Stereochemical Analysis of Higher α,α-Dialkylglycine Containing Peptides. Characterization of Local Helical Conformations at Dipropylglycine Residues and Observation of a Novel Hydrated Multiple β-Turn Structure in Crystals of a Glycine Rich Peptide

Isabella L. Karle,*,§ Ramesh Kaul,⊥,‡ R. Balaji Rao,⊥ S. Raghothama,† and P. Balaram*,‡

Contibution from the Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375-5341, Department of Chemistry, Banaras Hindu University, Varanasi 221005, India, Sophisticated Instruments Facility, Indian Institute of Science, Bangalore 560012, India, and Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Received February 24, 1997[⊗]

Abstract: The peptide Boc-Gly-Dpg-Gly-Gly-Dpg-Gly-NHMe (**1**) has been synthesized to examine the conformational preferences of Dpg residues in the context of a poor helix promoting sequence. Single crystals of **1** were obtained in the space group $P2_1/c$ with $a = 13.716(2)$ Å, $b = 12.960(2)$ Å, $c = 22.266(4)$ Å, and $\beta = 98.05(1)°$; $R = 6.3\%$ for 3660 data with $|F_o| > 4\sigma$. The molecular conformation in crystals revealed that the Gly(1)-Dpg(2) segment adopts ϕ, ψ values distorted from those expected for an ideal type II′ β-turn ($\phi_{Gly(1)} = +72.0°$, $\psi_{Gly(1)} = -166.0°$; $\phi_{Dpg(2)} = -54.0°$, $\psi_{Dpg(2)} = -46.0°$) with an inserted water molecule between Boc-CO and Gly(3)NH. The Gly-(3)-Gly(4) segment adopts ϕ, ψ values which lie broadly in the *right handed* helical region ($\phi_{Gly(3)} = -78.0°$, $\psi_{Gly(3)} = -9.0°$; $\phi_{Gly(4)} = -80.0°$, $\psi_{Gly(4)} = -18.0°$). There is a chiral reversal at Dpg(5) which takes up ϕ, ψ values in the *left handed* helical region. The Dpg(5)-Gly(6) segment closely resembles an ideal type I′ β-turn ($\phi_{Dpg(5)} = +56.0°$, $\psi_{Dpg(5)} = +32.0°$; $\phi_{Gly(6)} = +85.0°$, $\psi_{Gly(6)} = -3.0°$). Molecules of both chiral senses are found in the centrosymmetric crystal. The C-terminus forms a hydrated Schellman motif, with water insertion into the potential $6 \rightarrow 1$ hydrogen bond between Gly(1)CO and Gly(6)NH. NMR studies in $CDCl_3$ suggest substantial retention of the multiple turn conformation observed in crystals. In solution the observed NOEs support *local helical* conformation at the two Dpg residues.

Introduction

α,α-Dialkylated amino acids have been widely used to introduce backbone conformational restraints in synthetic peptides.[1,2] The prototype residue α-aminoisobutyric acid (Aib)[3] stabilizes helical conformations in diverse sequences.[4,5] The higher dialkyl glycines like diethylglycine (Deg), di-*n*-propylglycine (Dpg), and di-*n*-butylglycine (Dbg) have been less extensively investigated. Initial interest in the use of higher α,α-dialkyl residues in peptide design stemmed from the fact that fully extended ($\phi \sim \psi \sim 180°$, C_5) conformations were observed in short homooligopeptides,[6−9] an observation rationalized by theoretical calculations which suggested energy minima in both helical ($\phi = \pm 60°$, $\psi = \pm 30°$) and fully extended ($\phi \sim \psi \sim 180°$) regions of ϕ, ψ space.[10] This situation was in marked contrast to the prototype α,α-dialkyl residue Aib, which appeared an almost obligatory helical residue.[4,5] Subsequent studies with heteromeric sequences of variable length revealed several examples of Dpg and its homolog α,α-dibutylglycine (Dbg) in helical conformations.[11−15] A particular interesting example is the coexistence of both conformations in single crystals of the tripeptide Boc-Leu-Dpg-Val-OMe.[16] Theoretical calculations suggest that for higher dialykylglycines pronounced energy minima exist in both the helical and fully extended regions of the conformational space. There appears

† Sophisticated Instruments Facility, Indian Institute of Science.
‡ Molecular Biophysics Unit, Indian Institute of Science.
§ Naval Research Laboratory.
⊥ Banaras Hindu University.
⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1997.

