4236 *J. Am. Chem. Soc.* **1998,** *120,* 4236−4237

# Rules for Antiparallel $\beta$-Sheet Design: D-Pro-Gly Is Superior to L-Asn-Gly for $\beta$-Hairpin Nucleation[1]

Heather E. Stanger and Samuel H. Gellman*

*Department of Chemistry, University of Wisconsin*
*Madison, Wisconsin 53706*

*Received October 27, 1997*

Folded proteins display considerable conformational diversity at the tertiary structural level, but only a few secondary structures with long-range order are observed: helices ($\alpha$ and $3_{10}$) and sheets (parallel and antiparallel).[2a] Understanding the origins of helix and sheet conformational stability is therefore a prerequisite for understanding the origins of protein folding preferences. The $\alpha$-helix has been extensively examined with model systems, because there are well-established rules for designing peptides that adopt helical conformations in solution.[2b] These model systems allow one to determine the effects of sequence, length, side chain−side chain interactions, solvent, and other variables on helix stability in the absence of a tertiary context. Relationships between $\beta$-sheet stability and sequence, length, or other factors are much less well-documented, because it has been difficult to generate peptides that adopt $\beta$-sheet conformations without aggregating in aqueous solution.[3,4] This difficulty has also hindered de novo design of $\beta$-sheet-rich protein tertiary structures.[5]

The prospects for rational design of small, soluble $\beta$-sheet increments have been advanced by the identification of several short, linear peptides ($\leq$16 residues) that display partial $\beta$-hairpin formation in aqueous solution.[6] A $\beta$-hairpin contains two antiparallel strands and a short connecting loop, and it must be possible to specify the location and size of the loop if $\beta$-hairpins are to be useful model systems for exploring antiparallel $\beta$-sheet stability. Two strategies have recently been reported for induction of $\beta$-hairpin folding with a two-residue loop at a specific site: use of L-Asn-Gly[7] or D-Pro-Xxx[8,9] as the loop sequence. Both sequences are expected to favor "mirror image" $\beta$-turns (type I′ or type II′), which are strongly correlated with two-residue $\beta$-hairpin loops in crystalline proteins.[10] We pursued the D-proline strategy[8a−c] because proline in the $i$+1 position strongly stabilizes $\beta$-turn formation.[11] Ramírez-Alvarado et al.[7a] and de Alba et al.[7b] examined L-Asn-Gly because this segment (at positions $i$+1 and $i$+2) has the highest statistical correlation with type I′ $\beta$-turns in crystalline proteins.[12] Here we report a direct comparison of $\beta$-hairpin promotion by L-Asn-Gly and D-Pro-Gly segments. Our results indicate that D-proline-based strategies for $\beta$-hairpin nucleation are particularly well suited for design of small, stable $\beta$-sheet increments.

Three 12-residue peptides RYVEV$\underline{\text{X}}$G$\underline{\text{O}}$KILQ-$NH_2$ were employed for evaluation of $\beta$-hairpin promotion; residues other than X have L-configuration, and X = L-Asn ("NG"), D-Pro ("$^{\mathrm{D}}$PG"), or L-Pro ("$^{\mathrm{L}}$PG"). Under the conditions employed for our studies (100 mM sodium acetate buffer, pH 3.8), all three 12mers should bear a charge of $\geq$+3, which discourages aggregation. Analytical ultracentrifugation (1.2 mM) suggested molecular weights of ca. 1260 for $^{\mathrm{D}}$PG, 1320 for $^{\mathrm{L}}$PG, and 1330 for NG. In each case, the deduced MW is somewhat below the expected value (1416 for PG, 1431 for NG). This type of deviation has previously been reported for designed $\beta$-hairpins[6c] and may reflect nonideality arising from the net charge on the peptides.[13]

Figure 1 summarizes ROESY[14] data obtained for $^{\mathrm{D}}$PG in aqueous solution; these data indicate that the intended $\beta$-hairpin conformation is highly populated. Long-range NH−NH NOEs between Val-5 and Orn-8, and between Val-3 and Ile-10, along with $C_{\alpha}H-C_{\alpha}H$ NOEs between Glu-4 and Lys-9, and between Tyr-2 and Leu-11, provide strong evidence for the proposed backbone conformation (Figure 1A). Further support for the $\beta$-hairpin folding pattern is provided by a network of side chain−side chain NOEs between Tyr-2 and Leu-11, and between Tyr-2 and Lys-9 (Figure 1B). In the flat rendering of Figure 1, Tyr-2 and Lys-9 seem to be far apart, but these side chains should be brought together by the right-handed twist commonly observed between adjacent $\beta$-sheet strands.[15] Only one of the long-range NOEs observed for $^{\mathrm{D}}$PG appears to be inconsistent with the proposed $\beta$-hairpin conformation, between Ile-10 $\gamma$-$CH_2$ and either Val-3 $\beta$-CH, Val-5 $\beta$-CH, or Glu-4 $\beta$-$CH_2$ (these three resonances overlap; this NOE is not indicated in Figure 1B). This NOE presumably arises from equilibration between the highly populated $\beta$-hairpin and alternative conformations.

