Protein Folding

DOI: 10.1002/anie.200703376

Spin Relaxation Enhancement Confirms Dominance of Extended Conformations in Short Alanine Peptides**

Kang Chen, Zhigang Liu, Chunhui Zhou, W. Clay Bracken, and Neville R. Kallenbach**

Mounting spectroscopic evidence indicates the presence of local order in unfolded proteins, including polyproline II (PII).[1–3] The data supporting this order rely on short-chain models, while the evidence for random-coil behavior is derived from measurements on a different length scale. Several theoretical papers have addressed this so-called "reconciliation problem", showing that excluded volume and local conformational preferences can account for the seemingly discrepant descriptions.[4,5] We reexamine herein the dimensional properties of short Ala chains, which serve as general models for the protein backbone,[6–9] using paramagnetic proton spin relaxation enhancement to evaluate the intramolecular intermediate-range distances (r). Mesostate-based Monte Carlo samplings[6] were performed to interpret the results. The salient conclusion of the study is that for short Ala sequences, over 90% occupation of extended conformations, that is, PII and β, is required to reproduce the experimentally observed $\langle r^{-6}\rangle$ averaging. Compact structures including αR and αL basins as well as turns are present but cannot be dominant (the geometry of each conformation is specified in the Supporting Information). This study presents the first long-range experimental evidence for the dominance of extended conformations in short Ala models.

Intramolecular distance measurements on biomolecules have been made using fluorescence resonant energy transfer (FRET),[10,11] double-quantum-filtered (DQF) ESR spectroscopy,[12] and NMR spin relaxation enhancement.[13–15] Our approach is to lock the spin label onto the peptide backbone[16] to minimize local motions. Longitudinal (R_1) and transverse (R_2) relaxation rates were measured on H^N of ^{15}N-labeled Ala residues using heteronuclear correlation detection. This method affords accurate determination of the r^{-6} averaged distances on the nanometer scale.

Two sets of peptides were employed in these studies, OO-T*/Aib-PPPA*PPPA*-OO and OO-T*/Aib-PPPAAAA*-OO, where O is ornithine, T* is Toac (2,2,6,6-tetramethylpiperidine-1-oxyl-4-amino-4-carboxylic acid), A* is a ^{15}N-labeled Ala residue and Aib is α-amino isobutyric acid (see Scheme S1 in the Supporting Information). Toac is a spin-labeled residue tightly coupled to the main chain; it has been extensively used by the Toniolo group.[16] The corresponding Aib peptides serve as controls for relaxation enhancement. The identical ^1H–^{15}N HSQC spectra of Aib and Toac peptides ensure the validity of using relaxation-rate data to calculate electron–proton distances (Table 1 and Figure 1).

Table 1: Experimental results.

Peptide	Pair	ΔR_1 [s^{-1}]	ΔR_2 [s^{-1}]	Distance [Å][a]	X^2 [b]
$O_2T^*P_3A^*P_3A^*O_2$	T*3–A*7	2.17 ± 0.08	13.6 ± 1.6	12.3 ± 0.1	0.01
	T*3–A*11	0.636 ± 0.070	3.49 ± 0.34	15.3 ± 0.2	0.07
$O_2T^*P_3A_3A^*O_2$	T*3–A*10	1.30 ± 0.06	4.78 ± 0.27	13.9 ± 0.1	0.30

[a] Distance determined by NMR spectroscopy is calculated according to $\langle r^{-6}\rangle^{-1/6}$. Error ranges were estimated using Monte Carlo methods. [b] According to Equation (3) in the Supporting Information.