(1) Balaram, P. *Curr. Opin. Struct. Biol.* **1992**, *2*, 845−851.
(2) Benedetti, E.; Toniolo, C. *Polymeric Materials Encyclopedia* **1996**, *8*, 6472−6481.
(3) (a) Abbreviations used: Aib, α-aminoisobutyric acid; Deg, α,α-diethylglycine; Dpg, α,α-di-*n*-propylglycine; Dbg, α,α-di-*n*-butylglycine; Boc, *tert*-butyloxycarbonyl; OMe, methyl ester; OBzl, benzyl ester; NHMe, methylamide. (b) IUPAC-IUB Commission on Biochemical Nomenclature. Biochemistry **1970**, *9*, 3471−3479.
(4) Prasad, B. V. V.; Balaram, P. *CRC Crit. Rev. Biochem.* **1984**, *16*, 307−347.
(5) Karle, I. L.; Balaram, P. *Biochemistry* **1990**, *29*, 6747−6756.
(6) Benedetti, E.; Toniolo, C.; Hardy, P. M.; Barone, V.; Bavoso, A.; Di Blasio, B.; Grimaldi, P.; Lelj, F.; Pavone, V.; Pedone, C.; Bonora, G. M.; Lingham, I. *J. Am. Chem. Soc.* **1984**, *106*, 8146−8152.
(7) Bonora, G. M.; Toniolo, C.; DiBlasio, B.; Pavone, V.; Pedone, C.; Benedetti, E.; Lingham, I.; Hardy, P. M. *J. Am. Chem. Soc.* **1984**, *106*, 8152−8156.
(8) Benedetti, E.; Barone, V.; Bavoso, A.; DiBlasio, B.; Lelj, F.; Pavone, V.; Pedone, C.; Bonora, G. M.; Toniolo, C.; Leplawy, M. T.; Kaczmarek, K.; Redlinski, A. *Biopolymers* **1988**, *27*, 357−371.
(9) Toniolo, C.; Benedetti. E. *Macromolecules* **1991**, *24*, 4004−4009.
(10) Barone, V.; Lelj, F.; Bavoso, A.; DiBlasio, B.; Grimaldi, P.; Pavone, V.; Pedone, C. *Biopolymers* **1985**, *24*, 1759−1767.
(11) DiBlasio, B.; Pavone, V.; Isernia, C.; Pedone C.; Benedetti, E.; Toniolo, C.; Hardy, P. M.; Lingham, I. *J. Chem. Soc, Perkin Trans. 2* **1992**, *2*, 523−526.
(12) Karle, I. L.; Rao R. B.; Prasad, S.; Kaul, R.; Balaram, P. *J. Am. Chem. Soc.* **1994**, *116*, 10355−10361.
(13) Karle, I. L.; Gurunath, R.; Prasad, S.; Kaul, R.; Rao R. B.; Balaram, P. *J. Am. Chem. Soc.* **1995**, *117*, 9632−9637.
(14) Karle, I. L.; Rao R. B.; Kaul, R.; Prasad, S.; Balaram, P. *Biopolymers* **1996**, *39*, 75−84.
(15) Karle, I. L.; Gurunath, R.; Prasad, S.; Rao R. B.; Balaram, P. *Int. J. Peptide Protein Res.* **1996**, *47*, 376−382.
(16) Prasad S.; Mitra S.; Subramanian, E.; Velmurugan, D.; Rao, R. B.; Balaram, P. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 424−430.

S0002-7863(97)00596-9 CCC: $14.00

to be a pronounced dependence on the bond angle N−C$^{\alpha}$−CO, τ (with $\tau \sim 110°$ in helices and $\tau \sim 103°$ in fully extended structures).[2,8] Available data thus suggests that the conformations of the higher dialkylglycines may be modulated by subtle environmental effects. In order to probe systematically the relationship between backbone conformation and sequence context, we have chosen to conformationally characterize a wide range of Deg/Dpg/Dbg containing peptides. In this report, we describe the conformation in crystals and in solution of the glycine rich protected hexapeptide Boc-Gly-Dpg-Gly-Gly-Dpg-Gly-NHMe (**1**).

Glycine rich sequences have been chosen since glycine is the most conformationally flexible residue and has a relatively low propensity for occurring in helical structures.[17] The juxtaposition of Gly and the α,α-dialkylated residue Aib in short sequences has led to interesting stereochemical consequences in both designed[18] and natural sequences like trichogin A IV[19] and synthetic fragments.[20,21] X-ray diffraction studies establish that both Dpg residues at position 2 and 5 favor *local helical* conformations in crystals. The peptide adopts a novel, hydrated multiple turn structure. In the apolar solvent $CDCl_3$, NMR data are consistent with a major population of conformers resembling that observed in crystals, while in the strongly interacting solvent $(CD_3)_2SO$, a mixed population of partially folded and extended structures is supported by the experimental evidence.