No long-range NOEs were observed for $^{\mathrm{L}}$PG, which indicates that simply switching the proline configuration completely disrupts $\beta$-hairpin folding. For NG, significant population of a $\beta$-hairpin with a two-residue loop at NG is indicated by the observation of

---

[1] In memory of Professor Harlan L. Goering, 1921−1997.

[2] (a) Creighton, T. E. *Proteins: Structures and Molecular Principles*, 2nd ed.; Freeman: New York, 1993. (b) For leading references on $\alpha$-helical model systems, see: Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95. Kemp, D. S.; Boyd, J. G.; Mundel, C. *Nature* **1991**, *352*, 451. For an important cautionary note, see: Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, *359*, 653.

[3] Small, engineered globular proteins have been used to examine residue propensities and context effects in $\beta$-sheet formation, but in these systems $\beta$-sheet formation depends on tertiary folding. (a) Kim, C. A.; Berg, J. M. *Nature* **1993**, *362*, 267. (b) Minor, D. L.; Kim, P. S. *Nature* **1994**, *367*, 660. (c) Minor, D. L.; Kim, P. S. *Nature* **1994**, *371*, 264. (d) Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33*, 5510. (e) Smith, C. K.; Regan, L. *Science* **1995**, *270*, 980. (f) Blasie, C. A.; Berg, J. M. *Biochemistry* **1997**, *36*, 6218.

[4] (a) Hartman, R.; Schwaner, R. C.; Herman, J. *J. Mol. Biol.* **1974**, *90*, 415. (b) Brack, A.; Orgel, L. E. *Nature* **1975**, *256*, 383. (c) Osterman, D. G.; Kaiser, E. T. *J. Cell. Biochem.* **1985**, *29*, 57. (d) DeGrado, W. F.; Lear, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7684.

[5] For leading references on $\beta$-sheet protein design efforts, see: Yan, Y.; Erickson, B. W. *Protein Sci.* **1994**, *3*, 1069. Quinn, T. P.; Tweedy, N. B.; Williams, R. W.; Richardson, J. S.; Richardson, D. C.*Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8747. Ilynia, E.; Roongta, V.; Mayo, K. H. *Biochemistry* **1997**, *36*, 5245.

[6] (a) Blanco, F. J.; Jiménez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887. (b) Constantine, K. L.; Mueller, L.; Andersen, N. H.; Tong, H.; Wandler, C. F.; Friedrichs, M. S.; Bruccoleri, R. E. *J. Am. Chem. Soc.* **1995**, *117*, 10841. (c) Blanco, F. J.; Rivas, G.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 584. (d) Searle, M. S.; Williams, D. H.; Rackman, L. C. *Nat. Struct. Biol.* **1995**, *2*, 999. (e) For hairpin folding induced by a non-peptide turn segment in aqueous solution, see: Nesloney, C. L.; Kelly, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 5836.

[7] (a) Ramírez-Alvarado, M.; Blanco, F. J.; Serrano, L. *Nat. Struct. Biol.* **1996**, *3*, 604. (b) de Alba, E.; Jimenez, M. A.; Rico, M. *J. Am. Chem. Soc.* **1997**, *119*, 175. (c) Maynard, A. J.; Searle, M. S. *J. Chem. Soc., Chem. Commun.* **1997**, 1297.

[8] (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) Karle, I. L.; Awasthi, S. K.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8189.

[9] For the use of D-residues in $\beta$-sheet protein design, see: (a) Reference 5c. (b) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342.

[10] (a) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170. (b) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221. (c) Sibanda, B. L.; Thornton, J. M. *J. Mol. Biol.* **1993**, *229*, 428.

[11] Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1.

[12] Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 2207.