The distances determined by NMR spectroscopy (Table 1) cannot be directly explained in terms of a completely rigid PII conformation for either Ala or Pro. Because the determined distance is inverse sixth-power averaged over conformations with flexible Ala residues in the middle, a small population of extremely short distances would be weighted heavily in the final result ($\int P(r)r\mathrm{d}r$ vs. $[\int P(r)r^{-6}\mathrm{d}r]^{-1/6}$, where $P(r)$ is the probability density function). The extreme weighting problem has been recognized by Jacob et al.[15] and by Zagrovic and van Gunsteren.[17]

Our approach to this weighting problem is to simulate ensembles of peptide structures through mesostate Monte Carlo sampling.[6] The dihedral angles of Pro and Ala were derived from the coil library (see Figure 2 and Table S1 in the Supporting Information) provided by Avbelj and Baldwin.[18] Two mesostates, PII weighted at 60% and turn/αR weighted at 40%, were identified in the coil library for Pro. Four

[*] Dr. K. Chen
Laboratory of Molecular Biophysics
National Heart, Lung and Blood Institute
National Institutes of Health
50 South Drive, Bethesda, MD 20892 (USA)
Fax: (+1) 301-402-3404
E-mail: chenk2@mail.nih.gov
Z. Liu, C. Zhou, Prof. N. R. Kallenbach
Department of Chemistry
New York University
100 Washington Square East, New York, NY 10003 (USA)
Fax: (+1) 212-260-7905
E-mail: nrk1@nyu.edu
Prof. W. C. Bracken
Department of Biochemistry
Weill Medical College of Cornell University
New York, NY 10021 (USA)

[**] The work was supported by a grant from ONR. The NMR data was collected on the spectrometer supported by the facility grant from NCRR/NIH (C06 RR-16572) to NYU. We thank Franc Avbelj and Buzz Baldwin for providing us the coil library data and George Rose for helpful discussion.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