Experimental Section

The peptide was assembled by conventional solution phase procedure using a racemization free, fragment condensation strategy. The Boc group was used for N-terminal protection and the C-terminus was protected as a methyl ester (OMe). Glycine was protected as a benzyl ester (OBzl). Deprotections were performed using 98% formic acid, while saponifications were carried out using 4 N sodium hydroxide solution and methanol. Couplings were mediated by dicyclohexylcarbodiimide-1-hydroxybenzotriazole (DCC/HOBT). Dpg and Dpg·OMe·HCl were synthesized as described previously.[22] All intermediate peptides were characterized by ^{1}H NMR (80 MHz, 400 MHz), thin layer chromatography (TLC) and used without further purification.

Synthesis of Peptides. Boc-Gly-Dpg-OMe (1). Boc-Gly-OH (5.25 g, 30 mmol) was dissolved in dichloromethane (DCM) (25 mL). H_2N-Dpg-OMe (5.20 g, 30 mmol) obtained from its ester hydrochloride was added, followed by DCC (6.0 g, 30 mmol). The reaction mixture was stirred at room temperature for 2 days. DCM was removed in vacuo. The residue was taken up in ethyl acetate (about 25 mL). The precipitated dicyclohexylurea (DCU) was filtered. The organic layer was washed with an excess of brine solution, 2 N HCl (2 × 50 mL), 1 M sodium carbonate solution (2 × 50 mL), and again with brine. The solution was then dried over anhydrous sodium sulfate and evaporated in vacuo. The dipeptide **1** was obtained as light yellow gum weighing 7.6 g (23 mmol, 76%): 80 MHz ^{1}H NMR ($CDCl_3$, δ ppm) 0.8, 0.93 (6H, t, Dpg C$^{\delta}$H$_3$), 1.26 (4H, m, Dpg C$^{\gamma}$H$_2$), 1.5 (9H, s, Boc CH_3), 2.4, 2.3 (4H, m, Dpg C$^{\beta}$H$_2$), 3.70 (3H, s, OCH_3), 3.75 (2H, d, Gly C$^{\alpha}$H), 5.50 (1H, t, Gly NH), 6.75 (1H, s, Dpg NH).

Boc-Gly-Dpg-OH (2). Peptide **1** (7.6 g, 23 mmol) was saponified using MeOH (25 mL) and 4 N NaOH (10 mL). The reaction mixture was stirred at room temperature, and its course followed by TLC. After 4 days, MeOH was evaporated, and the residue taken in water. The aqueous solution was washed with ether (2 × 40 mL). The aqueous layer was neutralized with 2 N HCl and extracted with ethyl acetate. The ethyl acetate extract was dried over anhydrous sodium sulfate and evaporated *in vacuo* to yield 5.68 g (18 mmol, 78%) of dipeptide acid **2**.

Boc-Gly-Dpg-Gly-OBzl (3). 3.16 g (10 mmol) of peptide **2** was dissolved in DMF (5 mL). Glycine benzyl ester-*p*-toluene sulfonate (6.74 g, 20 mmol) was added to it followed by DCC (2.0 g,10 mmol) and HOBt (1.35 g, 10 mmol). The reaction was stirred at room temperature for 3 days. DCU was filtered, and the work up was similiar to one described for **1**. The peptide **3** was obtained as a yellowish solid weighing 3.9 g (8.5 mmol, 85%): 400 MHz ^{1}H NMR ($CDCl_3$, δ ppm) 0.8, 0.95 (6H, t, Dpg C$^{\delta}$H$_3$), 1.12, 1.25 (4H, m, Dpg C$^{\gamma}$H$_2$), 1.45 (9H, s, Boc CH_3), 1.6, 2.45 (4H, m, Dpg C$^{\beta}$H$_2$), 3.75 (2H, d, Gly C$^{\alpha}$H), 4.12 (2H, d, Gly C$^{\alpha}$H), 5.14 (1H, t, Gly NH), 5.20 (2H, s, $-CH_2$), 6.5 (1H, t, Gly NH), 7.18 (1H, s, Dpg NH), 7.35 (5H, m, Phe).

Boc-Gly-Dpg-Gly-OH (4). Peptide **3** (1.85 g, 4 mmol) was saponified using 15 mL of MeOH and 4 N NaOH (8 mL). The reaction was monitored by TLC. After 3 days the reaction was worked up as described for **2** to yield 1.3 g (3.5 mmol, 87.5%) of the tripeptide acid **4** as a white solid.

