformationally heterogeneous.[25−33] SH3 domains have also served as a paradigm for equilibrium folding, emerging as the prototypical two state folder.[34−40] However, at least with some SH3 domains this has recently been called into question. For example, a concentration-dependent unfolding midpoint was taken as evidence of a more complex associative process with Spc and α-II spectrin SH3 domains,[41,42] and mutation- or condition-dependent intermediates have been detected during the equilibrium unfolding of Fyn[43−46] and PI3 SH3

Special Issue: Peter G. Wolynes Festschrift

Received: March 20, 2013
Revised: June 20, 2013
Published: July 9, 2013

domains.[47−49] Thus, many SH3 domains possess conformational heterogeneity within the folded state or upon slight destabilization of the folded state. In contrast, NMR studies of the N-terminal SH3 domain from the murine adapter protein Crk-II (hereafter referred to as nSH3) have revealed no evidence of conformational heterogeneity,[36,37,50,51] which challenges the idea that it is generally required for SH3 function. However, using site-selectively incorporated C−D bonds, we reported that the $C_\alpha D_2$ stretching absorptions of (d_2)Gly incorporated at position Gly180 of nSH3 were doubled,[15] suggesting that nSH3 might possess at least some backbone conformational heterogeneity involving exchange between states that is too fast to detect by NMR.

To more rigorously explore the electrostatic and conformational heterogeneity of nSH3, eleven sites were selected for deuteration and their FT IR spectra were characterized within the folded and thermally unfolded proteins (Figure 1). The

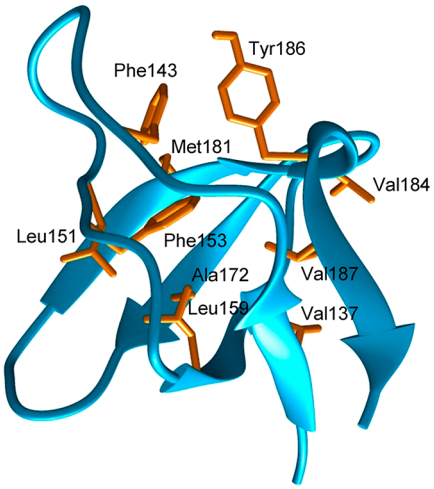


Figure 1. Structure of nSH3 (PDB 1CKA). Side chains of deuterated residues are shown in orange.

spectra at several sites of the folded protein show unique, site-dependent absorptions indicative of microenvironmental heterogeneity. To differentiate between electrostatic and conformational origins of the heterogeneity, as well as to characterize the equilibrium folding of nSH3, we also characterized the FT IR spectra as a function of temperature. We find that most sites probed experience the same thermally induced two-state transition to the unfolded state and that the spectra associated with the folded state are themselves independent of temperature, strongly arguing against conformational heterogeneity as the origin of the observed microenvironmental heterogeneity. Thus, the data suggest that, unlike other characterized SH3 domains, nSH3 is relatively rigid, but that, despite its small size, it possesses a wide range of electrostatically distinct microenvironments that likely contribute to its biological functions.

EXPERIMENTAL METHODS

Synthesis of nSH3. nSH3 was synthesized manually as described previously[15,52] using *in situ* neutralization protocols for Boc solid phase peptide synthesis with side-chain protected amino acids, including Arg(Tos), Asn(Xan), Asp(OcHx), Gln(Xan), Glu(OcHx), Lys(2-ClZ), Ser(OBzl), Trp(formyl) and Tyr(2-BrZ). Proteo and (d_3)Met181 nSH3 were synthesized on Boc-Arg(Tos)-Pam resin, while (d_8)Val137, (d_8)Phe143, (d_7)Leu151, (d_8)Phe153, (d_7)Leu159, (d_1)Leu159, (d_3)Ala172, (d_8)Val184, (d_4)Tyr186, and (d_8)Val187 nSH3 were synthesized on *p*-methylbenzhydrylamine resin.

Synthesis was carried out on 0.4 mmol of polystyrene resin. All nondeuterated amino acids were coupled for 20 min with 2.2 mmol of amino acid, 4 mL of 0.5 M HCTU (2.0 mmol), and 700 μL of DIEA (∼4.1 mmol) except for Gln. Boc deprotection was performed via 2 min incubation in neat TFA. DMF flow washes (30 s) were used between coupling/deprotection and between deprotection/coupling. Gln couplings were performed for 30 min, and a DCM wash was performed before and after Boc deprotection. The quality of the synthetic product was improved through a TFA flow wash (15 mL) before the 2 min TFA Boc deprotection, rapid removal of TFA after Boc deprotection by DMF flow wash, and constant mixing of resin during all procedures by N_2 or Ar gas.

Deuterated residues (1.25 to 1.5 equiv, 0.5 to 0.6 mmol) were coupled for at least 1 h with ∼900 to 1100 μL of 0.5 M HCTU and ∼160 to 190 μL of DIEA. In cases where the total solution volume was not sufficient to cover the resin bed, DMF was added. Couplings were monitored by ninhydrin. The N-terminal Boc, side-chain protected peptide resins were stored at −20 °C. As needed, ∼400 mg of resin was swollen in DMF for 20 min and Boc deprotected with neat TFA, and then the Trp formyl protecting groups were removed via incubation with 10% piperidine/DMF (v/v) for 1 h on ice. The peptide was side-chain deprotected and cleaved from the dry resin using 10 mL of anhydrous HF with 10% anisole (v/v) for 1 h at 0 °C. The crude peptide was precipitated, washed with chilled diethyl ether, dissolved in 10% aqueous (v/v) AcOH, and lyophilized.

HPLC was performed using a linear gradient of solvent A (0.1% (v/v) aqueous TFA) and solvent B (90% acetonitrile (v/v), 10% water (v/v), 0.1% TFA (v/v)) and a C12 or C18 reverse phase column. Preparative HPLC was performed with a solvent gradient of 30% B to 40% B over 60 min with a constant 15 mL/min flow rate, while monitoring at 230 nm. Analytical HPLC with the pure peptide was performed with a solvent gradient of 20% B to 50% B over 40 min with a constant 1.5 mL/min flow rate while monitoring at 220 nm, and purity was confirmed by mass spectrometry (Supporting Information).

