NMR-Based Quantification of β-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States

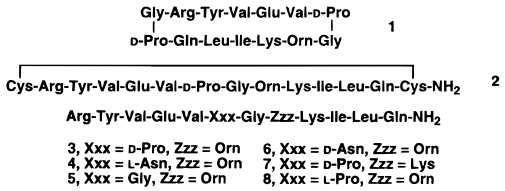
Faisal A. Syud, Juan F. Espinosa, and Samuel H. Gellman*

Department of Chemistry, University of Wisconsin Madison, Wisconsin 53706

Received August 2, 1999
Revised Manuscript Received October 19, 1999

Stabilization of small increments of β-sheet in aqueous solution is critical for fundamental research on protein folding preferences[1−8] and for development of biologically active small molecules.[9−11] Two different strategies for promoting double-stranded antiparallel β-sheet are found among natural and synthetic peptides: side-chain linkage, usually involving cysteine residues,[10] and backbone linkage.[11] Here, we report a direct comparison between side-chain disulfide cyclization and backbone cyclization for generation of reference compounds that can be used to determine β-sheet populations of flexible peptides. NMR data suggest that both cyclization modes in our peptide system stabilize a very highly populated native-like antiparallel β-sheet conformation, although backbone cyclization appears to confer somewhat greater conformational stability. These reference compounds allow quantitative evaluation of sequence-stability relationships among linear peptides that adopt double-stranded β-sheet ("β-hairpin") conformations in aqueous solution.

Cyclic peptides **1** and **2** (Figure 1) are potential models for the fully folded state of linear peptide **3**. Both **1** and **2** contain the



12-residue sequence of **3**, with two additional residues to allow ring closure. Our previous studies of **3** and related linear peptides have shown that the D-Pro-Gly segment is a very strong inducer of β-sheet formation in adjacent strands,[2,12] and we therefore used a second D-Pro-Gly segment to achieve backbone cyclization in **1**.[13] For **2**, the disulfide-linked cysteines were placed at a "non-hydrogen-bonded" pairing because this juxtaposition is common among proteins.[14] We have previously concluded from NMR data that a single β-sheet conformation is significantly populated for linear peptide **3** in aqueous solution.[2d] Like other short peptides, however, **3** equilibrates rapidly on the NMR time scale between the unfolded and β-sheet conformations. A high population in **1** or **2** of a β-sheet conformation analogous to that observed for **3** would allow determination of the β-sheet population of **3** from conformationally averaged NMR data.[15]

Two-dimensional NMR data for **1** and **2** in aqueous solution revealed numerous NOEs between residues that are not adjacent in sequence.[16] The NMR data for **1** and **2** were used for NOE-restrained structure determinations with the program DYANA (Figure 1).[17] For each peptide, all NOEs were consistent with a single family of conformations in which the two five-residue strand segments, Arg-2 to Val-6 and Orn-9 to Gln-13, display the expected antiparallel β-sheet structure. (We number the residues of **1** beginning with Gly adjacent to Arg to facilitate comparison with **2**.) The best 20 out of 500 annealed structures were compared for each peptide; RMSD among the 20 structures (backbone only) was 0.75 ± 0.23 Å for **1** and 1.40 ± 0.39 Å for **2**. The 20-structure average for **1** showed a 1.20 Å RMSD relative to the 20-structure average for **2**, indicating that the two cyclization modes promote similar conformations.

Amide H/D exchange data (100 mM sodium deuterioacetate buffer, pD 3.8 (uncorrected), 4 °C) provided further evidence that cyclic peptides adopt a regular β-sheet conformation; the exchange data also suggested that backbone cyclization induces a somewhat more stable folded state than does disulfide cyclization. Retardation of amide H/D exchange is generally interpreted to indicate a steric and/or electronic impediment to chemical catalysis of exchange, e.g., from intramolecular hydrogen bonding.[18] For **1**, there was a regular alternation of pseudo-first-order rate constants along the strands: each residue expected to orient its N−H away from the opposite strand showed a rate constant $>10^{-3}$ min^{-1}, and each residue expected to be engaged in interstrand hydrogen

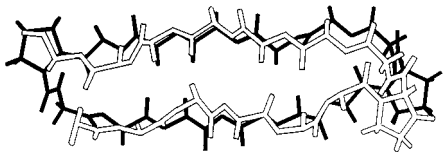


Figure 1. Comparison of NOE-restrained dynamics simulations (DYANA[17]) for cyclic peptides **1** and **2**. The 20 best structures were averaged for each peptide; shown here is an overlay of the two averaged structures (backbone only; **1** is black and **2** is white). The proline side chains are shown for orientation; the D-Pro-Gly turn common to **1** and **2** is at the right.