H_2N-Gly-Dpg-Gly-OBzl (5). **3** (1.4 g, 3 mmol) was taken in 8 mL of 98% formic acid, and the reaction mixture tightly stoppered. The reaction was monitored by TLC. After 8 h, formic acid was evaporated in vacuo, and the residue was dissolved in water. The solution was washed with ether (2 × 30 mL). The aqueous layer was neutralized with sodium carbonate solution and extracted with ethyl acetate. The organic extract was dried over anhydrous sodium sulfate and evaporated *in vacuo* to yield 0.72 g (2 mmol, 66%) of free amine tripeptide **5** as a gum.

Boc-Gly-Dpg-Gly-Gly-Dpg-Gly-OBzl (6). **4** (0.75 g, 2 mmol) was dissolved in DMF (4 mL) and to it was added 0.72 g (2 mmol) of **5**, followed by DCC (0.4 g, 2 mmol) and HOBt (0.27 g, 2 mmol). The reaction was stirred at room temperature for 5 days. DCU was filtered and the work up was similar to one described for **1** to yield 1.14 g (1.6 mmol, 80%) of the hexapeptide ester **6** as a white solid: mp 136−138 °C; 400 Mhz ^{1}H NMR ($CDCl_3$, δ ppm) 0.88−0.9 (12H, t, Dpg 2,5 C$^{\delta}$H$_3$), 1.2−1.3 (8H, m, Dpg 2,5 C$^{\gamma}$H$_2$), 1.49 (9H, s, Boc CH_3), 1.85 (4H, m, Dpg 2 C$^{\beta}$H$_2$), 2.05 (4H, m, Dpg 5 C$^{\beta}$H$_2$), 3.65 (2H, d, Gly 6 C$^{\alpha}$H), 3.85−3.95 (2H, d, Gly 3/Gly 4 C$^{\alpha}$H), 4.10 (2H, d, Gly 1 C$^{\alpha}$H), 5.25 (2H, s, -CH_2), 5.65 (1H, t, Gly 1 NH), 6.78 (1H, s, Dpg 2 NH), 6.89 (1H, s, Dpg 5 NH), 6.92 (1H, t, Gly 6 NH), 7.45 (5H, m, Phe), 7.41 (1H, t, Gly 3 NH), 7.82 (1H, t, Gly 4 NH).

Boc-Gly-Dpg-Gly-Gly-Dpg-Gly-NHMe (7). Peptide **6** (0.5 g, 7 mmol) was dissolved in dry MeOH (10 mL), and methyl amine (CH_3-NH_2) gas was passed till saturation (about 2 h). MeOH was evaporated, and the peptide **7** was obtained as a yellow solid: yield 0.3 g (0.5 mmol, 71%); mp = 141−143 °C.

The crude peptide was purified on a reverse phase MPLC C_{18} column (40−60 μ) using a gradient of MeOH/H_2O. Homogeneity of the peptide was susequently demonstrated by analytical HPLC on a reverse phase (C18, 5 μ) column. The peptide was characterized by complete assignment of 400 MHz ^{1}H NMR spectra, by 125 MHz ^{13}C NMR, and by its FAB mass spectrum: 400 MHz ^{1}H NMR ($CDCl_3$, δ ppm) 0.89−0.93 (12H, t, Dpg 2/5 C$^{\delta}$H$_3$), 1.21−1.3 (8H, m, Dpg 2/5C$^{\gamma}$H$_2$), 1.45 (9H, s, Boc CH_3), 1.85, 1.75 (4H, m, Dpg 2 C$^{\beta}$H$_2$), 2.10, 1.83 (4H, m, Dpg 5 C$^{\beta}$H$_2$), 2.78 (3H, d, NH 7 **Me**), 3.88−3.86 (2H, d, Gly 3/Gly 4/Gly 6 C$^{\alpha}$Hs), 3.66 (2H, d, Gly 1 C$^{\alpha}$H), 5.60 (1H, t, Gly 1 NH), 6.68 (1H, s, Dpg 2 NH), 6.96 (1H, s, Dpg 5 NH), 7.13 (1H, t, Gly 6 NH), 7.53 (1H, t, Gly 3 NH), 7.92 (1H, t, Gly 4 NH), 7.24 (NH 7 Me). 125 MHz ^{13}C NMR ($CDCl_3$, δ ppm): 14.06−14.09 (4C, C$^{\delta,\delta'}$ Dpg 2/5), 16.45−16.56 (4C, C$^{\gamma,\gamma}$ Dpg 2/5), 26.06 (1C, $NHCH_3$), 28.11 (3C, Boc CH_3s), 35.11, 35.71 (4C, C$^{\beta,\beta'}$ Dpg 2/5), 43.19, 43.84, 44.05, 45.30 (4C, C$^{\alpha}$ Gly 1/3/4/6), 62.60, 63.06 (2C, C$^{\alpha}$ Dpg 2/5), 81.15 [1C, $(CH_3)_3$C], 157.3 (1C, Boc C=O), 170.22, 170.59, 170.63, 171.37, 173.97, 174.67 (6C, C=O); mass spectral data M + Na^+ = 665, M_{calcd} = 642.