Pure lyophilized peptide was washed and concentrated with 4 × 15 mL of sodium phosphate buffer (5 mM, pH 7.3) using a 3000 MWCO centrifugal filter (Millipore). For each concentration step, the final volume of the solution did not drop below ∼1.5 mL. After the final washing step, the concentration of the peptide was determined by UV−vis absorption (using an extinction coefficient of 15220 M^{-1} cm^{-1} for SH3[15]). Lyophilized aliquots of 0.1 μmol of protein were stored at −20 °C and were dissolved in 50 or 10 μL of buffer (100 mM sodium phosphate, 200 mM NaCl, pH 7.3) immediately before use to provide 2 or 10 mM protein solutions.

FT IR Spectra. Approximately 8 μL of the protein sample of either 2 or 10 mM nSH3 was loaded into a temperature-controllable demountable liquid cell with CaF_2 windows and a 75 μm Teflon spacer (Harrick Scientific Products, Inc., model TFC-M13, connected to a Watlow Series 989 1/8 DIN temperature controller). A Bruker Equinox 55 FT IR spectrometer equipped with liquid N_2 cooled MCT detector was used to record the FT IR spectra and was continuously purged with dry N_2 gas. 8,000 scans with 2 cm^{-1} resolution were averaged, and a Blackman−Harris 3-term apodization

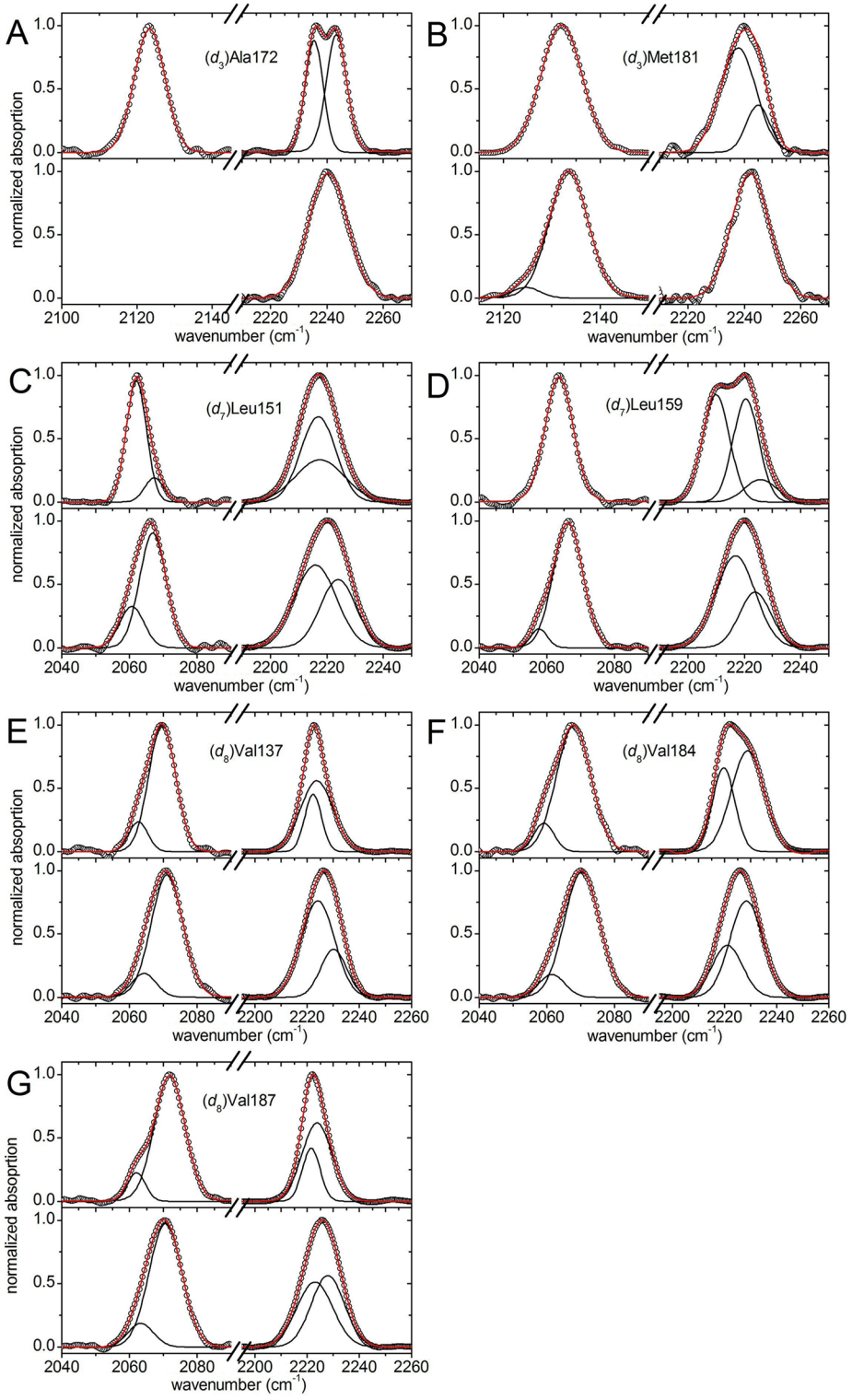


Figure 2. Room temperature (top panels) and high temperature (bottom panels), symmetric (lower frequency) and asymmetric (higher frequency) absorption bands of (A) Ala-, (B) Met-, (C, D) Leu-, and (E−G) Val-labeled proteins. Spectra are shown as circles, while individual pseudo-Voigt profiles and their sum are shown as black and red lines, respectively.

function was used in the Fourier transform. Labview (National Instruments Corp.) was used to control the temperature cell and spectrometer for automation. For background correction, samples of proteo nSH3 were handled identically, and their FT IR spectra were collected under identical conditions.

