J|A|C|S
ARTICLES
Published on Web 12/01/2006

Conformational Preference of Polyglycine in Solution to Elongated Structure

Satoshi Ohnishi,*[,†,§] Hironari Kamikubo,[‡] Masayoshi Onitsuka,[‡] Mikio Kataoka,[‡] and David Shortle[†]

Contribution from the Department of Biological Chemistry, The Johns Hopkins University, School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, and Graduate School of Materials Science, Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara 630-0101, Japan

Received August 27, 2006; E-mail: ohnishi@gsc.riken.jp

Abstract: Do polypeptide chains ever behave like a random coil? In this report we demonstrate that glycine, the residue with the fewest backbone restrictions, exhibits a strong preference for an extended conformation in solution when polymerized in short segments of polyglycine. A model peptide system comprised of two unique tripeptide units, between which 1 to 18 glycine residues are inserted, is characterized by NMR and by small-angle X-ray scattering (SAXS). The residual dipolar coupling (RDC) values of the two tripeptide units are insensitive to changes in number of intervening glycines, suggesting that extension of the linker does not alter the average angular relationship between the tripeptides. Polyglycine segments longer than nine residues form insoluble aggregates. SAXS measurements using synchrotron radiation provide direct evidence that polyglycine peptides adopt elongated conformations. In particular, the construct with a linker with six glycines showed a scattering profile indicative of a monomeric state with a radius of gyration and the maximum dimension of 9.1 Å and ~34 Å, respectively. The ensemble averaged global structure of this 12-mer peptide can best be approximated by a cylinder with a radius of 4 Å and a length of ~33 Å, making it intermediate in extension between a β strand and an α helix.

1. Introduction

The random coil was originally introduced in polymer science as an idealized chain with conformations that vary randomly in three-dimensional space in a manner described by a Gaussian probability function.[1] Flory later extended this simple model to polypeptide chains by proposing the isolated-pair hypothesis: the backbone conformation of each amino acid, defined by its backbone dihedral angles, is independent of the conformation of its neighboring residues.[2] In practical terms, it has proven difficult to experimentally determine the level of randomness achieved by peptides and by unfolded proteins because of their dynamic and heterogeneous nature. The experimental support for this hypothesis is more or less limited to the finding of very good agreement between polymer length and ensemble averaged radius of gyration (R_g) as predicted by Flory's power law relationship

$$R_g = R_0 N^{\nu} \quad (1)$$

where N is the number of monomers in the polymer chain, R_0 is a constant related to persistence length, and ν is the scaling factor of interest that depends on solvent quality. The logarithm of R_g of denatured proteins determined by SAXS shows a linear correlation with the logarithm of chain length with $\nu = 0.598$, which is very similar to $\nu = 0.588$, predicted for the random coil.[1,3]

However, the radius of gyration is a scalar quantity representing an extreme level of averaging over a three-dimensional structure. Just as two folded proteins with the same radius of gyration do not necessarily have the same structure, multiple models of the denatured state ensemble with different levels of randomness can yield the same radius of gyration. For example Fitzkee and Rose generated ensembles of conformations from folded proteins of different sizes by varying only 7% of phi/psi angles.[4] These minimally unfolded ensembles displayed a remarkably good correlation between R_g and chain length, with $\nu = 0.602$. Clearly, more experimental data are needed to attain a better understanding of randomness and residual structure present in unfolded polypeptides.

Recent progress in NMR spectroscopy with measuring residual dipolar couplings (RDCs) has brought a new perspective to the study of unstructured polypeptides and denatured proteins.[5,6] RDCs reflect angular information of each dipole

† The Johns Hopkins University.
‡ Nara Institute of Science and Technology.
§ Current address: RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.

(1) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953.
(2) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Wiley: New York, 1969.
(3) Kohn, J. E.; Millett, I. S.; Jacob, J.; Zagrovic, B.; Dillon, T. M.; Cingel, N.; Dothager, R. S.; Seifert, S.; Thiyagarajan, P.; Sosnick, T. R.; Hasan, M. Z.; Pande, V. S.; Ruczinski, I.; Doniach, S.; Plaxco, K. W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12491.
(4) Fitzkee, N. C.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12497.
(5) Bax, A. *Protein Sci.* **2003**, *12*, 1.
(6) Prestegard, J. H.; al-Hashimi, H. M.; Tolman, J. R. *Q. Rev. Biophys.* **2000**, *33*, 371.

10.1021/ja066008b CCC: $33.50

vector of interest aligned with respect to the magnetic field. This angular information is independent of distance between the dipole vectors, and thus, RDCs can potentially provide long-range or global structural information.[7−11]

Here we describe a peptide system with tripeptide segments of Tyr-Glu-Ser and Ala-Thr-Asp connected by varying lengths of polyglycine linkers. Since glycine has the greatest freedom in backbone dihedral angles among the natural amino acids, polyglycine can be assumed to be the most "flexible" polypeptides. Unfortunately, NMR study of polyamino acids is usually quite difficult due to serious resonance signal overlaps. The present constructs also showed severe signal overlaps for glycine residues in the linker segment, but resonances from six amino acid residues in the terminal tripeptide segments were cleanly separated and unambiguously assigned. Backbone $^{1}H-^{15}N$ and $^{1}H_{\alpha}-^{13}C_{\alpha}$ RDCs from these residues were measured by using these isotopes at natural abundance and compared between the constructs with different linker lengths. Thus, the relative orientation of the tripeptide segments with respect to each other was probed as a function of the length of the intervening polyglycine linker.