Spectroscopic Studies. All NMR studies were carried out on a Bruker AMX-400 spectrometer. Peptide concentrations were in the range of 7−8 mM and the probe temperature was maintained at 298 K. Resonance assignments were done using two-dimensional ROESY

(17) Richardson, J. S.; Richardson, D. C. In Prediction of Protein Structure and Principles of Protein Conformation. Fasman, G. D., Ed.; Plenium Press: New York, pp 1−95.

(18) Karle, I. L.; Banerjee, A.; Bhattacharjya, S.; Balaram, P. *Biopolymers* **1996**, *38*, 515−526.

(19) Toniolo, C.; Peggion, C.; Crisma, M.; Formaggio, F.; Shui, X.; Eggleston, D. S. *Nature Struct. Biol*. **1994**, *1*, 908−914.

(20) Gurunath, R.; Balaram, P. *Biopolymers* **1995**, *35*, 21−29.

(21) Monaco, V.; Formaggio, F.; Crisma, M.; Toniolo, C.; Shui, X.; Eggleston, D. S. *Biopolymers* **1996**, *39*, 31−42.

(22) Prasad, S.; Rao, R. B.; Balaram, P. *Biopolymers* **1995**, *35*, 11−20.

Table 1. Crystal and Diffraction Parameters

empirical formula	$C_{30}H_{59}N_7O_{10}$
color, habit	colorless, very fragile wafers
crystal size (nm)	0.05 × 0.40 × 1.3
crystal system	monoclinic
space group	$P2_1/c$
unit cell dimensions	$a = 13.716(2)$ Å
	$b = 12.960(2)$ Å
	$c = 22.266$ Å
	$\beta = 98.050(10)°$
volume	3918.2(11) Å^3
Z	4
formula weight	677.8
density (calc)	1.149 Mg/m^3
absorption coeff	0.713 mm^{-1}
F(000)	1472

Table 2. Torsional Angles[a] (deg) for Peptide **1**

residue	ϕ	ψ	ω	χ^1	χ^2
Boc			175		
Gly(1)	+72	−166	−171		
Dpg(2)	−54	−46	−173	177, +63	178, −170
Gly(3)	−78	−9	172		
Gly(4)	−80	−18	−179		
Dpg(5)	+56	+32	179	179, −56	−179, 177
Gly(6)	+85	−3	−179		

[a] The torsional angles for rotation about bonds of the peptide backbone (ϕ,ψ,ω) and about the bonds in the Dpg side chains (χ^1, χ^2) are described in IUPAC-IUB commission on Biochemical Nomenclature.[3b]

spectra.[23] 2D data were collected in phase sensitive mode using the time proportional phase incrementation method. Data points numbering 512 with 64 transients were collected. Spectral widths were in the range of 4500 Hz with a spin lock time of 300 m s. Zero filling was done to yield data sets of 1K × 1K using square sine bell window.

X-ray Studies. Single crystals of peptide **1** were grown from methanol−water solution by slow evaporation. The crystal and diffraction parameters are listed in Table 1. X-ray data were collected on a Siemens automated diffractometer with Cu radiation in the $\theta/2\theta$ mode, constant scan speed of 10 deg/m, scan width of $1.3° + 2\theta$ (α_1−α_2), $2\theta_{max} = 116°$ (resolution 0.9 Å) for a total of 5363 unique reflections and 3360 reflections with intensities greater than $4\sigma(F)$. The structure was solved by direct phase determination. Full-matrix anisotropic least-squares refinement was performed on the parameters for the C, N, and O atoms. The positions of the four H atoms on W1 and W2 were found in a difference map and refined isotropically. The remaining 55 H atoms were placed in idealized positions and allowed to ride with C or N atom to which each was bonded. The final R values were $R = 6.33\%$ for 3360 data and $R_W = 6.40\%$.

Results and Discussion

Crystal Structure of 1. Figure 1 shows a view of the molecular conformation determined in crystals. The backbone conformational parameters are summarized in Table 2. While the hydrogen bond parameters are listed in Table 3. Since peptide **1** is *achiral* it crystallizes in a centrosymmetric space group and contains molecules of both hands (chiral senses). Only *one sign* of the dihedral angles is listed in Table 2, and only one hand is shown in Figure 1.