A series of FT IR spectra for both deuterated and proteo samples were collected between 24 and 93 °C at 5 °C intervals, and analyzed as described previously.[53,54] Briefly, the corresponding proteo spectrum was autosubtracted from the deutero spectrum using the OPUS software (Bruker). The resultant difference spectra were further background corrected using Matlab (Mathworks, Inc.) by fitting and subsequently subtracting a polynomial of varying order to the spectra while excluding the spectral areas that correspond to the C−D

dx.doi.org/10.1021/jp402772x | J. Phys. Chem. B 2013, 117, 13082−13089

absorptions. Linear combinations of the low and high temperature spectra were used to fit the spectra at intermediate temperatures, varying only their total amplitudes, $A_i(T)$, where i corresponds to the low and high temperature state, respectively. For all the deuterated sites, this approach provided adequate fits at all intermediate temperatures. We also collected spectra for the corresponding deuterated free amino acids in buffer as a function of temperature and found only miniscule spectral changes over the entire temperature range (Supporting Information).

The resulting $A_i(T)$ for each deuterated site was then fit according to

$$A_i(T) = a_i \frac{\exp[\Delta H^\circ_{Si\rightarrow H}(T^{-1} - T_{m,i}{}^{-1})/k_B]}{\sum_j^N \exp[\Delta H^\circ_{Si\rightarrow H}(T^{-1} - T_{m,i}{}^{-1})/k_B]}$$

yielding a_i, the temperature-independent absorbance of the ith species at 100% population, $\Delta H^\circ_{Si\rightarrow H}$, the site-specific enthalpy of unfolding, and $T_{m,i} = \Delta H^\circ_{Si\rightarrow H}/\Delta S^\circ_{Si\rightarrow H}$, the site-specific midpoint temperature of unfolding from state i, as fit parameters. The site-specific Gibbs free energy of unfolding was calculated according to $\Delta G^\circ_{Si\rightarrow H} = \Delta H^\circ_{Si\rightarrow H} - T\Delta S^\circ_{Si\rightarrow H}$ (assuming that $\Delta H^\circ_{Si\rightarrow H}$ and $\Delta S^\circ_{Si\rightarrow H}$ are temperature independent in the temperature range of interest).

In addition, the spectra at 24 and 93 °C were deconvoluted into pseudo-Voigt profiles $I(\nu)$ to detect homogeneous contributions to line width according to

$$I(\nu) = \frac{A}{m + (1-m)\sqrt{\pi \ln 2}}\left(m\frac{\text{fwhm}^2}{4(\nu-\nu_0)^2 + \text{fwhm}^2} + (1-m)\sqrt{\pi \ln 2}\,\exp\left(-4(\ln 2)\left(\frac{\nu-\nu_0}{\text{fwhm}}\right)^2\right)\right)$$

where A is the amplitude at center frequency ν_0, m varies between 0 and 1 and describes the character of the pseudo-Voigt profile, which is purely Lorentzian for $m = 1$ and purely Gaussian for $m = 0$, and fwhm is the full width at half-maximum of the pseudo-Voigt profile.

RESULTS AND DISCUSSION

We selected eleven sites that are spread throughout nSH3 for deuteration (Figure 1). Perdeuterated valine was incorporated at position 137, 184, or 187 ((d_8)Val137, (d_8)Val184, or (d_8)Val187, respectively), phenylalanine at position 143 or 153 ((d_8)Phe143 or (d_8)Phe153), and leucine at position 151 or 159 ((d_7)Leu151 or (d_7)Leu159), while side-chain perdeuterated tyrosine was incorporated at position 186 ((d_4)Tyr186), (methyl-d_3) alanine at position 172 ((d_3)Ala172), (methyl-d_3) methionine at position 181 ((d_3)Met181), and C_α–D leucine at position 159 ((d_1)Leu159).

Spectral Analysis of C–D Stretch Absorptions. Background-corrected IR spectra of 10 mM samples of each protein were collected at 24 °C, where the protein is folded, and at 93 °C, where the protein is unfolded. Each spectrum was then deconvoluted into the maximum number of pseudo-Voigt profiles that were statistically justified by an F-test. We found that the majority of pseudo-Voigt profiles have a predominantly Gaussian character ($m < 0.1$), suggesting that most absorptions are in the inhomogeneous line-broadening limit (Supporting Information). Throughout the following discussion, we will refer to pseudo-Voigt profiles with $m < 0.1$ as Gaussians.

For the proteins with deuterated Ala, Met, Leu, and Val residues, we analyzed the symmetric and asymmetric absorptions originating from the deuterated methyl group(s) (Figure 2). *A priori*, (d_3)Ala172 and (d_3)Met181, which have a single CD_3 group, are expected to show a symmetric and an asymmetric absorption band, originating from a single symmetric stretch and (nearly) degenerate asymmetric stretches, respectively.[14] For (d_3)Ala172 in the unfolded state, the symmetric stretch was too weak to be reproducibly characterized. A single asymmetric absorption band was observed that was well fit by a single Gaussian, suggesting that the two asymmetric stretches are degenerate. In the folded state, the symmetric band was observable and well fit by a single Gaussian. However, the asymmetric band is clearly doubled and is well fit with a Gaussian and a pseudo-Voigt profile ($m = 0.14$) of similar amplitudes and line widths, but center frequencies that are different by ~8 cm^{-1}. This doubling of the asymmetric band could result from conformational heterogeneity in the folded protein; however, it might also result from an electric field that lifts the degeneracy of the asymmetric vibrations.[55,56] Interestingly, the oscillator strength of the asymmetric stretch is ~3-fold increased in the folded compared to the unfolded state, which in analogy to surface enhancement effects is consistent with a significant local electric field in the folded protein.[57]

For (d_3)Met181 in the unfolded state, the symmetric band was well fit by two Gaussians with relative integrated intensities of 17:1, while the asymmetric band was fit by a single Gaussian, again revealing degeneracy between the asymmetric stretches. In the folded state, (d_3)Met181 shows a symmetric absorption band that was well fit by a single Gaussian that has approximately the same line width as the major symmetric Gaussian in the unfolded state but is somewhat red-shifted. In contrast to the unfolded state, fitting the asymmetric absorption band in the folded state required two Gaussians with a ratio of integrated intensities of ~3:1. The two asymmetric absorptions in the folded state might again result either from conformational heterogeneity or from an electrostatic field within the folded protein that lifts the degeneracy of the asymmetric stretches.

