J|A|C|S
COMMUNICATIONS

Published on Web 04/07/2009

Direct Observation of Structural Heterogeneity in a β-Sheet

Matthew E. Cremeens, Jörg Zimmermann, Wayne Yu, Philip E. Dawson, and Floyd E. Romesberg*

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

Received January 22, 2009; E-mail: floyd@scripps.edu

Structural heterogeneity, or intrinsic disorder, is increasingly thought to be inherent in many proteins, and its presence has important ramifications for biological function.[1−3] However, it is difficult to detect by conventional methods. For example, crystallization may be incompatible with heterogeneous proteins[3] or favor one conformation, and NMR spectroscopy may lack sufficient time resolution. In contrast, it is straightforward to detect discrete species, even those interconverting on the picosecond time scale, by IR spectroscopy, and IR absorptions are well-known to be sensitive to structure.[4−6] Although the use of IR spectroscopy to study protein dynamics has traditionally been complicated by the many overlapping vibrations inherent in any protein, our recent work has demonstrated that this limitation may be overcome by the use of site-selectively incorporated carbon−deuterium (C−D) bonds.[7−10] C−D bonds are nonperturbative, and their stretching absorption frequencies fall within a region of the IR spectrum that is free of other protein absorptions; this not only facilitates detection but also simplifies interpretation, as they are adiabatically decoupled from other vibrations.

Our efforts to date have focused on the characterization of protein side chains,[7−10] but backbone dynamics are expected to be an important source of the structural heterogeneity. Indeed, previous calculations and model dipeptide studies suggest that the C_α−D signals should be sensitive to local backbone structure.[4,5] Thus, we chose to explore the use of $C_\alpha D_2$ backbone-deuterated glycine [(d_2)Gly] as a probe of the dynamics and heterogeneity of the N-terminal Src homology 3 (nSH3) domain from the human CrkII adaptor protein (Figure 1).[11] SH3 domains are regulatory elements that mediate diverse signaling pathways via the recognition of a network of different proteins, which is likely to require some degree of conformational heterogeneity.[2,12]

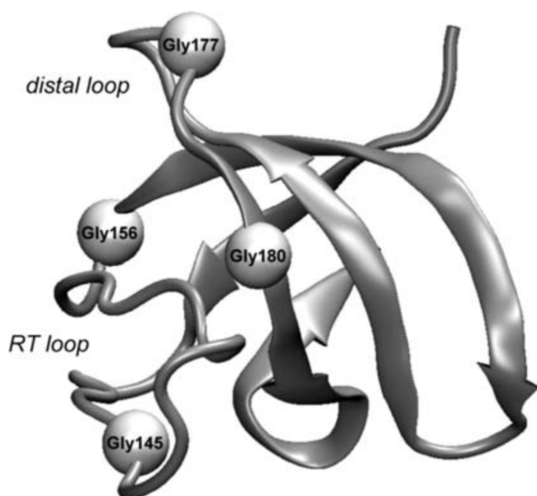

Figure 1. Structure of nSH3 (PDB entry 1CKA) with Gly residues indicated. The bound CG3 peptide has been omitted for clarity.

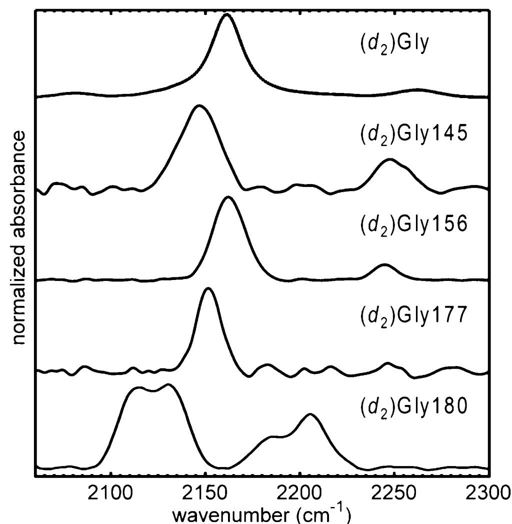

Figure 2. C−D IR spectra of (d_2)Gly and the deuterated SH3 variants.

We first predicted the dependence of the $C_\alpha D_2$ stretching frequencies on backbone conformation using density functional theory (DFT) calculations on *N*-formyl glycine amide [Fm−(d_2)Gly−NH_2] at the B3LYP/6-31G(d) level of theory [see the Supporting Information (SI)]. Three stable minima, corresponding to an extended β-strand, a γ-turn, and an α-helix-like turn, were identified. The calculations predicted a strong dependence of the $C_\alpha D_2$ IR absorptions on the Φ and Ψ angles, with symmetric/asymmetric absorption frequencies predicted to be 2116/2185, 2140/2270, and 2155/2252 cm^{-1} for the β-strand, γ-turn, and α-helix-like turns, respectively. Of particular significance is the predicted red shift for the β-strand-like conformation relative to the turn conformations.