Inspection of the torsional angles reveals that the Gly(1)-Dpg(2) segment adopts ϕ, ψ values somewhat distorted from those expected for an ideal type II′ β-turn ($\phi_{i+1} = +60°$, $\psi_{i+1} = -120°$; $\phi_{i+2} = -80°$, $\psi_{i+2} = 0°$). Gly(3)-Gly(4) adopts ϕ, ψ values which lie broadly in the *right handed* helical region. There is a chiral reversal at Dpg(5) which now adopts ϕ, ψ values in the *left handed* helical region. The Dpg(5)-Gly(6) segment closely resembles an ideal type I′ β-turn ($\phi_{i+1} = +60°$, $\psi_{i+1} = +30°$; $\phi_{i+2} = +90°$, $\psi_{i+2} = 0°$). Inspection of the hydrogen bonding pattern suggests that the Dpg(2)-Gly(3)-Gly-(4) segment forms one turn of a 3_{10} helix stabilized by two intramolecular 4 → 1 hydrogen bonds[Gly(1)CO···Gly(4)NH, N···O 3.09 Å; Dpg(2)CO···Dpg(5)NH, N···O 3.34 Å]. The Dpg(5)-Gly(6) segment forms a type I′ β-turn stabilized by a 4 → 1 hydrogen bond between Gly(4)CO···NHMe, N···O 2.95 Å. At the N-terminus, a water molecule (W1) is inserted into the Gly(1)-Dpg(2) type II β-turn resulting in the absence of a 4 → 1 intramolecular hydrogen bond [Boc CO···Gly(3)NH, N···O 4.30 Å]. Water insertion into β-turn structures is commonly observed in peptides.[28] The observed distortion in backbone dihedral angles from the idealized values expected in type II/II′ β-turns is clearly a consequence of the water insertion into the backbone. Interestingly, a potential 6 → 1 hydrogen bond between Gly(1)CO···Gly(6)NH is disrupted by insertion of a water molecule (W2) resulting in an N···O distance of 4.60 Å. The Dpg(2)-Gly(3)-Gly(4)-Dpg(5) segment forms an $\alpha_R\alpha_R\alpha_R\alpha_L$ motif (α_R-residue in *right handed* helical conformation and α_L-residue in *left handed* helical conformation). Such conformational features have been termed as *Schellman motifs* and are frequently found at the C-terminus end of α-helical structures in proteins.[29−32] In synthetic

Table 3. Hydrogen Bonds

type	donor	acceptor	N···O (Å) O···O (Å)	H···O (Å) H···O (Å)
lateral, intermol	N1	O4[a]	3.147	2.43
head-to-tail	N2	O5[b]	2.893	2.00
	N3	W1	2.854	2.01
4 → 1	N4	O1	3.091	2.30
4 → 1	N5	O2	3.345[c]	2.51[c]
	N6	W2	2.886	2.10
4 → 1	N7	O4	2.954	2.08
	W1	O0	2.850	2.01
	W1	O6[d]	2.809	1.96
	W2	O1	2.756	1.67
	W2	O3[a]	2.865	2.16

[a] Symmetry equivalent at $2-x$, $0.5+y$, $1.5-z$. [b] Symmetry equivalent at x, $0.5-y$, $-0.5+z$. [c] Long values for hydrogen bonds. [d] Symmetry equivalent at $2-x$, $-0.5+y$, $1.5-z$.

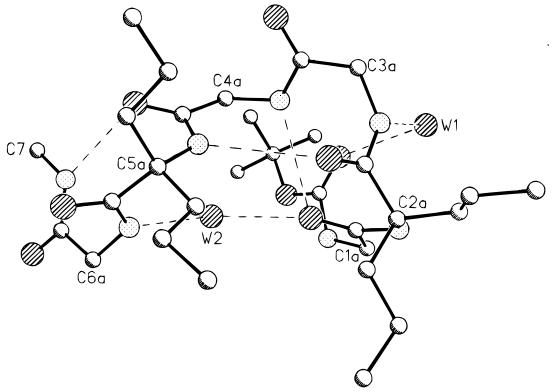

Figure 1. Conformation of the hexapeptide **1** in the crystal. Dashed lines indicate intramolecular hydrogen bonds. W1 and W2 are the water molecules inserted into the molecule.