[13] Sedimentation equilibrium studies of these three peptides at higher concentrations leads to decreased apparent molecular weight (3.7 mM $^{\mathrm{D}}$PG, 1050; 4.7 mM $^{\mathrm{L}}$PG, 1130; 4.8 mM NG, 1120), which supports our conclusion that the lower-than-expected apparent molecular weights of $^{\mathrm{D}}$PG, $^{\mathrm{L}}$PG, and NG result from electrostatic repulsion between peptide molecules. Plots of ln(abs) vs radius$^2$ were linear at 42 K rpm and 56 K rpm, at peptide concentrations in the range of 1.2−4.8 mM.

[14] Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811.

[15] Salemme, F. R. *Prog. Biophys. Mol. Biol.* **1983**, *42*, 95.

S0002-7863(97)03704-9 CCC: $15.00

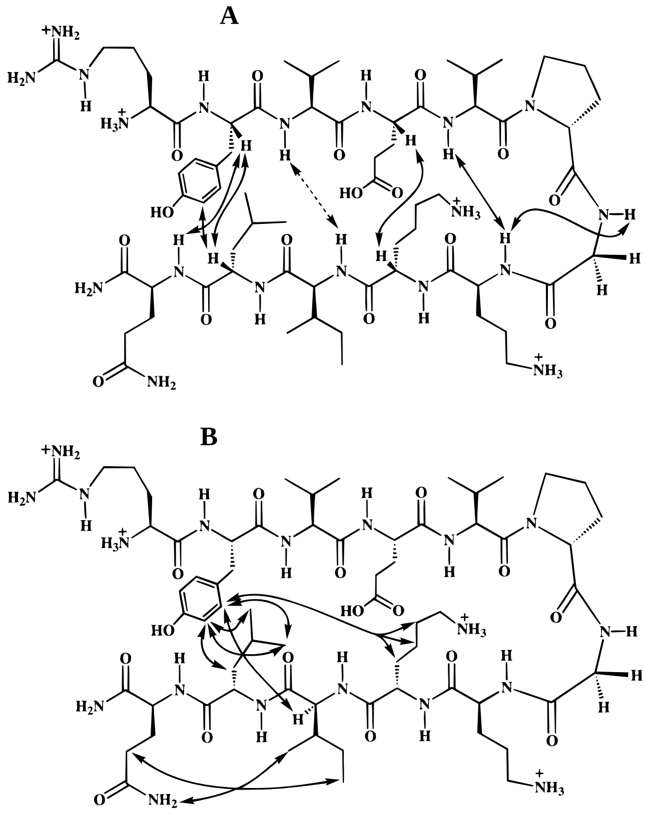

**Figure 1.** Long-range NOEs observed in ROESY analysis for 3.5 mM $^{\mathrm{D}}$PG in 100 mM aqueous sodium deuterioacetate buffer, pH 3.8 (uncorrected), 4 °C. (A) NOEs involving NH and/or $H_\alpha$ resonances ($H_\alpha$−$H_\alpha$ NOEs were observed in $D_2O$, and the remaining NOEs in 9:1 $H_2O$:$D_2O$). The dotted line for the NH−NH NOE between Val-3 and Ile-10 indicates that this NOE was weak. (B) NOEs involving side chains (obtained in 9:1 $H_2O$:$D_2O$). Only unambiguously assigned NOEs are shown in A and B. One additional long-range NOE was observed (see text), but resonance overlap interferred with assignment. Resonance assignments were obtained from a combination of COSY and TOCSY data, and sequential NOEs from ROESY data.

long-range NOEs corresponding to many of those shown for $^{\mathrm{D}}$PG in Figure 1.[16] However, the absence of the $C_\alpha H$−$C_\alpha H$ NOE between Tyr-2 and Leu-11, and the smaller number of NOEs between side chain pairs Tyr-2/Leu-11 and Tyr-2/Lys-9 suggest that the $\beta$-hairpin adopted by NG is less stable than that adopted by $^{\mathrm{D}}$PG. This conclusion is reinforced by the observation that NG displays numerous long-range NOEs that are not consistent with $\beta$-hairpin folding, including NH−NH ($i$, $i$+1) NOEs for Glu-4/Val-5, Val-5/Asn-6, Asn-6/Gly-7, and Ile-10/Leu-11.