The two prochiral CD_3 groups of Leu and Val are expected to show a pair of symmetric bands and a pair of asymmetric bands, arising from a pair of (nearly) degenerate symmetric stretches and two pairs of (nearly) degenerate asymmetric stretches, respectively.[16] In the unfolded state, the spectra of each deuterated Leu are identical, as are the spectra of each deuterated Val, regardless of their position within the polypeptide. In each case, a single symmetric band was observed that is well fit by two Gaussians, as expected. Each asymmetric band was also well fit by two Gaussians, likely reflecting degeneracy between the pairs of asymmetric stretches of each methyl group. While the detailed Gaussian deconvolutions differ from each other, the overall line shapes are nearly identical. This may be due to the fact that the asymmetric absorption band, which is in fact a superposition of four (nearly) degenerate C–D stretches, was fit with only two Gaussians, which may make the deconvolution sensitive to noise and thus prone to fit artifacts.

In sharp contrast, in the folded state, the spectra of the Leu- and Val-labeled proteins are distinctly dependent upon their position within the polypeptide sequence. The symmetric bands are generally red-shifted relative to the unfolded protein, and required two Gaussians for fitting, except for (d_7)Leu159, which required a single pseudo-Voigt profile ($m = 0.11$). The

asymmetric bands of each Leu- and Val-labeled folded protein are generally also red-shifted relative to the unfolded state and, except for (d_7)Leu159, required two Gaussians for fitting, again likely reflecting degeneracy between the asymmetric stretches of each methyl group. While the number of functions used to fit the asymmetric absorptions in the folded and unfolded states were generally similar, the spectra at several sites are clearly more complex within the folded protein, being composed of more unique absorptions than in the unfolded state (in particular, at Leu159 and Val184, Figure 2). The additional absorptions in the folded state may again arise from conformational heterogeneity or from an electric field that lifts the degeneracy of the individual C−D absorptions. However, the absorption line widths are generally more narrow in the folded state, suggesting a more rigid and conformationally homogeneous environment, as has been observed for other proteins.[13]

The spectra of (d_8)Phe143 and (d_8)Phe153 are shown in Figure 3A,B. In the unfolded state, both (d_8)Phe143 and

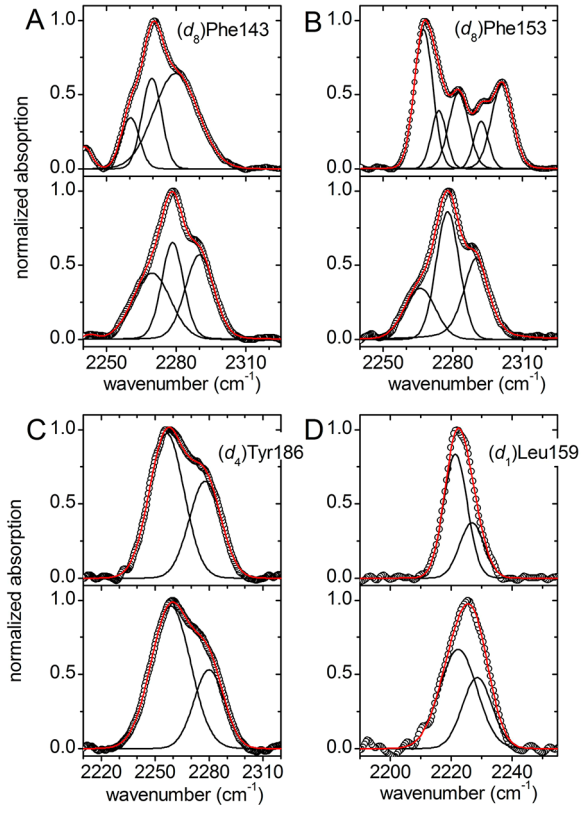


Figure 3. Low (top panels) and high temperature (bottom panels) spectra of (A) (d_8)Phe143, (B) (d_8)Phe153, (C) (d_4)Tyr186, and (D) (d_1)Leu159. Spectra are shown as circles, while individual pseudo-Voigt profiles and their sum are shown as black and red lines, respectively.

(d_8)Phe153 give rise to absorption bands that are similar to each other and to the free amino acid,[58] revealing that five aryl C−D stretches give rise to only three observable absorption bands due to degeneracy. These required one Gaussian and two pseudo-Voigt profiles (m = 0.14 to 0.53) for adequate fitting. In the folded state, (d_8)Phe143 shows a similar absorption that required three Gaussians for fitting; however, relative to the unfolded state the band is red-shifted by 9 cm^{-1}. In contrast, in the folded state, (d_8)Phe153 shows a very broad absorption band that required five Gaussians for fitting. The red-shifted absorption features are similar to the spectrum of (d_8)Phe143, while the remaining absorptions are significantly blue-shifted relative to the unfolded protein. The additional bands in the folded state may again arise from conformational heterogeneity or from an electric field that lifts the degeneracy of the individual C−D stretches. However, it is interesting to note that the increased number of absorptions in the folded state equals the number of C−D bonds present, consistent with their arising from the lifting of degeneracies.

The spectra of (d_4)Tyr186 are shown in Figure 3C. Both in the unfolded and the folded state, the spectra were well fit with two Gaussians with similar relative integrated intensities of ∼2:1. Moreover, in this case there were no significant differences in the spectra of the folded and unfolded proteins, which is consistent with the residue's solvent exposed position.[59]

Finally, we characterized the single backbone C_α−D absorption of (d_1)Leu159. In both the unfolded and folded states, (d_1)Leu159 shows a single absorption band, again somewhat red-shifted in the folded state, that required two Gaussians for adequate fitting (Figure 3D). Clearly, the single C_α−D bond experiences at least two unique environments in both the folded and unfolded states of the protein. While the unique backbone environments could result from the presence of multiple side-chain conformations, neither the number of Gaussians nor their relative intensities match those of the side-chain deuterated (d_7)Leu159 absorptions. Thus, we conclude that the two distinct environments of the C_α−D bond likely arise from backbone heterogeneity.