(23) Wuthrich, K. NMR of Proteins and Nucleic Acids; John Wiley and Sons: New York, 1986.
(24) Kopple, K. D.; Schamper, T. J. *J. Am. Chem. Soc.* **1972**, *94*, 3644−3646.
(25) Kopple, K. D.; Go, A.; Pilipauskas, D. R. *J. Am. Chem. Soc.* **1975**, *97*, 6830−6838.
(26) Pitner, T. P.; Urry, D. W. *J. Am. Chem. Soc.* **1972**, *94*, 1399−1400.
(27) Iqbal, M.; Balaram, P. *J. Am. Chem. Soc.* **1981**, *103*, 5548−5552.
(28) Karle, I. L. *Biopolymers* **1996**, *40*, 157−180.
(29) Schellman, C. In *Protein Folding*; Janenicke, R., Ed.; Elsevier/North Holland Biochemical Press: Amsterdam, 1980; pp 53−61.
(30) Milner-White, E. J. *J. Mol. Biol.* **1988**, *199*, 503−511.

Table 4. Observed Dpg Backbone Conformations in Peptide Crystal Structures

peptides		torsional angles[a] ϕ	ψ	τ	ref
Tfa-**Dpg**-DBH	mol. A	176.0	175.9	101.1	6
	mol. B	177.0	174.8	101.1	
Tfa-**(Dpg)$_2$**-DBH		179.5	169.4	103.0	6
		172.5	−179.4	102.3	
For-Met-**Dpg**-Phe-OMe		173.1	179.0	105.4	37
Tfa-**(Dpg)$_3$**-DBH		±54.8	±39.4	108.0	11
		±59.9	±10.2	116.0	
		±48.3	±32.9	111.0	
Boc Leu-**Dpg**-Val-OMe	mol. A	176.0	−180.0	102.0	16
	mol. B	62.8	39.6	111.0	
Boc-Ala-**Dpg**-Ala-OMe		66.2	19.3	113.4	40
Boc-Ala-**Dpg**-Ala-NHMe		−50.6	−14.2	113.4	40
Boc-Leu-**Dpg**-Leu-Ala-Leu-Aib-OMe	mol. A[b]	−55.0	−43.0	107.0	12
	mol. B[b]	−53.0	−43.0	105.0	
	mol. A[c]	−43.0	−37.0	113.0	
Boc-Aib-Ala-Leu-Ala-Leu-**Dpg**-Leu-Ala-Leu-Aib-OMe		−56.0	−47.0	109.0	12
Boc-Gly-**Dpg**-Ala-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe		−46.0	−35.0	112.7	13
Boc-Gly-**Dpg**-Leu-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe		−51.0	−47.0	110.6	13
Boc-Gly-**Dpg**-Pro-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe		−52.0	−51.0	109.9	13
Boc-Gly-**Dpg**-Gly-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe		−53.0	−50.0	109.1	14
Boc-Gly-**Dpg**-Gly-OH		178.0	171.0	104.1	39
Boc-Val-Val-Ala-Leu-Gly-**Dpg**-Gly-Val-Ala-Leu-Aib-Val-Ala-Leu-OMe		−52.0	−44.0	107.2	39
Boc-Gly-**Dpg**-Gly-Gly-**Dpg**-Gly-NHMe		−54.0	−46.0	109.9	this study
		56.0	32.0	110.8	

[a] Nomenclature follows ref 3b. τ is the bond angle N−C$^\alpha$−C′. [b] Triclinic. [c] Orthorhombic.

peptides, the presence of an achiral residue at the C-terminus of a helix frequently results in chiral reversal of ϕ, ψ values resulting in the formation of α_L terminated helices.[33−35] In the present structure, the use of achiral residues results in a reversal of screw sense at the C-terminus. Formally this structure may be viewed as a case where a right handed helical twist is immediately followed by a left handed helical segment. Indeed, such ambidextrous helical molecules have recently been characterized in crystals of a 14 residue peptide helix containing two contiguous helical segments of opposite chirality.[36]

Dpg Residue Conformations. Table 4 summarizes observed Dpg backbone conformations in crystals and also lists the bond angle at the C$^\alpha$ atom (τ). Theoretical calculations suggest a strong bond angle dependence of the backbone ϕ, ψ values with $\tau < 105°$ resulting in extended conformations and $\tau > 105°$ yielding helical conformations.[2,10] While the majority of examples listed in the table follow this correlation, τ values ∼105° appear to accomodate both types of local conformations. The number of examples of Dpg in helical structures is clearly growing in diverse sequence environments suggesting that Dpg and related residues may be largely helix stabilizers, with fully extended forms requiring specific "local factors" for stabilization.