Figure 2 compares $\Delta\delta_{\alpha H}$ (i.e., observed $\delta_{\alpha H}$ − random coil $\delta_{\alpha H}$) data for $^{\mathrm{D}}$PG, $^{\mathrm{L}}$PG, and NG. This measurement has previously been used to probe secondary structure in proteins and designed peptides.[6,7,17] Downfield shifts of $\delta_{\alpha H}$ relative to random coil ($\Delta\delta_{\alpha H} \geq +0.1$) indicate $\beta$-sheet,[17] and this criterion suggests that the segments Tyr-2 to Val-5 and Orn-8 to Ile-10 of $^{\mathrm{D}}$PG have adopted $\beta$-strand conformations. For $^{\mathrm{L}}$PG, in contrast, there is no clear $\Delta\delta_{\alpha H}$ trend, suggesting that this peptide does not favor a well-defined folding pattern in solution. The $\Delta\delta_{\alpha H}$ data for NG follow the $^{\mathrm{D}}$PG trends for many residues, but the downfield shifts of $\delta_{\alpha H}$ are always smaller (and sometimes nonexistent) for NG. Thus, the $\Delta\delta_{\alpha H}$ data support the conclusion that L-Asn-Gly is a weaker promoter of $\beta$-hairpin folding than is D-Pro-Gly.

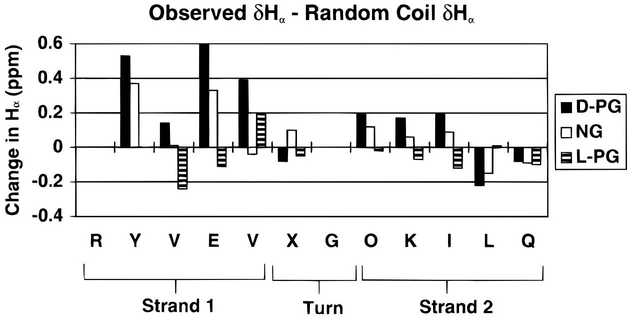

**Figure 2.** $\Delta\delta_{\alpha H}$ = observed $\delta_{\alpha H}$ − random coil $\delta_{\alpha H}$ for 3.5 mM $^{\mathrm{D}}$PG (solid bars), 4.6 mM NG (open bars) and 3.1 mM $^{\mathrm{L}}$PG (striped bars) in aqueous (9:1 $H_2O$:$D_2O$) sodium deuterioacetate buffer, pH 3.8 (uncorrected), 4 °C. (See ref 17 for origin of random coil values.) The reported random coil for lysine was used for ornithine. This extrapolation is supported by the observation that $\delta_{\alpha H}$ for Orn-8 in $^{\mathrm{L}}$PG is very near the lysine random coil value. No data are shown for Arg-1 because the N-terminus is uncapped. No data are shown for Gly-7 because there are two $H_\alpha$ resonances for $^{\mathrm{D}}$PG and NG; for all three peptides, the Gly-7 $\Delta\delta_{\alpha H}$ values lie between +0.07 and −0.21. Chemical shifts were externally referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS).

Circular dichroism data[16] provide further evidence that the D-Pro-Gly segment is a stronger $\beta$-hairpin inducer than is L-Asn-Gly. $^{\mathrm{D}}$PG shows a well-defined $\beta$-sheet signature, with a minimum at 217 nm and a maximum at 201 nm. The data for $^{\mathrm{L}}$PG suggest that random coil is predominant. The NG spectrum suggests the presence of significant proportions of both $\beta$-sheet (apparent minimum around 217 nm) and random coil (increasingly negative ellipticity below 200 nm).

Our data show that the D-Pro-Gly segment is a very strong promoter of $\beta$-hairpin formation and superior to the L-Asn-Gly segment in this regard. This finding has important implications for the design of short peptides intended to display $\beta$-hairpin folding in aqueous solution, which is crucial for development of model systems that elucidate $\beta$-sheet conformational preferences. The utility of the D-Pro-Gly segment for fundamental $\beta$-sheet analysis is enhanced by the fact that swapping L-proline for D-proline provides a negative control, since L-proline appears to be antithetical to formation of two-residue $\beta$-hairpin loops.[18]

**Supporting Information Available:** Summary of CD data and long-range NOEs for NG (4 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA973704Q

---

[16] Long-range NOEs observed for NG and CD data for all peptides are summarized in the Supporting Information.

[17] (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. (c) The "random coil" values used for calculation of $\Delta\delta_{\alpha H}$ in our study are given in: Wüthrich, K. *NMR of Proteins and Nucleic Acids*, Wiley: New York, 1986.

[18] This research was supported by the Army Research Office. H.E.S. is grateful for a National Research Service Award (T32 GM08923) from NIH. Peptides were characterized with a MALDI-TOF mass spectrometer purchased in part with NSF CHE-9520868. NMR spectrometers were purchased in part through grants NSF CHE-8813550 and NIH 1 S10 RR04981. We thank Dr. D. McCaslin, of the UW Biophysics Instrumentation Facility, for assistance with the CD and centrifugation measurements; this facility is supported by NSF BIR-9512577.