Thermally Induced Equilibrium Unfolding of nSH3. To characterize the thermally induced unfolding mechanism of nSH3 and to more rigorously address the potential contributions of electrostatic and conformational heterogeneity to the C−D absorptions of the folded protein, we characterized the absorption spectra as a function of temperature. The absorptions of the corresponding free, deuterated amino acids show no significant dependence on temperature (Supporting Information). Spectra of 10 mM samples of each protein were collected at 5 °C intervals between 24 and 93 °C. After heating and returning to 24 °C, the spectra and thermal melting curves were virtually identical to those acquired before heating, suggesting that the unfolding is reversible (Supporting Information). The spectra of each protein at all intermediate temperatures were well fit by a superposition of the low temperature (folded) and high temperature (unfolded) spectra, and the resulting fractional concentrations fit well to two-state transitions (Figure 4). All data were thus fit to two-state models, and the residue-specific T_m, $\Delta H°$, $\Delta S°$, and $\Delta G°$ of unfolding were determined (Table 1). Clearly the individual residues probed undergo a common thermally induced transition. The only exceptions are (d_8)Phe143, which appears to undergo an unfolding transition at higher temperature that we attribute to it being insensitive to the unfolding transition but sensitive to temperature-induced changes in the unfolded state, and (d_4)Tyr186, which does not show a transition because it is solvent exposed in both the folded and unfolded states. Thus, the data reveal that nSH3 unfolds via a single cooperative transition with T_m = 55.6 ± 0.1 °C, $\Delta H°$ = 37.2 ± 0.5 kcal mol^{-1}, and $\Delta S°$ = 113 ± 2 cal mol^{-1} K^{-1}, resulting in $\Delta G°$ = 3.4 ± 0.1 kcal mol^{-1} at 25 °C.

As other SH3 domains have been shown to unfold via a concentration-dependent midpoint, which has been taken as

dx.doi.org/10.1021/jp402772x | J. Phys. Chem. B 2013, 117, 13082−13089

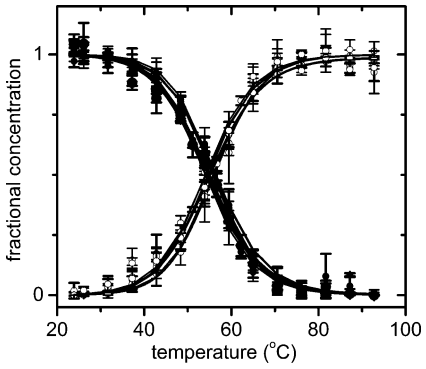


Figure 4. Overlay of the titration curves for all deuterated proteins except (d_8)Phe143 and (d_4)Tyr186 (see text). In each case, the strongest absorption band was analyzed. Solid symbols refer to the folded protein and open symbols to the unfolded protein; lines are fits to residue-specific two-state transitions (see Table 1).

Table 1. Thermodynamic Fit Parameters

	T_m (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal/mol/K)	$\Delta G°$(25 °C) (kcal/mol)
		10 mM		
(d_1)Leu159	56.3 ± 0.8	37.8 ± 5.5	115 ± 18	3.6 ± 0.6
(d_3)Ala172	54.4 ± 0.8	34.4 ± 2.4	105 ± 9	3.1 ± 0.3
(d_3)Met181	55.2 ± 0.8	34.6 ± 2.7	105 ± 10	3.2 ± 0.3
(d_4)Tyr186[a]	*a*	*a*	*a*	*a*
(d_7)Leu151	54.9 ± 0.7	40.3 ± 3.5	123 ± 12	3.7 ± 0.4
(d_7)Leu159	55.8 ± 0.1	40.0 ± 3.9	122 ± 12	3.7 ± 0.4
(d_8)Val137	54.2 ± 0.4	36.6 ± 1.1	112 ± 4	3.3 ± 0.1
(d_8)Val184	56.5 ± 1.5	35.4 ± 3.1	107 ± 12	3.4 ± 0.4
(d_8)Val187	54.2 ± 1.1	37.1 ± 0.7	113 ± 4	3.3 ± 0.1
(d_8)Phe143	60.0 ± 1.5	44.8 ± 2.8	135 ± 12	4.7 ± 0.4
(d_8)Phe153	54.6 ± 0.5	38.8 ± 2.5	118 ± 9	3.5 ± 0.3
		2 mM		
(d_3)Ala172	54.5 ± 0.2	37.7 ± 4.2	115 ± 13	3.4 ± 0.4
(d_7)Leu151	54.6 ± 0.9	39.9 ± 4.6	122 ± 16	3.6 ± 0.5
(d_7)Leu159	54.1 ± 0.4	38.2 ± 2.7	117 ± 9	3.4 ± 0.3
(d_8)Val137	55.2 ± 0.6	35.9 ± 4.0	109 ± 13	3.3 ± 0.4
(d_8)Val184	56.9 ± 1.8	35.8 ± 5.2	109 ± 19	3.5 ± 0.6
(d_8)Val187	55.1 ± 0.4	39.4 ± 2.3	120 ± 8	3.6 ± 0.2

[a]No transition observed (see text).

evidence of a more complex unfolding process,[41,42] we also examined the thermal denaturation of a subset of six of the labeled proteins at 5-fold lower concentration (2 mM) (Table 1 and Supporting Information). The data clearly reveal similar two-state transitions that are thermodynamically identical across all sites probed and identical to the same sites probed at higher protein concentrations. Thus, we conclude that nSH3 thermally unfolds via a simple two-state, cooperative mechanism without intermolecular association or the intervention of additional conformational changes. Considering the number of specific sites characterized, and the high time resolution of the IR probes, this is to our knowledge the most conclusive evidence reported to date for the two-state cooperative equilibrium (un)folding of a protein.