In addition, to gain global structural information more directly, we performed a small-angle X-ray scattering (SAXS) study on the present peptide system using a synchrotron radiation source at the SPring-8 facility, Japan. We could obtain high-resolution data for these flexible short peptides with the residue length from 8 to 15 at various peptide concentrations so that nonideality from intermolecular interactions could be eliminated. Careful analyses of the scattering profiles of the peptides provided the Guinier profiles, the Kratky profiles, and the pair distance distribution profiles that were all consistent with a remarkable preference of polyglycine for an elongated conformation.

2. Materials and Methods

Peptide Design and Preparation. In the design of two tripeptide units connected with a polyglycine linker, we first screened amino acids based on "random coil chemical shift"[12] for resolvable ^{1}HN, ^{15}N, $^{13}C_{\alpha}$, and $^{1}H_{\alpha}$ resonances to reduce signal overlap. Six unique amino acids were chosen: glutamic acid and asparatic acid to confer a net charge for enhanced solubility; serine, threonine, and alanine for their unique chemical shifts; and tyrosine to permit measurement of peptide concentration by UV absorption. The N-terminal amide group was acetylated so that the amide proton resonance of the first residue could be detected. Thus, the Ac-YES-G_n-ATD construct was designed as the primary motif. All the peptides were synthesized by the Peptide Synthesis Facility of Johns Hopkins University School of Medicine, Department of Biological Chemistry. The final products of the synthesis and their purity were checked by reversed-phase FPLC and MALDI-TOF mass spectroscopy.

NMR Spectroscopy. All NMR data were collected on Varian Inova 500 and 600 spectrometers at 20 °C. Amide and C_{α} proton resonances of the Ac-YES-G_n-ATD peptides and those of the relevant variants in 20 mM sodium acetate buffer (pH 5.5) were assigned by $^{1}H-^{1}H$ NOESY with a mixing time of 200 ms and $^{1}H-^{1}H$ TOCSY spectra with a 1−2 mM peptide concentration. The $^{1}H-^{15}N$ and $^{1}H_{\alpha}-^{13}C_{\alpha}$ HSQC (heteronuclear single quantum coherence) spectra were measured using natural abundance with an ∼15 mM peptide concentration in the same buffer, unless mentioned otherwise. Weak molecular alignment was achieved by compression of a 18% (v/v) polyacrylamide gel (1:20 bis-acrylamide) cylinder as described previously,[11] and the $^{1}H-^{15}N$ and $^{1}H_{\alpha}-^{13}C_{\alpha}$ IPAP HSQC spectra were recorded with 9.5 Hz/point resolution along the t1 dimension and processed with zero filling with a final resolution of ∼0.2 Hz/point. A duration of ∼18−24 h for each experiment allowed a large number of accumulated scans to achieve a *S*/*N* ratio of ∼30. We estimate the uncertainty in chemical shift reading as ∼0.5 Hz by five independent measurements of the same sample on different days, which is consistent with the previous report.[7]

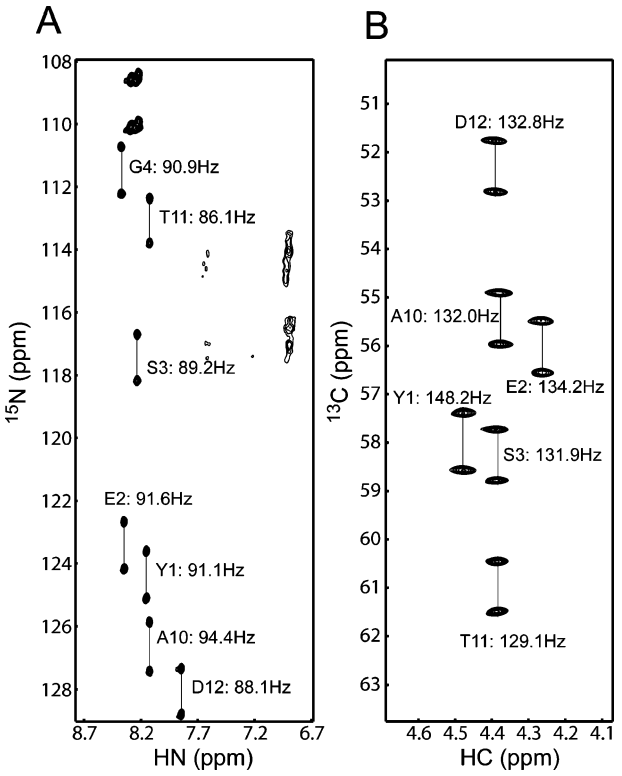

Figure 1. IPAP HSQC spectra of the Ac-YES-G_6-ATD at 20 °C, pH 5.5 constrained in the 18% polyacrylamide gel. (A) A $^{1}H-^{15}N$ IPAP HSQC spectrum recorded using a 600 MHz spectrometer. (B) A $^{1}H-^{13}C$ IPAP HSQC spectrum recorded using a 500 MHz spectrometer.

X-ray Fiber Diffraction. A suspension of the precipitate was placed in a 1 mm quartz capillary. Fiber diffraction data were collected using Cu Kα radiation from a Rigaku rotating anode operating at 50 kV and 100 mA. The diffraction pattern was recorded on an R-axis IV image plate (Rigaku) with a sample-to-detector distance of 80 mm, and the image was analyzed using the manufacturer's software.

SAXS Using Synchrotron Radiation Source. SAXS measurements were carried out using a 1.0 Å wavelength X-ray from an undulator source of the electron storage ring at the SPring-8 facility with a detector equipped with an R-Axis V imaging plate and a sample cell with a path length of 3 mm. Sample temperature was controlled at 20 °C, and the sample-to-detector distance was 1000 mm. AngiotensinII (SIGMA) was used for a control.