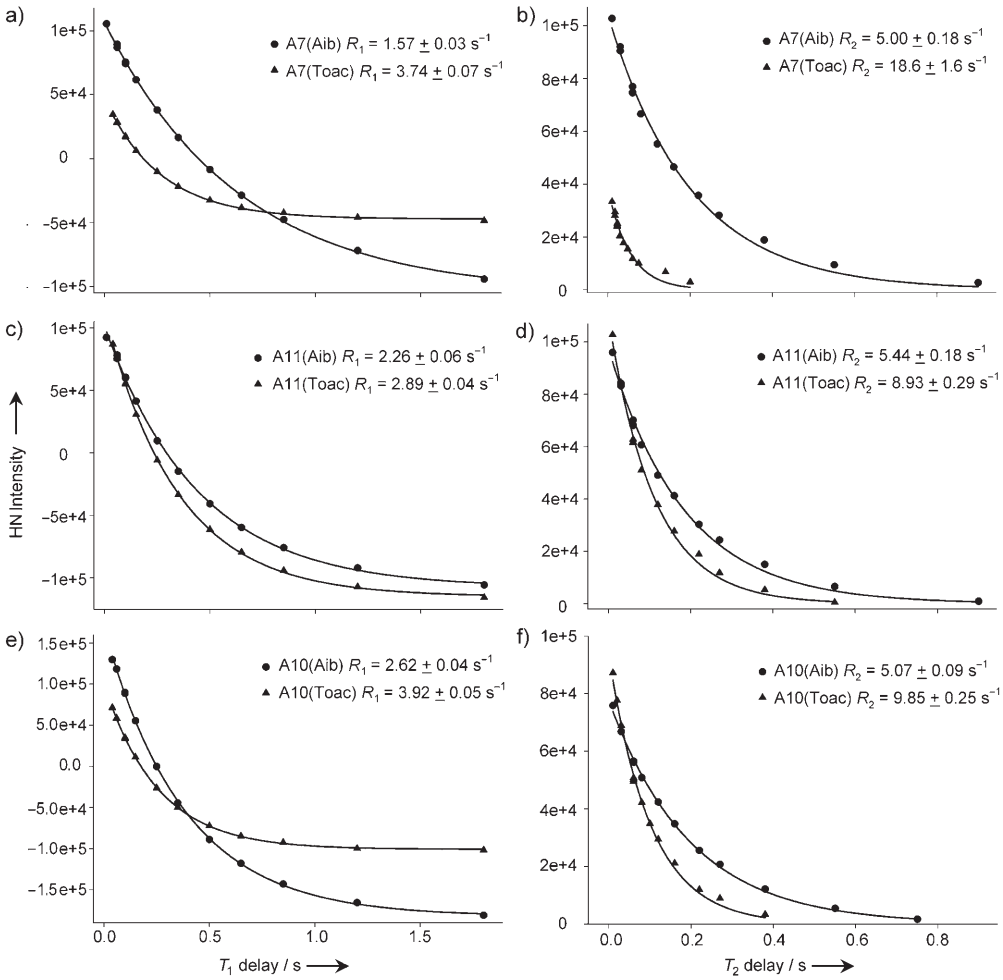

Figure 1. Relaxation data and fitting plots of H^N in ^{15}N-labeled Ala residues. a) R_1 of A*7. b) R_2 of A*7. c) R_1 of A*11. d) R_2 of A*11. e) R_1 of A*10. f) R_2 of A*10.

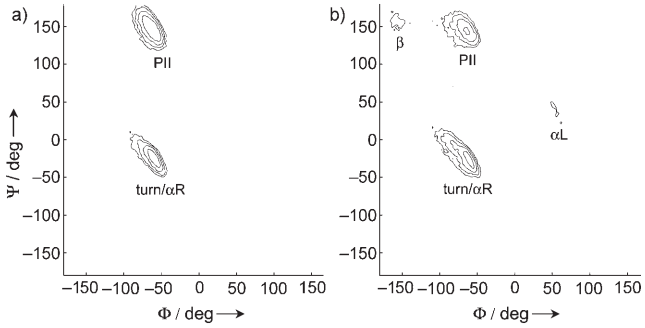

Figure 2. Ramachandran plots of Pro (a) and Ala (b) based on the coil library.[18] The mesostates are labeled as PII, β, turn/αR, and αL. The contours are drawn with statistical weights of 0.05%, 0.1%, 0.2%, and 0.4%.

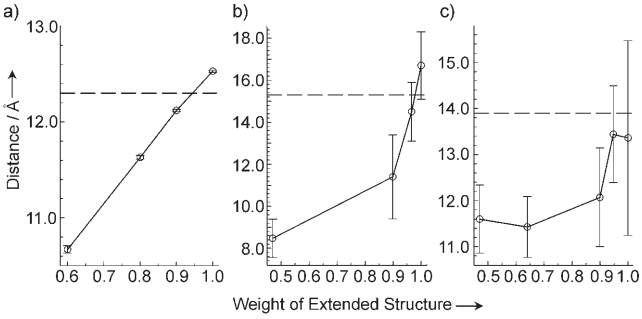

Figure 3. Monte Carlo simulations of distances determined by NMR spectroscopy. Statistical weights of extended structures were varied systematically in each panel. a) Distance T*3–A*7 with increased weight of extended conformations for Pro. b, c) Distance T*3–A*11 (b) and distance T*3–A*10 (c) with increased weight of extended conformations of Ala but fixed at 94% for Pro. The experimental distances by NMR spectroscopy are shown as horizontal dashed lines.

mesostates, turn/αR at 50%, PII at 38%, β at 8%, and αL at 4%, were identified for Ala. Both Pro and Ala possess significant populations of compact conformations such as turn/αR and αL in coil libraries.[5,18] Initial simulations based on the native distributions gave significantly shorter distance values than those detected by NMR spectroscopy, as shown by the first data point in each curve of Figure 3. We then systematically increased the weights of the extended conformers PII and β to approach the real conformational space in unfolded peptides.

The distance T*3–A*7 is determined solely by the structure of the three Pro residues between T*3 and A*7,

since H^N of A7 is rigidly anchored to the C=O group of residue P6. Thus, the Pro conformational space was calibrated from the distance of T*3–A*7 and fixed in subsequent simulations. We find that 94 ± 3% extended PII structure in Pro is needed to reproduce the distance of 12.3 Å determined by NMR spectroscopy (Figure 3a). We then grouped the PII and β structures as extended conformations and turn/αR and αL as compact structures for Ala. The statistical weight ratio of extended versus compact structures was varied during Monte Carlo sampling; the relative weight ratios of PII versus β and turn/αR versus αL were fixed at their native values. This procedure revealed that 96 ± 4% extended conformation in Ala is required to reproduce the distances determined by NMR spectroscopy (Figure 3b,c), with about 79% PII structure, 18% β, 3% turn/αR, and less than 1% αL. We note that this degree of PII conformation is based solely on the results of NMR spectroscopy, yet it agrees well with our earlier $J_{\alpha N}$ coupling-constant data[19] and with multiple sets of coupling constants measured by Graf et al.[3] Varying the PII weight in Pro from 90% to 98% did not alter this conclusion (data not shown). The simulated distributions of distances $P(r)$ are shown in Figure 4.