Importantly, with the exception of (d_1)Leu159, where weak intensity and significant spectral overlap of the two C−D absorption features unfortunately precluded further analysis, in all cases where multiple Gaussian or pseudo-Voigt functions were needed to fit the folded state spectrum, allowing the individual functions to change amplitude independently and fitting the resulting fractional concentrations to a three-state model did not lead to a statistically significant improvement of the fit at either 2 mM or 10 mM (Supporting Information). If conformational heterogeneity were the origin of the spectral complexity within the folded state, the relative population of the conformations would have to be independent of temperature, which seems unlikely. Thus, we conclude that the spectral complexity in the folded state of the side-chain deuterated proteins results in large part from electrostatic effects.

CONCLUSIONS

While we detected conformational heterogeneity in the backbone of nSH3 at Leu159, which adds to that previously detected at Gly180,[15] the conformational heterogeneity within nSH3 appears to be much less pronounced than in several other SH3 domains.[36,37,50,51] Instead, the protein appears to fold into a rigid structure that only samples different conformations when sufficiently destabilized to undergo a single, global two-state transition to the unfolded state. However, despite its relatively small size, the folded protein provides a wide range of electrostatic microenvironments, with folding inducing site-dependent shifts in C−D absorption frequencies of up to 10 cm^{-1} to the red and 14 cm^{-1} to the blue. Moreover, at several positions the degeneracies of the C−D absorptions of a given side chain are lifted in the folded protein, resulting in a significant level of spectral complexity. This likely results from different Stark shifts of the C−D absorptions due to varying orientation of their transition dipole moments relative to the direction of the local electric field. Interestingly, the data thus provide multiple experimental constraints that should prove useful for characterizing electric fields within proteins. When combined with calculations that establish the relative orientations of the electrostatic fields and the C−D bonds, site-selective labeling of proteins with C−D bonds should allow not only for the detection of electrostatic heterogeneity but also for its perturbation-free quantification.

ASSOCIATED CONTENT

Supporting Information

Experimental details and peptide purification; data analysis; additional tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: floyd@scripps.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This article is dedicated to Prof. Peter G. Wolynes with gratitude for the inspiration and vision he continues to provide to the field of protein folding. This work was supported by the National Science Foundation under Grant No. MCB 034697.

REFERENCES

(1) Zimmermann, J.; Oakman, E. L.; Thorpe, I. F.; Shi, X.; Abbyad, P.; Brooks, C. L., III; Boxer, S. G.; Romesberg, F. E. Antibody Evolution Constrains Conformational Heterogeneity by Tailoring Protein Dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13722−13727.