3. Results and Discussion

RDCs of the Ac-YES-G_6-ATD. The six glycine linker construct, Ac-YES-G_6-ATD, showed $^{1}H-^{15}N$ HSQC and $^{1}H-^{13}C$ HSQC spectra with clearly resolved resonances from the six unique amino acid residues and the first glycine (G4), so that $^{1}H-^{15}N$ and $^{1}H_{\alpha}-^{13}C_{\alpha}$ RDCs of these residues were unambiguously measured (Figure 1). The Ac-YES-G_6-ATD

(7) Ackerman, M. S.; Shortle, D. *Biochemistry* **2002**, *41*, 13791.
(8) Alexandrescu, A. T.; Kammerer, R. A. *Protein Sci.* **2003**, *12*, 2132.
(9) Fieber, W.; Kristjansdottir, S.; Poulsen, F. M. *J. Mol. Biol.* **2004**, *339*, 1191.
(10) Ohnishi, S.; Lee, A. L.; Edgell, M. H.; Shortle, D. *Biochemistry* **2004**, *43*, 4064.
(11) Ohnishi, S.; Shortle, D. *Proteins* **2003**, *50*, 546.
(12) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67.

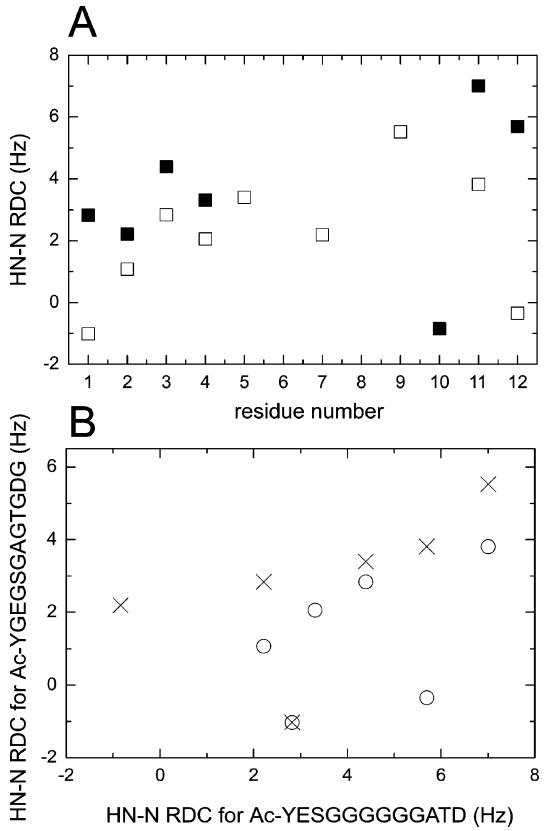

Figure 2. (A) $^{1}H-^{15}N$ RDC profile of the Ac-YES-G_6-ATD (■) and that of its variant (□), Ac-YGEGSGAGTGDG that has the identical amino acid composition and length but a different order in sequence. (B) Comparisons of $^{1}H-^{15}N$ RDC datasets between these peptides in terms of residue position (○) and residue type (×). Pearson's correlation coefficient *r* and *p* value for the residue position comparison were 0.46 and 0.36, respectively, and those for the residue type comparison were 0.58 and 0.23, respectively.

showed a range of RDCs from −0.8 to 7.0 Hz (black squares in Figure 2A), whereas theory predicts that random coils should show a bell-shaped RDC profile.[13] To examine the origin of the RDC values of AcYES-G_6-ATD, its shuffled analogue with the identical amino acid composition and length but a different sequence order (Ac-YGEGSGAGTGDG) was examined (white squares in Figure 2A). As shown in Figure 2B, RDC comparisons between these 12-mer peptides by residue position as well as by residue type demonstrate only poor correlations, indicating the amino acid sequence, not just amino acid composition, is the source of RDC profiles.

Changing the Length of the Polyglycine Linker. When the length of the polyglycine linker was changed to $n = 1$, 2, and 9, RDCs of the tripeptide units were minimally altered (Tables 1 and 2). In contrast, the RDC correlation of the tripeptide units was significantly perturbed when $n = 0$, and it became considerably lower when glycines in the linker were replaced with proline (Table 2). Thus, the type of linker affects the RDC profile of the tripeptides, suggesting that change in relative topology of the tripeptide segments can be monitored by change in RDCs. Therefore, the insensitivity of the RDC pattern to the length of the polyglycine linker can be interpreted as indicating the structure of each tri-peptide unit and relative angular relationship between the two ends of the polyglycine are insensitive to change in the length of the linker. Given the achiral nature of glycine and the lack of dependence on the number of intervening glycines, the simplest explanation for this situation is that the ensemble-averaged dimension of the linker would be a simple linear extended structure.

Table 1. Summary of RDC Data for the Ac-YES-G_n-ATD System Measured in Constrained Polyacrylamide Gels

$^{1}H-^{15}N$ RDC (Hz)	Y	E	S	A	T	D
$n = 0$	−0.4	2.5	4.9	3.7	5.9	4.5
$n = 1$	0.8	1.6	3.7	−0.7	5.9	3.9
$n = 2$	2.3	2.1	3.7	−0.8	6.4	5.8
$n = 6$	2.8	2.2	4.4	−0.8	7.0	5.7
$n = 9$	3.1	2.1	3.4	−0.5	6.4	5.1
Ac-YES-P-ATD	3.3	−0.4	4.7	3.1	5.7	3.7
$^{1}H_\alpha-^{13}C_\alpha$ RDC (Hz)	Y	E	S	A	T	D
$n = 0$	3.3	−12.4	−7.2	−11.3	−15.2	−8.4
$n = 1$	4.4	−9.4	−9.1	−8.3	−11.7	−10.4
$n = 2$	5.5	−10.3	−9.9	−9.1	−15.2	−10.4
$n = 6$	3.6	−11.0	−10.5	−8.2	−13.1	−12.2
$n = 9$	4.9	−9.9	−9.6	−9.1	−15.6	−10.6
Ac-YES-P-ATD	1.8	−13.2	−7.2	−7.7	−1.7	−8.5