* Author for correspondence. E-mail: gellman@chem.wisc.edu.

(1) Recent review: Gellman, S. H. *Curr. Opin. Chem. Biol.* **1998**, *7*, 717.
(2) (a) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4105. (b) Haque, T. S.; Little, J. C.; Gellman, S. H., *J. Am. Chem. Soc.* **1996**, *118*, 6975. (c) Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 2303. (d) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236. (e) Schenck, H. L.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4869.
(3) (a) Blanco, F. J.; Jiménez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887. (b) de Alba, E.; Rico, M.; Jiménez, M. A. *Protein Sci.* **1997**, *6*, 2548. (c) de Alba, E.; Jimenez, M. A.; Rico, M. *J. Am. Chem. Soc.* **1997**, *119*, 175.
(4) (a) Ramírez-Alvarado, M.; Blanco, F. J.; Serrano, L. *Nat. Struct. Biol.* **1996**, *3*, 604. (b) Ramírez-Alvarado, M.; Blanco, F. J.; Niemann, H.; Serrano, L. *J. Mol. Biol.* **1997**, *273*, 898. (c) Kortemme, T.; Ramírez-Alvarado, J.; Serrano, L. *Science* **1998**, *281*, 253.
(5) (a) Sharman, G. J.; Searle, M. S. *J. Chem. Soc., Chem. Commun.* **1997**, 1955. (b) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 5291. (c) Maynard, A. J.; Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 1996.
(6) Muñoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196.
(7) Ilynia, E.; Roongta, V.; Mayo, K. H. *Biochemistry* **1997**, *36*, 5245.
(8) Koepf, E. K.; Petrassi, M. H.; Sudol, M.; Kelly, J. W. *Protein Sci.* **1999**, *8*, 841.
(9) (a) Veber, D. F.; Freidinger, R. M.; Perlow, D. S.; Palaveda, W. J.; Holly, F. W.; Strachan, R. G.; Nutt, R. F.; Arison, B. J.; Homnick, C.; Randal W. C.; Glitzer, M. S.; Saperstein, R.; Hirschmann, R. *Nature* **1981**, *292*, 55. (b) Melacini, G.; Zhu, Q.; Goodman, M. *Biochemistry* **1997**, *36*, 1233.
(10) (a) Fahrner, R. L.; Dieckmann, T.; Harwig, S. S. L.; Lehrer, R. I.; Eisenberg, D. I.; Feigon, J. *Chem. Biol.* **1996**, *3*, 543. (b) McDonnell, J. M.; Fushman, D.; Cahill, S. M.; Sutton, B. J.; Cowburn, D. *J. Am. Chem. Soc.* **1997**, *119*, 5321.
(11) (a) Hull, S. E.; Karlsson, R.; Main, P.; Woolfson, M. M. *Nature* **1978**, *275*, 206. (b) Späth, J.; Stuart, F.; Jiang, L.; Robinson, J. A. *Helv. Chim. Acta* **1998**, *81*, 1726.
(12) For related work, see: (a) Karle, I. L.; Awasthi, S. K.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8189. (b) Raghothama, S. R.; Awasthi, S. K.; Balaram, P. *J. Chem. Soc., Perkin Trans. 2* **1998**, 137.
(13) Analytical ultracentrifugation showed that none of these peptides self-associates under the conditions used for NMR analysis.
(14) (a) Wouters, M. A.; Curmi, P. M. G. *Proteins: Struct., Funct. Genet.* **1995**, *22*, 119. (b) Hutchinson, E. G.; Sessions, R. B.; Thornton, J. M.; Woolfson, D. N. *Protein Sci.* **1998**, *7*, 2287.
(15) For other approaches to quantification of β-hairpin populations, see refs 3b, 4a,b and 5b,c.
(16) Details may be found in the Supporting Information.
(17) Guntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283.