We next characterized the free amino acid (d_2)Gly in phosphate buffer (Figure 2). The IR spectrum is composed of two absorption bands, each of which is well-fit by a quasi-Voigt function (see the SI). On the basis of the calculations, the absorption with a center frequency of 2161.1 ± 0.3 cm^{-1} and line width (fwhm) of 20.4 ± 0.3 cm^{-1} is assigned to the $C_\alpha D_2$ symmetric stretching vibration and the absorption at 2261.3 ± 0.3 cm^{-1} with a line width of 24.7 ± 9.0 cm^{-1} to the $C_\alpha D_2$ asymmetric stretching vibration.

We then synthesized and characterized four deuterium-labeled variants of nSH3 bound to the peptide ligand CG3 (Figure 2 and the SI). At position 145 or 156 [(d_2)Gly145 or (d_2)Gly156], which are both in the nSH3 RT loop (Figure 1), (d_2)Gly shows two absorptions, each well-fit by single Gaussians with center frequencies/line widths of 2146.7 ± 0.9 cm^{-1}/22.9 ± 7.1 cm^{-1} and 2248.6 ± 1.3 cm^{-1}/18.7 ± 6.2 cm^{-1} for (d_2)Gly145 and 2162.3 ± 1.6 cm^{-1}/19.4 ± 5.4 cm^{-1} and 2243.4 ± 1.6 cm^{-1}/13.1 ± 3.8 cm^{-1} for (d_2)Gly156. These absorptions are assigned to the $C_\alpha D_2$ symmetric and asymmetric stretches, respectively. For (d_2)Gly177, which is

10.1021/ja900505e CCC: $40.75
© 2009 American Chemical Society

located in the nSH3 distal loop (Figure 1), we observed a single absorption that was well-fit by a single Gaussian with a center frequency/line width of 2151.7 ± 1.8 cm^{-1}/15.4 ± 2.8 cm^{-1} and is assigned to the $C_\alpha D_2$ symmetric stretch (the asymmetric-stretch absorption was too weak to detect).

We also examined (d_2)Gly180, which is located in the antiparallel β-stranded core (Figure 1). The spectrum of (d_2)Gly180 is composed of two sets of two overlapping absorptions, each well-fit by a single Gaussian. The more red-shifted pair of absorptions (with center frequencies/line widths of 2112.4 ± 1.0 cm^{-1}/19.4 ± 4.2 cm^{-1} and 2132.4 ± 0.5 cm^{-1}/20.1 ± 2.4 cm^{-1}) are assigned to overlapping $C_\alpha D_2$ symmetric stretches and the less red-shifted pair (with center frequencies/line widths of 2183.3 ± 0.4 cm^{-1}/18.9 ± 3.6 cm^{-1} and 2206.1 ± 0.8 cm^{-1}/21.3 ± 2.8 cm^{-1}) to overlapping asymmetric stretches. While the data do not allow for an unambiguous assignment of which symmetric-stretch absorption is associated with which asymmetric-stretch absorption, assigning the 2132.4 and 2183.3 cm^{-1} absorptions to one symmetric/asymmetric pair and the 2112.4 and 2206.1 cm^{-1} absorptions to the other is perhaps most reasonable, as the average frequencies are conserved with only the splittings being different. Regardless, as predicted for an extended backbone conformation relative to a turn conformation, both sets of absorptions are significantly red-shifted relative to the absorptions of the loop residues.

Under denaturing conditions (6 M guanidine hydrochloride), the spectra of the four proteins are similar to each other and to that of the free amino acid, but they are not identical (see the SI). Each consists of a single absorption band that is well-fit with a single Gaussian and assigned to a symmetric $C_\alpha D_2$ stretch. The frequencies/line widths are 2164.9 ± 1.0 cm^{-1}/38.3 ± 11.7 cm^{-1} for (d_2)Gly145, 2160.5 ± 2.9 cm^{-1}/33.7 ± 6.1 cm^{-1} for (d_2)Gly156, 2163.6 ± 4.3 cm^{-1}/37.2 ± 12.8 cm^{-1} for (d_2)Gly177, and 2161.7 ± 3.5 cm^{-1}/38.4 ± 7.1 cm^{-1} for (d_2)Gly180. Interestingly, the consistently larger line widths than observed under native conditions suggest increased disorder in the denatured state. It also appears that small differences in frequency persist in the denatured state, although additional studies are required to determine the significance and generality of these differences. Nonetheless, these differences suggest that unique microenvironments exist at the different backbone positions.