Table 2. Summary of Correlation Coefficients in the RDC Comparisons with Respect to the Ac-YES-G_6-ATD Datasets

	$^{1}H-^{15}N$ RDC		$^{1}H_\alpha-^{13}C_\alpha$ RDC	
	r	*p*	*r*	*p*
Ac-YESATD	0.49	0.32	0.90	0.01
Ac-YES-P-ATD	0.56	0.25	0.63	0.18
Ac-YES-G-ATD	0.96	2.3×10^{-3}	0.99	$<0.1 \times 10^{-4}$
Ac-YES-G_2-ATD	0.99	$<0.1 \times 10^{-4}$	0.98	5.8×10^{-4}
Ac-YES-G_9-ATD	0.99	1.6×10^{-4}	0.98	8.9×10^{-4}

[a] *r* and *p* represent Pearson's correlation coefficient and *p*-value, respectively.

When the linker was extended beyond six glycines, the Ac-YES-G_n-ATD constructs displayed reduced solubility. Ac-YES-G_9-ATD showed poor solubility at low pH, while it could be dissolved when pH was increased above 5. However, as described below, SAXS analysis revealed that this 15-mer peptide exists in an oligomeric state. The Ac-YES-G_{18}-ATD was extremely insoluble, both at acidic and basic pH's, and the addition of chaotropic agents failed to dissolve the precipitates. The insoluble precipitates of the Ac-YES-G_{18}-ATD as well as those of Ac-YES-G_9-ATD at low pH displayed a characteristic 4.2 Å X-ray diffraction periodicity, a hallmark for the polyglycine II assembly, in which the polyglycine chain is modeled as an extended conformation with phi and psi angles of −80° and 150°, respectively.[14] It should be noted that polypeptide conformation in a solid does not necessarily correspond to that in solution. Nevertheless, the marked tendency of the present constructs with longer polyglycine linkers for forming the insoluble polyglycine II conformation is consistent with an elongated ensemble-averaged structure of polyglycine in solution.

Dimensional Characterization by SAXS. To obtain more direct evidence for the remarkable preference of polyglycine for an extended conformation, SAXS analysis using the synchrotron radiation source was carried out. We collected SAXS profiles for Ac-YES-G_2-ATD, Ac-YES-G_6-ATD, Ac-YES-G_9-ATD, and Angiotensin II (8 residues; DRVYIHPF) as a control. As judged by the absolute intensity extrapolated at *Q*

(13) Louhivuori, M.; Paakkonen, K.; Fredriksson, K.; Permi, P.; Lounila, J.; Annila, A. *J. Am. Chem. Soc.* **2003**, *125*, 15647.

(14) Crick, F. H.; Rich, A. *Nature* **1955**, *176*, 780.

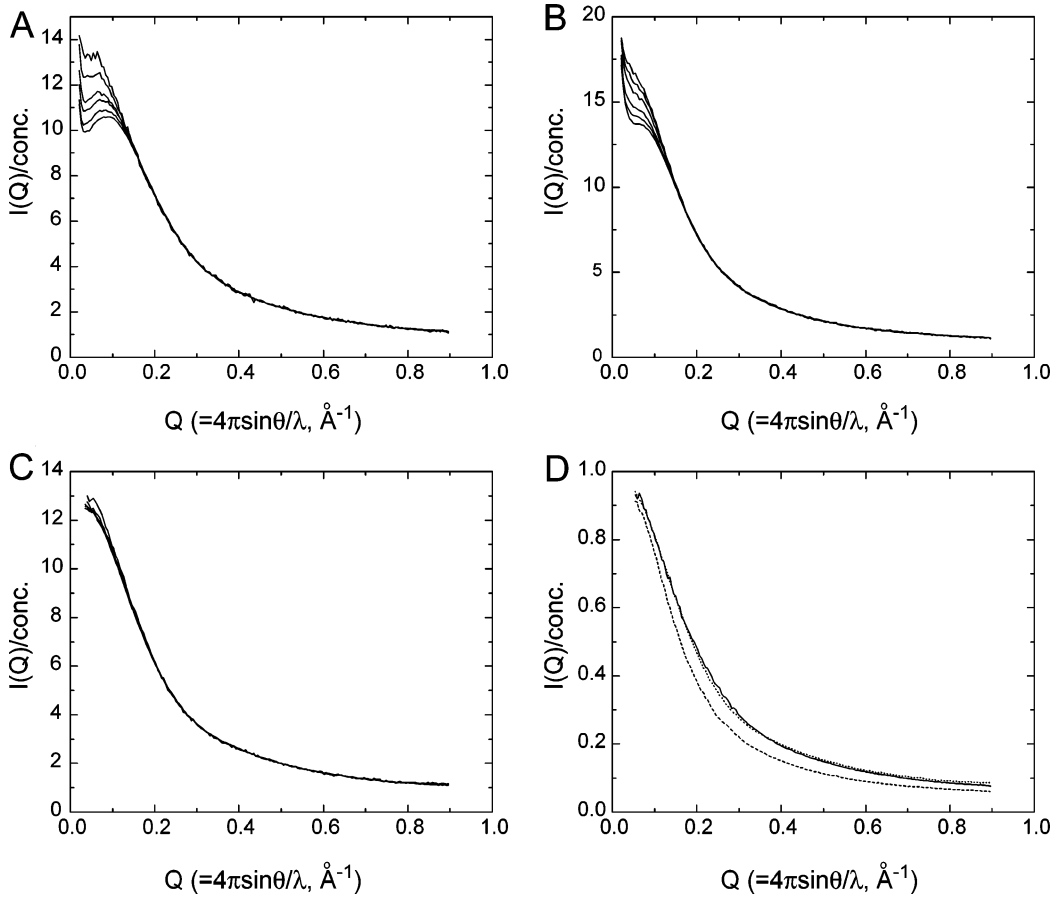

Figure 3. X-ray scattering profiles of the Ac-YES-G_2-ATD (A), Ac-YES-G_6-ATD (B), and Angiotensin II (C) at various concentrations between 2 and 10 mg/mL. (D) Normalized scattering profiles at the zero peptide concentration via singular value deconvolution analyses for Ac-YES-G_2-ATD (solid), Ac-YES-G_6-ATD (dash), and Angiotensin II (dot).