10.1021/ja992733t CCC: $18.00
Published on Web 11/30/1999

bonding showed a rate constant $<10^{-3}$ min^{-1} (for two residues the rate constant was $<10^{-5}$ min^{-1}). For **2**, a similar distinction was seen, but the differences among rate constants were less pronounced for **2** than for **1**. The backbone amide protons of Arg-2 and Gln-13 of **2** exchanged relatively rapidly, suggesting that the disulfide linkage is less effective than a D-Pro-Gly linkage at stabilizing local β-sheet structure.

The conformational rigidity implied for **1** by the H/D exchange data suggests that this cyclic peptide is an excellent reference compound for quantitative NMR-based determination of β-sheet populations in flexible peptides such as **3**−**7**. We analyze the flexible peptides in terms of a two-state conformational equilibrium, unfolded (U) vs folded (F), where the latter is the β-sheet. K_{eq} can be estimated from any conformationally sensitive proton chemical shift via eq 1, where δ_{obs} is the chemical shift of the

$$K_{eq} = (\delta_{obs} - \delta_U)/(\delta_F - \delta_{obs}) \quad (1)$$

proton in question within the equilibrating peptide, δ_U is the chemical shift of that proton in the fully unfolded state, and δ_F is the chemical shift of that proton in the fully folded state. We have previously shown that **8**, the L-Pro diastereomer of **3**, is completely unfolded in water;[2d] therefore, we obtained δ_U values from **8**. Results presented above show that **1** is a good source of δ_F values.

Protons on the α-carbons of amino acid residues are very sensitive to backbone conformation; residues in β-sheets display downfield shifts relative to the random coil state.[19] We used Val-3, Orn-8 (or Lys-8), and Ile-10 to estimate K_{eq} for β-sheet formation by linear peptides **3**−**7** in aqueous solution, based on eq 1. These three residues were selected because all occur at hydrogen-bonded positions, which means that H_α is oriented away from the opposite strand in the folded state.[20] This H_α orientation should minimize the effect of any differences in side-chain packing between **1** and the folded state of the linear peptides (because of Tyr ring current effects, such packing differences could influence $\delta_{H\alpha}$ values). $\Delta G°$ for β-sheet formation was calculated from K_{eq}, and $\Delta\Delta G°$ values (Table 1) were calculated for **4**−**7** by comparison with D-proline peptide **3** (systematic error is presumably minimized in the $\Delta\Delta G°$ values).

The reasonably good agreement among the β-sheet population values deduced for Val-3, Orn-8 (or Lys-8), and Ile-10 of **3**, **4**, and **7** (Table 1) supports our hypothesis that these peptides display a two-state folding equilibrium. The small variations in population values within each peptide might indicate real differences in the folded population at each of the indicator residues or systematic error arising from the reference peptides. In particular, it is possible that cyclic peptide **1** and the folded state of **3** have different ensembles of side-chain orientations, even though **1** is an excellent model for the backbone of **3** in the folded state. The $\Delta\Delta G°$ values for **4** agree well with one another, as do the $\Delta\Delta G°$ values for **7**, but there is some variation of $\Delta\Delta G°$ in **5** and **6**, which could indicate more than two states in equilibrium. Nevertheless, comparisons among the average $\Delta\Delta G°$ values for **3**−**6** provide insight into the way in which variations in the two-residue loop influence overall β-sheet stability.

Table 1. Population (%) of the β-Hairpin State for Each Linear Peptide and $\Delta\Delta G°$ Values (kcal/mol) for Each Mutant Relative to Peptide **3**[a]

	Val-3		Orn-8[b]		Ile-10		
peptide	pop. (±2)	$\Delta\Delta G°$	pop. (±3)	$\Delta\Delta G°$	pop. (±2)	$\Delta\Delta G°$	av $\Delta\Delta G°$
3	70	–	65	–	59	–	–
4	46	+0.56 ±0.09	41	+0.53 ±0.14	39	+0.46 ±0.10	+0.52 ±0.11
5	30	+0.95 ±0.10	21	+1.1 ±0.1	30	+0.68 ±0.10	+0.91 ±0.11
6	31	+0.91 ±0.10	26	+0.90 ±0.16	31	+0.64 ±0.10	+0.81 ±0.12
7	70	0.0 ±0.10	62	+0.07 ±0.14	59	+0.01 ±0.10	+0.03 ±0.11

[a] The population of the β-hairpin state for each linear peptide was calculated at each indicator residue from $\delta_{H\alpha}$ data as $(\delta_{obs} - \delta_U)/(\delta_F - \delta_U) \times 100$. $\Delta\Delta G°$ for each mutant relative to peptide **3** was calculated in the usual way (K_{eq} for β-hairpin formation from eq 1; $\Delta G° = -RT \ln K_{eq}$; $\Delta\Delta G° = \Delta G°(\text{mutant}) - \Delta G°(\mathbf{3})$). Experimental uncertainties reflect error propagation from the ±0.01 ppm uncertainty in the $\delta_{H\alpha}$ measurements. [b] In peptide **7**, residue 8 is lysine rather than ornithine.