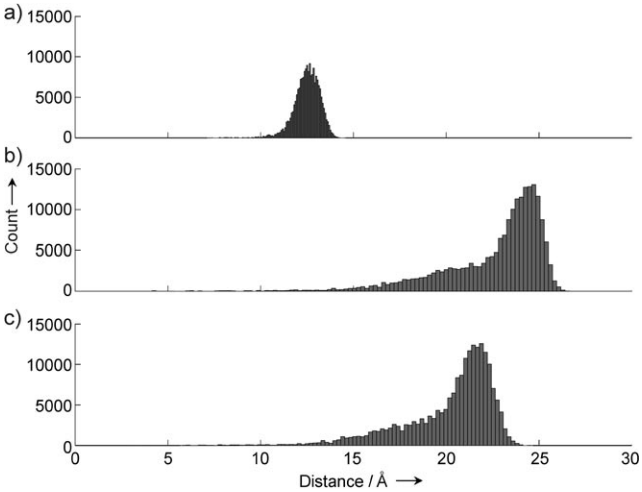

Figure 4. The linear distributions of simulated distances. a) T*3–A*7. b) T*3–A*11. c) T*3–A*10.

The tight coupling of the Toac spin label to the backbone enables estimates of intramolecular distances that are sufficiently accurate to define the conformational manifold of these peptides. We conclude that the conformational space of Ala in its unfolded state possesses over 90% extended structure, which is not induced by neighboring units. The maximum length of polyAla in this study is three residues (T*3–A*10), and we have shown previously that the PII structure is non-cooperative in tri-Ala peptides.[20] Pro next to Ala will not promote the PII structure of Ala,[21] and theoretical studies indicate that the Flory isolated pair theory holds for extended conformation.[22,23]

Small-angle X-ray scattering (SAXS) determination of the dimensions of XAO (a peptide similar to $O_2A_7O_2$) by Zagrovic et al.[24] pointed out that a blend of extended PII and β conformations should result in values of the radius of gyration much larger than the value they measured (7.4 ± 0.5 Å). We simulated an $(Ala)_{11}$ peptide ensemble containing 20000 structures using the Monte Carlo sampling method described above. The average radius of gyration we obtain is 8.1 ± 1.1 Å assuming 81% extended structures, which is consistent with the SAXS data. The reduction in population of extended conformation from 96% to 81% is consistent with incipient nucleation of α-helix structure in the longer Ala segment.[25] About 18% α-helix structure ("Helixcoil", http://www.nyu.edu/projects/kallenbach/) is predicted for a sequence of seven Ala residues.[26] Helicity builds up, or folding initiates, sharply with increasing peptide length in an appropriate context.[3]

This analysis shows that local order, and PII conformation in particular, is in fact compatible with random chain dimensions in unfolded proteins. The new data address the reconciliation problem from a different perspective.[27,28] A minor population of compact conformations at the residue level readily leads to global random-coil behavior for the whole chain. This idea has been expressed earlier;[6] the population distribution is now characterized experimentally in this work and in the important study of Graf et al.[3] Consistency with random-coil chain statistics is a weak constraint on local order.[29] It is important to note that other proposed models for short Ala peptides, with a roughly balanced distribution between extended and compact conformations,[30,31] do not agree with intramolecular distances determined by NMR spectroscopy.

Experimental Section

Peptide synthesis, NMR spin relaxation measurements, Monte Carlo sampling, and the geometry of each mesostate are described in the Supporting Information.