$= 0$ Å^{-1}, i.e., $I(0)$, Ac-YES-G_9-ATD is in an oligomeric state with association number of ~3 to 4, while the others are monodisperse and monomeric. Consequently, we focus on the SAXS results on the three monomeric peptides.

As shown in Figure 3, the scattering profile of Ac-YES-G_2-ATD and that of Ac-YES-G_6-ATD showed a significant concentration dependence in the lower Q region ($Q < 0.2$ Å^{-1}), indicative of intermolecular interactions that likely arise from a -3 net charge (glutamate, asparatate, and the uncapped carboxyl terminus). We collected scattering data of these peptides at six different concentrations and obtained their normalized scattering profiles at zero peptide concentration via singular value deconvolution analyses (Figure 3D). All subsequent analyses use these normalized scattering profiles.

The Guinier fitting analyses on the scattering profiles (Supporting Information Figure S1) yielded the R_g values for Angiotensin II, Ac-YES-G_2-ATD, and Ac-YES-G_6-ATD as 7.94 Å, 7.90 Å, and 9.10 Å, respectively. These values are significantly larger than those expected from the well-known power law relationship (eq 1, ~4.7 Å for 8-mer and ~6.0 Å for 12-mer), supporting that the ensemble-averaged conformation of these peptides in solution must be elongated.

As shown in Figure 4A, the Kratky plots of Ac-YES-G_2-ATD, Ac-YES-G_6-ATD, and Angiotensin II all show a linear increase with Q in the $Q > 0.6$ Å^{-1} region, suggesting that these polypeptide chains are constituted with rigid and straight segments. Figure 4B and 4C show comparisons of Kratky plots of Ac-YES-G_2-ATD and Ac-YES-G_6-ATD between experimentally obtained curves and simulated curves with model conformers (i.e., (β-strand) phi $-140°$, psi 135°; (α-helix) phi $-60°$, psi $-40°$; (polyglycine II) phi $-80°$, psi 150°). Obviously, the above models cannot reproduce the experimental profiles of the peptides.

Figure 5 shows the pair distance distribution function $P(r)$ calculated for the three peptides. The profile of Ac-YES-G_2-ATD shows a larger error in the longer distance region, and that of Angiotensin II shows a negative value around $R = 30$ Å. For smaller molecules like those for the present peptides, it is quite difficult to totally exclude artifacts from negative electron density contrast and/or intermolecular interactions. Consequently, the maximum dimensions (D_{max}) of the present peptides contain modest uncertainty (Table 3). Nevertheless, the $P(r)$ profiles in Figure 5 clearly demonstrate that solution conformations of the Ac-YES-G_n-ATD peptides are distinct from β-strand, α-helix, and polyglycine II, as suggested by the analyses of the Kratky plots.

As summarized in Table 3, the dimensions of the Ac-YES-G_n-ATD peptides are significantly smaller than those calculated from the polyglycine II model, which is consistent with a previous report on a polyalanine peptide system.[15] For example, R_g of an 11-mer polyalanine peptide (Ac-(diaminobutyric acid)$_2$-A_7-(ornithine)$_2$-NH_2) was determined as 7.4 Å, while that of this peptide in the polyproline II conformation (identical to the polyglycine II conformation in terms of phi and psi angles) was

(15) Zagrovic, B.; Lipfert, J.; Sorin, E. J.; Millett, I. S.; van Gunsteren, W. F.; Doniach, S.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11698.

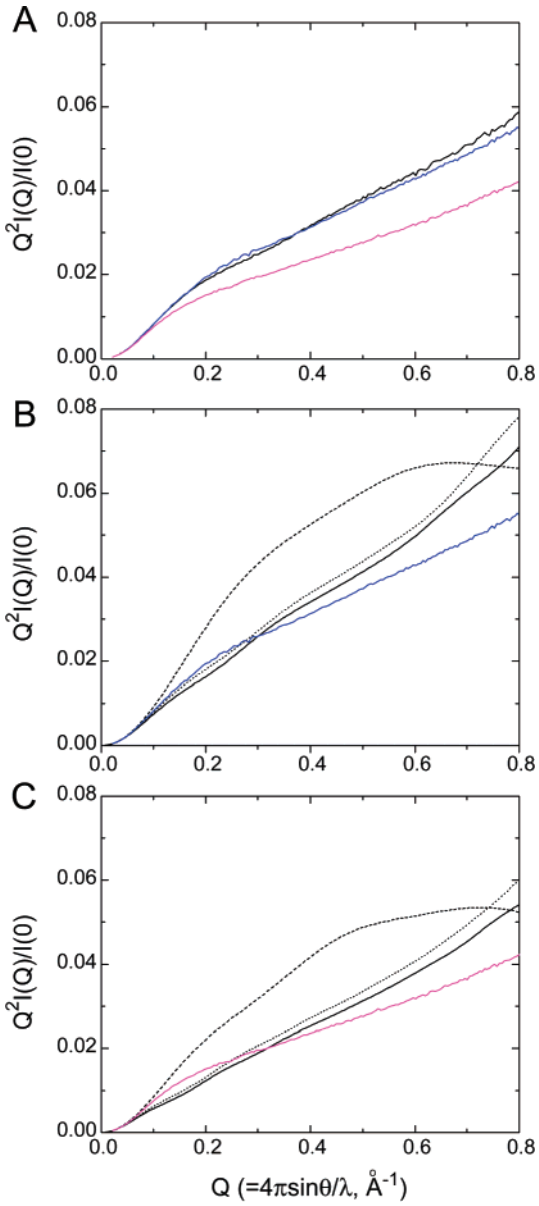

Figure 4. (A) Kratky plots of the Ac-YES-G_2-ATD (blue), the Ac-YES-G_6-ATD (magenta), and Angiotensin II (black). (B, C) Simulated Kratky plots of the Ac-YES-G_2-ATD and Ac-YES-G_6-ATD in α-helix (dash), β-strand (solid), and polyglycine II (dot).