Several groups have shown that the L-Asn-Gly loop promotes formation of small β-sheets in aqueous solution.[3−5] The data for **3** and **4** in Table 1 show that D-Pro-Gly provides ca. 0.5 kcal/mol stabilization relative to L-Asn-Gly, which is consistent with previous qualitative conclusions.[2d] Replacing D-Pro in **3** with Gly in **5** destabilizes the β-sheet conformation by ca. 0.9 kcal/mol.

We examined the D-Asn-Gly loop (**6**) in order to elucidate the origin of β-sheet promotion by L-Asn-Gly. Two-residue loops should adopt type I′ or II′ β-turn conformations in order to promote sheet interactions between flanking segments, because these β-turns are compatible with the right-handed twist of the strands.[21] In general, type I′ and II′ β-turns are less stable than type I and II β-turns for segments containing L-residues.[22] L-Asn-Gly was identified as a potential β-sheet promoter[3−5] because this segment is the most likely of all proteinogenic sequences to be involved in a type I′ turn.[23] This trend could arise because L-Asn-Gly intrinsically prefers a type I′ turn conformation, or because L-Asn-Gly is least resistant to type I′ turn formation among proteinogenic sequences. Our observation that D-Asn-Gly peptide **6** forms a less stable β-sheet than does diastereomer **4** demonstrates that L-Asn-Gly intrinsically prefers the type I′ turn.

The stabilities of the β-hairpins formed by **3** and by the ornithine → lysine mutant **7** are indistinguishable. This result shows that the non-proteinogenic ornithine residue does not exert any unusual conformational effect. The population and $\Delta\Delta G°$ values for **3**−**7** calculated using δ_F values from disulfide-cyclized **2** rather than from **1** are indistinguishable from those in Table 1.

Our comparison of disulfide and backbone cyclization shows that both strategies for end-to-end connection can stabilize a small increment of antiparallel β-sheet sufficiently to provide a "pure" folded state reference in aqueous solution. Backbone cyclization via a D-Pro-Gly segment, as in **1**, produces a particularly stable two-stranded β-sheet. The ability to lock-in the β-sheet conformation provides a basis for quantifying the folding equilibrium of flexible peptides, which in turn allows thermodynamic analysis of the factors that control β-sheet folding preferences.[24]

Supporting Information Available: Resonance assignments and NOEs for peptides **1** and **2** (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA992733T

(18) (a) For leading references on use of H/D exchange to analyze peptide and protein conformation, see: Englander, S. W.; Kallenbach, N. R. *Q. Rev. Biophys.* **1984**, *16*, 521. Perrin, C. L.; Dwyer, T. J.; Rebek, J.; Duff, R. J. *J. Am. Chem. Soc.* **1990**, *112*, 3122. Arrington, C. B.; Teesch, L. M.; Robertson, A. D. *J. Mol. Biol.* **1999**, *285*, 1265. (b) For application of H/D exchange to β-sheet model systems, see: Nesloney, C. L.; Kelly, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 5836.

(19) (a) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311. (b) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647.

(20) Val-5 also occurs in a hydrogen-bonded position but cannot be used because this residue is followed by proline in reference peptides **1** and **8** but not in equilibrating peptides **4**−**6**. Proline profoundly influences the backbone conformation of the preceding residue: Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221.

(21) (a) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170. (b) Gunasekaran, K.; Ramakrishnan, F.; Balaram, P. *Protein Eng.* **1997**, *10*, 1131 and references therein.

(22) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1.

(23) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 2207.

(24) This research was supported by the National Science Foundation (CHE-9820952). J.F.E. was supported by a fellowship from the Ministerio de Educacion y Cultura (Spain) and the Fulbright Commission. The NMR spectrometers were purchased in part with grant NIH 1 S10 RR04981. The CD spectrometer and analytical ultracentrifuge are part of the UW Biophysics Instrumentation Facility (NSF BIR-9512577).