Received: July 27, 2007
Revised: September 8, 2007
Published online: October 17, 2007

Keywords: intramolecular distances · NMR spectroscopy · peptides · protein folding

[1] Z. Shi, K. Chen, Z. Liu, N. R. Kallenbach, *Chem. Rev.* **2006**, *106*, 1877.
[2] Z. Shi, R. W. Woody, N. R. Kallenbach, *Adv. Protein Chem.* **2002**, *62*, 163.
[3] J. Graf, P. H. Nguyen, G. Stock, H. Schwalbe, *J. Am. Chem. Soc.* **2007**, *129*, 1179.
[4] H. T. Tran, X. Wang, R. V. Pappu, *Biochemistry* **2005**, *44*, 11369.
[5] A. K. Jha, A. Colubri, K. F. Freed, T. R. Sosnick, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13099.
[6] H. Gong, P. J. Fleming, G. D. Rose, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16227.
[7] T. Head-Gordon, F. H. Stillinger, M. H. Wright, D. M. Gay, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11513.
[8] S. Marqusee, V. H. Robbins, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5286.
[9] E. J. Spek, C. A. Olson, Z. Shi, N. R. Kallenbach, *J. Am. Chem. Soc.* **1999**, *121*, 5571.
[10] L. Stryer, R. P. Haugland, *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 719.

[11] B. Schuler, E. A. Lipman, P. J. Steinbach, M. Kumke, W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2754.
[12] P. P. Borbat, J. H. Davis, S. E. Butcher, J. H. Freed, *J. Am. Chem. Soc.* **2004**, *126*, 7746.
[13] J. R. Gillespie, D. Shortle, *J. Mol. Biol.* **1997**, *268*, 158.
[14] M. M. Dedmon, K. Lindorff-Larsen, J. Christodoulou, M. Vendruscolo, C. M. Dobson, *J. Am. Chem. Soc.* **2005**, *127*, 476.
[15] J. Jacob, B. Baker, R. G. Bryant, D. S. Cafiso, *Biophys. J.* **1999**, *77*, 1086.
[16] C. Toniolo, E. Valente, F. Formaggio, M. Crisma, G. Pilloni, C. Corvaja, A. Toffoletti, G. V. Martinez, M. P. Hanson, G. L. Millhauser, C. George, J. L. Flippen-Anderson, *J. Pept. Sci.* **1995**, *1*, 45.
[17] B. Zagrovic, W. F. van Gunsteren, *Proteins Struct. Funct. Genet.* **2006**, *63*, 210.
[18] F. Avbelj, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5742.
[19] Z. Shi, C. A. Olson, G. D. Rose, R. L. Baldwin, N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9190.
[20] K. Chen, Z. Liu, N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15352.
[21] M. A. Kelly, B. W. Chellgren, A. L. Rucker, J. M. Troutman, M. G. Fried, A. F. Miller, T. P. Creamer, *Biochemistry* **2001**, *40*, 14376.
[22] R. V. Pappu, G. D. Rose, *Protein Sci.* **2002**, *11*, 2437.
[23] R. V. Pappu, R. Srinivasan, G. D. Rose, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12565.
[24] B. Zagrovic, J. Lipfert, E. J. Sorin, I. S. Millett, W. F. van Gunsteren, S. Doniach, V. S. Pande, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11698.
[25] B. H. Zimm, J. K. Bragg, *J. Chem. Phys.* **1959**, *31*, 526.
[26] J. X. Yang, K. Zhao, Y. X. Gong, A. Vologodskii, N. R. Kallenbach, *J. Am. Chem. Soc.* **1998**, *120*, 10646.
[27] I. S. Millett, S. Doniach, K. W. Plaxco, *Adv. Protein Chem.* **2002**, *62*, 241.
[28] E. R. McCarney, J. H. Werner, S. L. Bernstein, I. Ruczinski, D. E. Makarov, P. M. Goodwin, K. W. Plaxco, *J. Mol. Biol.* **2005**, *352*, 672.
[29] N. C. Fitzkee, G. D. Rose, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12497.
[30] J. Makowska, S. Rodziewicz-Motowidlo, K. Baginska, J. A. Vila, A. Liwo, L. Chmurzynski, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1744.
[31] R. Schweitzer-Stenner, T. J. Measey, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 6649.