calculated as 13.1 Å. Of particular interest here is that the experimentally determined R_g of the alanine-based peptide is significantly smaller than that of the glycine-based peptide shown in Table 3, suggesting that polyglycine prefers a more extended conformation than polyalanine.

The overall scattering profile of the Ac-YES-G_2-ATD is quite similar to that of Angiotensin II (Figure 2), which indicates that the ensemble-average structures of these 8-mer peptides resemble each other. Atomic coordinates of Angiotensin II are found in the PDB databank; the entry 1N9V represents the NMR structure determined in an organic solvent,[16] and 3CK0 represents the crystal structure complexed with Fab131. Both Angiotensin II structures are well defined and similar with an L-shaped backbone trace (Supporting Information Figure S2). The structure in the Fab131 complex is more extended than that in organic solvent, and the ensemble-averaged structure of Angiotensin II in aqueous solution in free form would be furthermore expanded upon loosing contacts with the protein as well as crystal packing, which is consistent with our SAXS data summarized in Table 3. Therefore, the high degree of similarity in the SAXS profile between Angiotensin II and the Ac-YES-G_2-ATD suggests that the structure ensemble of Ac-YES-G_2-ATD is composed of elongated conformers.

Molecular Shape of Ac-YES-G_n-ATD in Solution. Results from both NMR and SAXS studies support a marked preference of the Ac-YES-G_n-ATD peptides for an extended conformation. To characterize the elongated conformation, we first analyzed the SAXS data using a cylinder model. The scattering from a cylinder is represented by the equation

$$I(Q)/I(0) = \int_0^{\pi/2} f^2(Q,\alpha) \sin\alpha \, d\alpha$$

$$f(Q,\alpha) = 2j_0(QH\cos\alpha)\frac{J_1(Qr\sin\alpha)}{(Qr\sin\alpha)}$$

$$j_0 = \sin(x)/x \qquad (2)$$

where $J_1(x)$ is the first-order Bessel function, and $2H$ and r are the length along long axis and radius of the cylinder, respectively. Geometric parameters thus obtained are summarized in Table 4. The ratio of the long-axis/short-axis of the cylinder is approximately 4−5, providing quantitative support for an extended conformation. Similar analyses assuming an ellipsoidal model could not reproduce the experimental curve well (data not shown).

We also analyzed the scattering profiles by assuming the semiflexible chain (wormlike chain) model as follows. In the small angle region, $Q < 0.4$ Å^{-1}, we used the Kratky−Porod's equation:

$$I(Q)/I(0) = \frac{2}{x^2}(x - 1 + \exp(-x)) + \frac{b}{L}\left[\frac{4}{15} + \frac{7}{15x} - \left(\frac{11}{15} + \frac{7}{15x}\right)\exp(-x)\right]$$

$$x = \frac{Q^2Lb}{6} \qquad (3)$$

where L and b denote the Contour length (the maximum end-to-end distance) and the statistical chain element length, respectively, since this relation holds a good approximation in the smaller Qb region ($Qb < 3-4$). On the other hand, for the larger angle region, $Q > 0.4$ Å^{-1}, we used the des Cloiseaux's equation:

$$I(Q)/I(0) = \frac{b}{L}\left[\frac{11.933}{(Qb)^2} - \frac{0.015\,77}{Qb} + 0.2988 - 0.00925(Qb)\right] \qquad (4)$$

This model shows a good fit in the larger Qb region ($Qb > 3-4$).[17] Table 5 shows the b and L values for Ac-YES-G_2-ATD, Ac-YES-G_6-ATD, and Angiotensin II. The analyses show a good correspondence between the results obtained from the two models; however it should be noted that approximation with

(16) Matsoukas, J. M.; Hondrelis, J.; Keramida, M.; Mavromoustakos, T.; Makriyannis, A.; Yamdagni, R.; Wu, Q.; Moore, G. J. *J. Biol. Chem.* **1994**, *269*, 5303.

(17) Garcia, P.; Serrano, L.; Durand, D.; Rico, M.; Bruix, M. *Protein Sci.* **2001**, *10*, 1100.