(2) Conant, G. C.; Wolfe, K. H. Turning a Hobby into a Job: How Duplicated Genes Find New Functions. *Nat. Rev. Genet.* **2008**, *9*, 938−950.
(3) Tokuriki, N.; Tawfik, D. S. Stability Effects of Mutations and Protein Evolvability. *Curr. Opin. Struct. Biol.* **2009**, *19*, 596−604.
(4) Fraser, J. S.; van den Bedem, H.; Samelson, A. J.; Lang, P. T.; Holton, J. M.; Echols, N.; Alber, T. Accessing Protein Conformational Ensembles Using Room-Temperature X-Ray Crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16247−16252.
(5) Bertini, I.; Calderone, V.; Cosenza, M.; Fragai, M.; Lee, Y. M.; Luchinat, C.; Mangani, S.; Terni, B.; Turano, P. Conformational Variability of Matrix Metalloproteinases: Beyond a Single 3D Structure. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5334−5339.
(6) Ding, K.; Louis, J. M.; Gronenborn, A. M. Insights into Conformation and Dynamics of Protein Gb1 During Folding and Unfolding by NMR. *J. Mol. Biol.* **2004**, *335*, 1299−1307.
(7) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. The Dynamic Energy Landscape of Dihydrofolate Reductase Catalysis. *Science* **2006**, *313*, 1638−1642.
(8) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106*, 3210−3235.
(9) Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer, S. G. Electric Fields at the Active Site of an Enzyme: Direct Comparison of Experiment with Theory. *Science* **2006**, *313*, 200−204.
(10) Romesberg, F. E. Multidisciplinary Experimental Approaches to Characterizing Biomolecular Dynamics. *ChemBioChem* **2003**, *4*, 563−571.
(11) Chin, J. K.; Jimenez, R.; Romesberg, F. E. Direct Observation of Protein Vibrations by Selective Incorporation of Spectroscopically Observable Carbon-Deuterium Bonds in Cytochrome *c*. *J. Am. Chem. Soc.* **2001**, *123*, 2426−2427.
(12) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. A High-Resolution Probe of Protein Folding. *J. Am. Chem. Soc.* **2004**, *126*, 3384−3385.
(13) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. Direct and High Resolution Characterization of Cytochrome *c* Equilibrium Folding. *J. Am. Chem. Soc.* **2006**, *128*, 14232−14233.
(14) Sagle, L. B.; Zimmermann, J.; Matsuda, S.; Dawson, P. E.; Romesberg, F. E. Redox-Coupled Dynamics and Folding in Cytochrome *c*. *J. Am. Chem. Soc.* **2006**, *128*, 7909−7915.
(15) Cremeens, M. E.; Zimmermann, J.; Yu, W.; Dawson, P. E.; Romesberg, F. E. Direct Observation of Structural Heterogeneity in a *β*-Sheet. *J. Am. Chem. Soc.* **2009**, *131*, 5726−5727.
(16) Zimmermann, J.; Gundogdu, K.; Cremeens, M. E.; Bandaria, J. N.; Hwang, G. T.; Thielges, M. C.; Cheatum, C. M.; Romesberg, F. E. Efforts toward Developing Probes of Protein Dynamics: Vibrational Dephasing and Relaxation of Carbon-Deuterium Stretching Modes in Deuterated Leucine. *J. Phys. Chem. B* **2009**, *113*, 7991−7994.
(17) Zimmermann, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E. Cyano Groups as Probes of Protein Microenvironments and Dynamics. *Angew. Chem., Int. Ed.* **2011**, *50*, 8333−8337.
(18) Groff, D.; Thielges, M. C.; Cellitti, S.; Schultz, P. G.; Romesberg, F. E. Efforts toward the Direct Experimental Characterization of Enzyme Microenvironments: Tyrosine100 in Dihydrofolate Reductase. *Angew. Chem., Int. Ed.* **2009**, *48*, 3478−3481.
(19) Thielges, M. C.; Case, D. A.; Romesberg, F. E. Carbon-Deuterium Bonds as Probes of Dihydrofolate Reductase. *J. Am. Chem. Soc.* **2008**, *130*, 6597−6603.
(20) Pawson, T.; Nash, P. Assembly of Cell Regulatory Systems through Protein Interaction Domains. *Science* **2003**, *300*, 445−452.
(21) Nguyen, J. T.; Turck, C. W.; Cohen, F. E.; Zuckermann, R. N.; Lim, W. A. Exploiting the Basis of Proline Recognition by SH3 and WW Domains: Design of N-Substituted Inhibitors. *Science* **1998**, *282*, 2088−2092.
(22) Scott, J. D.; Pawson, T. Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart. *Science* **2009**, *326*, 1220−1224.
(23) Gmeiner, W. H.; Horita, D. A. Implications of SH3 Domain Structure and Dynamics for Protein Regulation and Drug Design. *Cell Biochem. Biophys.* **2001**, *35*, 127−140.
(24) Kobashigawa, Y.; Sakai, M.; Naito, M.; Yokochi, M.; Kumeta, H.; Makino, Y.; Ogura, K.; Tanaka, S.; Inagaki, F. Structural Basis for the Transforming Activity of Human Cancer-Related Signaling Adaptor Protein Crk. *Nat. Struct. Mol. Biol.* **2007**, *14*, 503−510.
(25) Zhang, O.; Forman-Kay, J. D. Structural Characterization of Folded and Unfolded States of an SH3 Domain in Equilibrium in Aqueous Buffer. *Biochemistry* **1995**, *34*, 6784−6794.
(26) Hansson, H.; Mattsson, P. T.; Allard, P.; Haapaniemi, P.; Vihinen, M.; Smith, C. I.; Hard, T. Solution Structure of the SH3 Domain from Bruton's Tyrosine Kinase. *Biochemistry* **1998**, *37*, 2912−2924.
(27) Ferreon, J. C.; Volk, D. E.; Luxon, B. A.; Gorenstein, D. G.; Hilser, V. J. Solution Structure, Dynamics, and Thermodynamics of the Native State Ensemble of the Sem-5 C-Terminal SH3 Domain. *Biochemistry* **2003**, *42*, 5582−5591.
(28) Stoll, R.; Renner, C.; Buettner, R.; Voelter, W.; Bosserhoff, A. K.; Holak, T. A. Backbone Dynamics of the Human Mia Protein Studied by (15)N NMR Relaxation: Implications for Extended Interactions of SH3 Domains. *Protein Sci.* **2003**, *12*, 510−519.
(29) Sadqi, M.; Casares, S.; Abril, M. A.; Lopez-Mayorga, O.; Conejero-Lara, F.; Freire, E. The Native State Conformational Ensemble of the SH3 Domain from Alpha-Spectrin. *Biochemistry* **1999**, *38*, 8899−8906.
(30) Wales, T. E.; Engen, J. R. Partial Unfolding of Diverse SH3 Domains on a Wide Timescale. *J. Mol. Biol.* **2006**, *357*, 1592−1604.
(31) Crowhurst, K. A.; Tollinger, M.; Forman-Kay, J. D. Cooperative Interactions and a Non-Native Buried Trp in the Unfolded State of an SH3 Domain. *J. Mol. Biol.* **2002**, *322*, 163−178.
(32) Goudreau, N.; Cornille, F.; Duchesne, M.; Parker, F.; Tocque, B.; Garbay, C.; Roques, B. P. NMR Structure of the N-Terminal SH3 Domain of Grb2 and Its Complex with a Proline-Rich Peptide from Sos. *Nat. Struct. Biol.* **1994**, *1*, 898−907.
(33) Dyson, H. J.; Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev.* **2004**, *104*, 3607−3622.
(34) Viguera, A. R.; Martinez, J. C.; Filimonov, V. V.; Mateo, P. L.; Serrano, L. Thermodynamic and Kinetic Analysis of the SH3 Domain of Spectrin Shows a Two-State Folding Transition. *Biochemistry* **1994**, *33*, 2142−2150.
(35) Grantcharova, V. P.; Baker, D. Folding Dynamics of the Src SH3 Domain. *Biochemistry* **1997**, *36*, 15685−15692.
(36) Camarero, J. A.; Fushman, D.; Sato, S.; Giriat, I.; Cowburn, D.; Raleigh, D. P.; Muir, T. W. Rescuing a Destabilized Protein Fold through Backbone Cyclization. *J. Mol. Biol.* **2001**, *308*, 1045−1062.
(37) Muralidharan, V.; Cho, J.; Trester-Zedlitz, M.; Kowalik, L.; Chait, B. T.; Raleigh, D. P.; Muir, T. W. Domain-Specific Incorporation of Noninvasive Optical Probes into Recombinant Proteins. *J. Am. Chem. Soc.* **2004**, *126*, 14004−14012.
(38) Muralidharan, V.; Dutta, K.; Cho, J.; Vila-Perello, M.; Raleigh, D. P.; Cowburn, D.; Muir, T. W. Solution Structure and Folding Characteristics of the C-Terminal SH3 Domain of c-Crk-II. *Biochemistry* **2006**, *45*, 8874−8884.
(39) Di Nardo, A. A.; Larson, S. M.; Davidson, A. R. The Relationship between Conservation, Thermodynamic Stability, and Function in the SH3 Domain Hydrophobic Core. *J. Mol. Biol.* **2003**, *333*, 641−655.
(40) Knapp, S.; Mattson, P. T.; Christova, P.; Berndt, K. D.; Karshikoff, A.; Vihinen, M.; Smith, C. I.; Ladenstein, R. Thermal Unfolding of Small Proteins with SH3 Domain Folding Pattern. *Proteins* **1998**, *31*, 309−319.
(41) Casares, S.; Sadqi, M.; Lopez-Mayorga, O.; Conejero-Lara, F.; van Nuland, N. A. Detection and Characterization of Partially Unfolded Oligomers of the SH3 Domain of Alpha-Spectrin. *Biophys. J.* **2004**, *86*, 2403−2413.
(42) Petzold, K.; Ohman, A.; Backman, L. Folding of the AlphaII-Spectrin SH3 Domain under Physiological Salt Conditions. *Arch. Biochem. Biophys.* **2008**, *474*, 39−47.