The crystal structure of nSH3 reveals that the carbonyl groups of the loop residues Gly145, Gly156, and Gly177 are solvent-exposed with similar Φ angles (81, 88, and 103°, respectively) but rather different Ψ angles (147, −24, and 4°, respectively).[11] In contrast, Gly180 is positioned in the antiparallel β-stranded core of the protein, is not solvent-exposed, and has Φ and Ψ angles of 174 and 176°, respectively. These data, along with the results of the calculations and previous solvent polarity studies,[10] suggest that the dramatic red shift of the (d_2)Gly180 absorptions result from the unique Φ angle associated with its position in a β-sheet. We attribute this dependence to variable orbital overlap between the C−D bonds and the adjacent amide π bonds (see the SI), which is consistent with previous conclusions that hyperconjugative effects make important contributions to conformation-dependent C−D frequency shifts.[7] Although small, the frequency differences that persist in 6 M guanidine hydrochloride suggest that even under these conditions, nSH3 is not a random coil.

In addition to its red-shifted absorptions, the absorption spectrum of (d_2)Gly180 is unique in that absorptions from two sets of symmetric and asymmetric stretches are observed. On the basis of previous studies, the red shift is too large to result from through-space interactions,[10] and neither absorption corresponds to the unfolded protein, which has been suggested to contribute to the conformational heterogeneity of other SH3 domains.[12] Thus, the two absorptions likely result from the population of unique β-strand conformations. Previous NMR studies have revealed heterogeneity within SH3 loops[13−16] but not within the β-sheet structure, although it is thought to be present in other proteins.[17,18] This suggests that the conformations interconvert on a time scale that is fast on the NMR time scale but slow on the IR time scale (i.e., microseconds to nanoseconds).

The data reveal, as predicted previously,[4,5] that C_α−D probes are sensitive to their local backbone structure within a protein. We have used this sensitivity to show that the β-sheet of nSH3 adopts at least two folded conformations and that the protein is likely not a random coil when denatured by guanidine hydrochloride. Thus, the folded core of the protein may be less ordered than previously suspected, and the unfolded protein may be more ordered. It seems possible that these factors may influence both the folding and activity of the protein. Because the high temporal and structural resolution of the C_α−D bond absorptions make them well-suited to detect transient conformations as well as to help elucidate structure, this method should prove useful in the study of a wide variety of proteins in both their folded and unfolded states.

Acknowledgment. This work was supported by the National Science Foundation under Grant MCB 034697 (to F.E.R.).

Supporting Information Available: Detailed methods, additional data, and details of the DFT calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1998**, *293*, 321–331.
(2) Dunker, A. K.; Obradovic, Z. *Nat. Biotechnol.* **2001**, *19*, 805–806.
(3) Tunistra, R. L.; Peterson, F. C.; Kutlesa, S.; Elgin, E. S.; Kron, M. A.; Volkman, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5057–5062.
(4) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2007**, *111*, 5300–5303.
(5) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. B* **2004**, *108*, 10923–10924.
(6) Barth, A.; Zscherp, C. *Q. Rev. Biophys.* **2002**, *35*, 369–430.
(7) Chin, J. K.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 2426–2427.
(8) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 14232–14233.
(9) Weinkam, P.; Zimmermann, J.; Sagle, L. B.; Matsuda, S.; Dawson, P. E.; Wolynes, P. G.; Romesberg, F. E. *Biochemistry* **2008**, *47*, 13470–13480.
(10) Thielges, M. C.; Case, D. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 6597–6603.
(11) Wu, X.; Knudsen, B.; Feller, S. M.; Zheng, J.; Sali, A.; Cowburn, D.; Hanafusa, H.; Kuriyan, J. *Structure* **1995**, *3*, 215–226.
(12) Gmeiner, W. H.; Horita, D. A. *Cell. Biochem. Biophys.* **2001**, *35*, 127–140.
(13) Hansson, H.; Mattsson, P. T.; Allard, P.; Haapaniemi, P.; Vihinen, M.; Smith, C. I.; Hard, T. *Biochemistry* **1998**, *37*, 2912–2924.
(14) Ferreon, J. C.; Volk, D. E.; Luxon, B. A.; Gorenstein, D. G.; Hilser, V. J. *Biochemistry* **2003**, *42*, 5582–5591.
(15) Martín-García, J. M.; Luque, I.; Mateo, P. L.; Ruiz-Sanz, J.; Cámara-Artigas, A. *FEBS Lett.* **2007**, *581*, 1701–1706.
(16) Stoll, R.; Renner, C.; Buettner, R.; Voelter, W.; Bosserhoff, A. K.; Holak, T. A. *Protein Sci.* **2003**, *12*, 510–519.
(17) Svennsson, L. A.; Sjölin, L.; Gilliland, G. L.; Finzel, B. C.; Wlodawer, A. *Proteins: Struct., Funct., Genet.* **1986**, *1*, 370–375.
(18) Frimurer, T. M.; Peters, G. H.; Sørensen, M. D.; Led, J. J.; Olsen, O. H. *Protein Sci.* **1999**, *8*, 25–34.

JA900505E