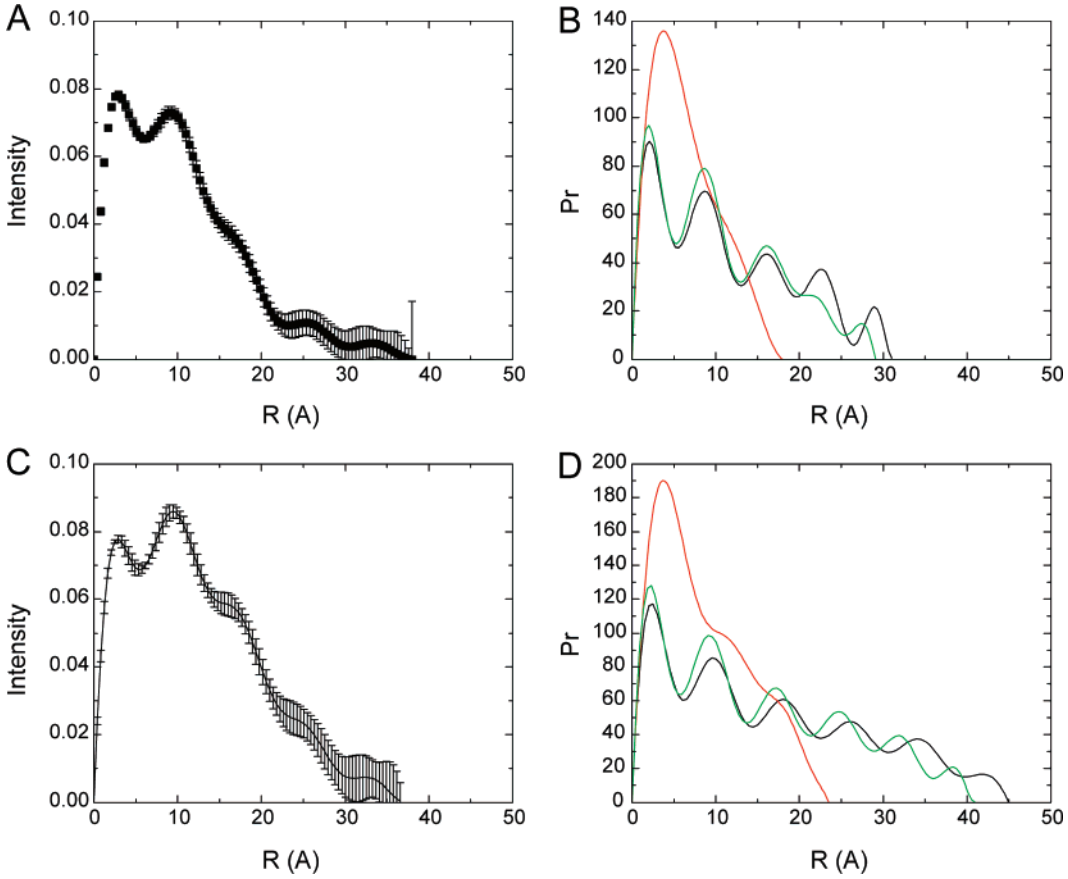

Figure 5. (A) $P(r)$ profile of the Ac-YES-G_2-ATD obtained from the scattering profile and (B) those from modeled conformations in the α-helix (red), β-strand (black), and polyglycine II (green). (C) $P(r)$ profile of the Ac-YES-G_6-ATD obtained from the scattering profile and (D) those from modeled conformations in the α-helix (red), β-strand (black), and polyglycine II (green).

Table 3. Summary of Dimensions of the Ac-YES-G_2-ATD and the Ac-YES-G_6-ATD and Angiotensin II

		R_g (Å)	D_{max} (Å)
AcYES-G_2-ATD	experimental	7.90	28−38
	β-strand model[a]	10.01	∼31
	α-helix model[a]	5.92	∼18
	PPII helix model[a]	9.10	∼29
AcYES-G_6-ATD	experimental	9.10	31−37
	β-strand model[a]	14.58	∼45
	α-helix model[a]	7.72	∼23.5
	PPII helix model[a]	13.19	∼41
Angiotensin II	experimental	7.94	28−38
	1N9V[b]	7.02	∼20
	3CK0[c]	7.82	∼25

[a] Models were defined with phi and psi angles of −140° and 135° for the β-strand model, −60° and −40° for the α-helix model, and −80° and 150° for the polyglycine II model, respectively. [b] Calculated values based on the PDB coordinate of Angiotensin II (1N9V) that is determined in an organic solvent by solution NMR.[16] [c] Calculated values based on the PDB coordinate of Angiotensin II (3CK0) that is determined in a crystal structure of a complex with Fab131.

Table 4. Geometric Parameters Obtained Assuming the Rigid Cylinder Model[a]

	χ^2	r (Å)	$2H$ (Å)	scale factor
Ac-YES-G_2-ATD	0.002 457 8	2.99	27.8	1.015
Ac-YES-G_6-ATD	0.000 882 5	4.00	32.8	1.008
Angiotensin II	0.000 260 8	3.23	28.2	1.014

[a] r: radius of the cylinder. $2H$: full length of the cylinder.

Table 5. Geometric Parameters Obtained Assuming the Wormlike Model Using the Kratky−Porod's Equation and the des Cloiseaux Equation

		b (Å)	L (Å)
AcYES-G_2-ATD	Kratky−Porod	12.5	47.0
	des Cloiseaux	11.8	46.6
AcYES-G_6-ATD	Kratky−Porod	12.9	62.0
	des Cloiseaux	12.7	61.5
Angiotensin II	Kratky−Porod	13.8	46.9
	des Cloiseaux	13.0	44.5

these models is generally appropriate for long chain molecules but not for short peptides.[18] The statistic chain segment length b is constant around 12−13 Å for the present polyglycine peptides, which is obviously too short for the molecule with a D_{max} of 30−38 Å. Thus, one straightforward interpretation from this analysis is that the whole chain should be regarded as a perturbed chain that has interactions between the rigid segment units.