(43) Korzhnev, D. M.; Salvatella, X.; Vendruscolo, M.; Di Nardo, A. A.; Davidson, A. R.; Dobson, C. M.; Kay, L. E. Low-Populated Folding Intermediates of Fyn SH3 Characterized by Relaxation Dispersion NMR. *Nature* **2004**, *430*, 586−590.
(44) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundstrom, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an Intermediate State in Protein Folding and Aggregation. *Science* **2012**, *336*, 362−366.
(45) Neudecker, P.; Zarrine-Afsar, A.; Choy, W. Y.; Muhandiram, D. R.; Davidson, A. R.; Kay, L. E. Identification of a Collapsed Intermediate with Non-Native Long-Range Interactions on the Folding Pathway of a Pair of Fyn SH3 Domain Mutants by NMR Relaxation Dispersion Spectroscopy. *J. Mol. Biol.* **2006**, *363*, 958−976.
(46) Korzhnev, D. M.; Neudecker, P.; Zarrine-Afsar, A.; Davidson, A. R.; Kay, L. E. Abp1p and Fyn SH3 Domains Fold through Similar Low-Populated Intermediate States. *Biochemistry* **2006**, *45*, 10175−10183.
(47) Dasgupta, A.; Udgaonkar, J. B. Four-State Folding of a SH3 Domain: Salt-Induced Modulation of the Stabilities of the Intermediates and Native State. *Biochemistry* **2012**, *51*, 4723−4734.
(48) Wani, A. H.; Udgaonkar, J. B. Native State Dynamics Drive the Unfolding of the SH3 Domain of PI3 Kinase at High Denaturant Concentration. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20711−20716.
(49) Wani, A. H.; Udgaonkar, J. B. Revealing a Concealed Intermediate That Forms after the Rate-Limiting Step of Refolding of the SH3 Domain of PI3 Kinase. *J. Mol. Biol.* **2009**, *387*, 348−362.
(50) Anafi, M.; Rosen, M. K.; Gish, G. D.; Kay, L. E.; Pawson, T. A Potential Sh3 Domain-Binding Site in the Crk SH2 Domain. *J. Biol. Chem.* **1996**, *271*, 21365−21374.
(51) Camarero, J. A.; Fushman, D.; Cowburn, D.; Muir, T. W. Peptide Chemical Ligation inside Living Cells: In Vivo Generation of a Circular Protein Domain. *Bioorg. Med. Chem.* **2001**, *9*, 2479−2484.
(52) Dawson, P. E.; Fitzgerald, M. C.; Muir, T. W.; Kent, S. B. H. Methods for the Chemical Synthesis and Readout of Self-Encoded Arrays of Polypeptide Analogues. *J. Am. Chem. Soc.* **1997**, *119*, 7917−7927.
(53) Weinkam, P.; Zimmermann, J.; Sagle, L. B.; Matsuda, S.; Dawson, P. E.; Wolynes, P. G.; Romesberg, F. E. Characterization of Alkaline Transitions in Ferricytochrome *c* Using Carbon-Deuterium Infrared Probes. *Biochemistry* **2008**, *47*, 13470−13480.
(54) Yu, W.; Dawson, P. E.; Zimmermann, J.; Romesberg, F. E. Carbon-Deuterium Bonds as Probes of Protein Thermal Unfolding. *J. Phys. Chem. B* **2012**, *116*, 6397−6403.
(55) Kahn, R.; Cohen De Lara, E.; Möller, K. D. Effect of an Electric Field on a Methane Molecule. II. Calculation of the Degeneracy Splitting of the ν_3 Band. Expression of the Second Derivatives of the CH_4 Dipole Moment and Evaluation of the Second Derivative of the C−H Bond Polarizability. *J. Chem. Phys.* **1985**, *83*, 2653−2561.
(56) Cohen De Lara, E.; Kahn, R.; Seloudoux, R. Effect of an Electric Field on a Methane Molecule. I. Infrared Analysis of Methane (CH_4−CD_4) Adsorbed in NaA Zeolite in the Temperature Range 150−20 K. *J. Chem. Phys.* **1985**, *83*, 2646−2652.
(57) Osawa, M. Near-Field Optics and Surface Plasmon Polaritons. *Top. Appl. Phys.* **2001**, *81*, 163−187.
(58) Hickert, A. S.; Durgan, A. C.; Patton, D. A.; Blake, S. A.; Cremeens, M. E. A B3LYP Investigation of the Conformational and Environmental Sensitivity of Carbon−Deuterium Frequencies of Aryl-Perdeuterated Phenylalanine and Tryptophan. *Theor. Chem. Acc.* **2011**, *130*, 883−889.
(59) Wu, X.; Knudsen, B.; Feller, S. M.; Zheng, J.; Sali, A.; Cowburn, D.; Hanafusa, H.; Kuriyan, J. Structural Basis for the Specific Interaction of Lysine-Containing Proline-Rich Peptides with the N-Terminal SH3 Domain of c-Crk. *Structure* **1995**, *3*, 215−226.