Physical Basis of an Elongated Conformation. Despite no obvious long-range interactions, the present results revealed a pronounced conformational preference of the Ac-YES-G_n-ATD peptides to an elongated structure. One simple explanation is that this conformational preference might be induced by like-charge repulsions between the glutamate in the N-terminal unit and the C-terminal asparatate and/or the uncapped carboxyl tail. Although precise evaluation of surface electrostatic potential

(18) Glatter O., Kratky, O. *Small-Angle X-Ray Scattering*; Academic Press: London, 1982.

for flexible molecules is extremely difficult, a simple consideration can provide a perspective. The strength of electrostatic interaction between two charges is proportional to the inverse of the distance between them, and the Bjerrum length that refers to the approximate distance within which electrostatic interactions dominate thermal motions for two charges is typically assumed to 7 Å in pure water. In a buffered solvent system like the present case, the Bjerrum length is reduced. Therefore, single charge repulsion between Glu2 and Asp12 in the Ac-YES-G_6-ATD appears unable to restrict this peptide conformation, which has maximum dimensions around 31−37 Å (Table 3). Indeed, the RDC profile of Ac-YES-G_6-ATD was insensitive between pH 5.5 and pH 3.5 (RDC correlation was $r = 0.98$, $p < 0.0001$). Thus, it is quite unlikely that charge repulsion induces the conformational preference of the Ac-YES-G_n-ATD peptides.

RDC comparisons between Ac-YES-G_6-ATD and its shuffled analogue indicate the amino acid sequence is responsible for the characteristic RDC profiles of these peptides and, thus, for the conformational preference of them in solution. In other words, local steric and electrostatic interactions with neighboring residues are not negligible in defining the solution conformations of unfolded polypeptides, which argues against the Flory's isolated-pair hypothesis.[2]

In addition, there might be additional factors that contribute to conformational preferences in unfolded polypeptides. For instance, significant effects of hydration on the conformational behavior of denatured polypeptides have been reported,[19−23] suggesting that hydration effects via direct hydrogen bonds between water molecule(s) and peptide backbone might restrict freedom in dihedral angles.

4. Conclusions

Despite the fact that glycine possesses the greatest freedom in phi/psi values, polyglycine exhibits a remarkable preference for an extended conformation. The ensemble-averaged conformation of Ac-YES-G_6-ATD could be approximated well by a cylinder model with a radius of 4 Å and a length of ~33 Å. As shown in Figure 6, this conformation is significantly more elongated than the ideal α-helix (length = ~23.5 Å) while more compact than the ideal β-strand (length = ~45 Å) and polyglycine II (length = ~41 Å). When the length of a glycine linker exceeds 9, polyglycine favors intermolecular association through backbone hydrogen bonding in a stretched conformation (polyglycine II). Thus the average solution conformation of polyglycine may be similar to that of polyglycine II in the solid state, with only small changes being necessary for intermolecular associations and loss of solubility.

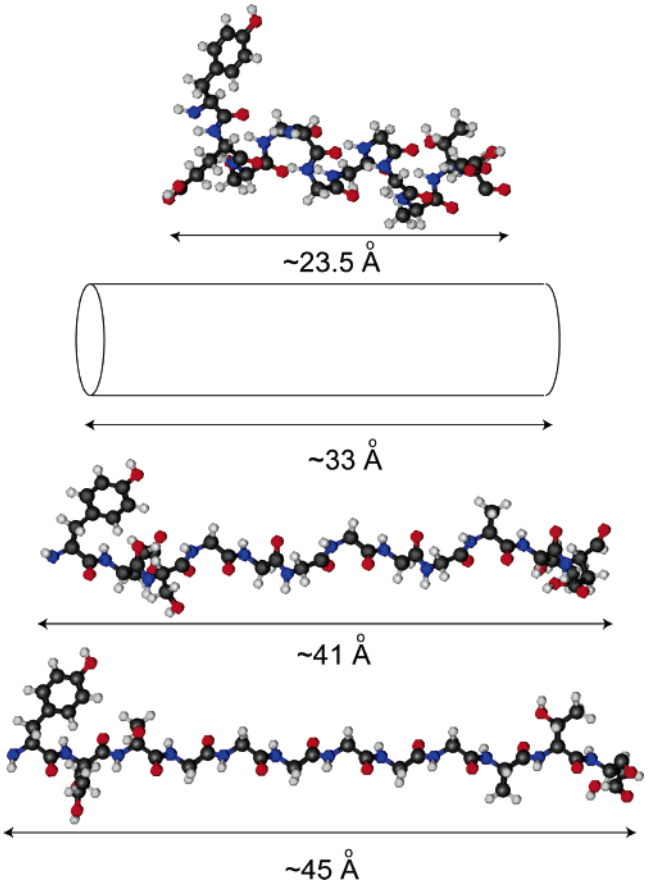

Figure 6. Models of Ac-YES-G_6-ATD to scale. From top to bottom: the 12-mer peptide in an α-helix, a cylinder with the dimension consistent with the SAXS data, the peptide in a polyglycine II, and that in a β-strand.

Acknowledgment. The authors thank Joel Tolman, George Rose, and Shohei Koide for helpful discussions; Ananya Majumdar for assistance on NMR measurements; and Michael Cosgrove for assistance in X-ray fiber diffraction measurements. This work was partly supported by NIH Grants GM34171 (D.S.). The SAXS work was performed under the approval of the JASRI Program Advisory Committee (Proposal No. 2005B0012).

Supporting Information Available: Figure S1: the Guinier plots of Angiotensin II, the Ac-YES-G_2-ATD, and the Ac-YES-G_6-ATD obtained from the normalized scattering profiles. Figure S2: solution structure of Angiotensin II determined in an organic solvent[16] (1N9V) and crystal structure of Angiotensin II in a complex with Fab131 (3CK0) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA066008B

(19) Wang, Y.; Jardetzky, O. *J. Am. Chem. Soc.* **2002**, *124*, 14075.
(20) Garcia, A. E.; Hummer, G.; Soumpasis, D. M. *Proteins* **1997**, *27*, 471.
(21) Li, Y.; Picart, F.; Raleigh, D. P. *J. Mol. Biol.* **2005**, *349*, 839.
(22) Hackel, M.; Konno, T.; Hinz, H. *Biochim. Biophys. Acta* **2000**, *1479*, 155.
(23) Avbelj, F.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5742.