Crystal Packing. The molecules of **1** contain multiple turns of varying conformational types, resulting in a structure that is more compact than a helix of comparable length. Nevertheless, the peptide CO and NH groups which are internally hydrogen bonded are the same as those that would have participated in such interactions in a 3_{10}/α-helical structures. The molecules

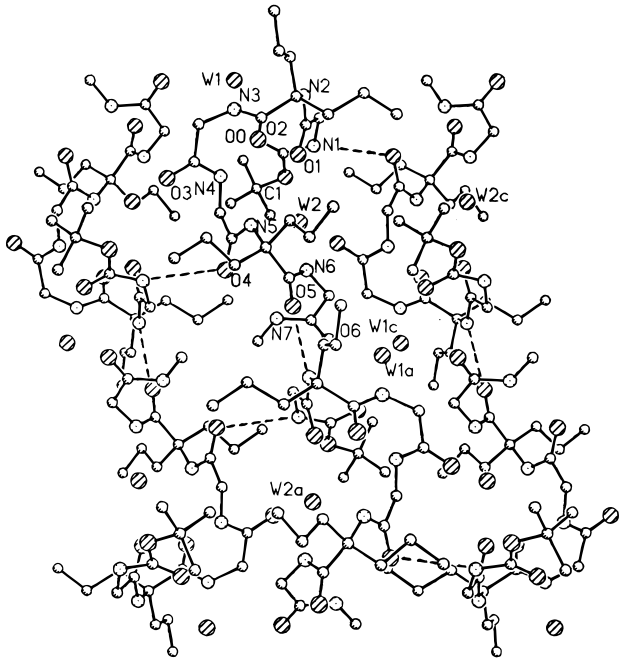

Figure 2. Packing of the distorted helices into sheets. Projection down the a axis. Dashed lines show the head-to-tail N2−O5 hydrogen bonds and the lateral N1−N4 hydrogen bonds between helices. The water molecules W1 and W2 also participate in forming lateral bridges of hydrogen bonds between helices (not shown).

assemble into columns by head-to-tail hydrogen bonding between N2 and O5, a motif adopted by almost all 3_{10}- and α-helices (Figure 2).[5] The columns pack into layers with antiparallel columns and lateral hydrogen bonds between N1 and O4. Water molecules W1 and W2, in addition to forming bridges between O0 and N3 as well as O1 and N6 in individual helices, also connect neighboring helices by lateral hydrogen bonds with O6 and O3, respectively. There are no hydrogen bonds between layers, that is, in the direction perpendicular to the view in Figure 2.

NMR Studies. Sequence specific assignments were easily achieved using sequential interresidue nuclear Overhauser effects

(31) Nagarajaram, H. A.; Sowdhamini, R.; Ramakrishnan, C.; Balaram, P. *FEBS Lett.* **1993**, *321*, 79−83.
(32) Aurora, R.; Srinivasan, R.; Rose, G. D. *Science* **1994**, *264*, 1126−1130.
(33) Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Balaram, P. *Int. J. Pept. Protein Res.* **1993**, *42*, 401−410.
(34) Rajashankar, K. R.; Ramakumar, S.; Mal, T. K.; Jain, R. M.; Chauhan, V. S. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 765−768.
(35) Banerjee, A.; Datta, S.; Pramanik, A.; Shamala, N.; Balaram, P. *J. Am. Chem. Soc.* **1996**, *118*, 9477−9483.
(36) Banerjee, A.; Raghothama, S.; Karle, I. L.; Balaram, P. *Biopolymers* **1996**, *39*, 279−285.

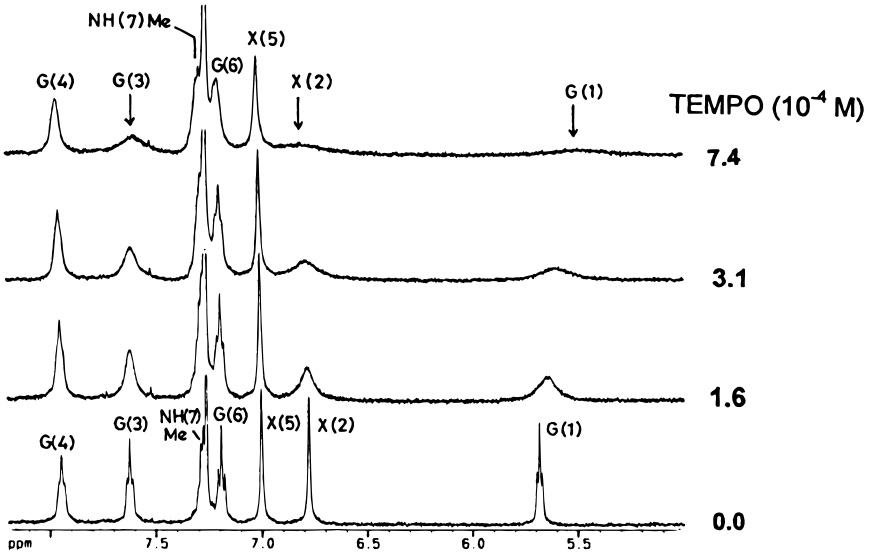

Figure 3. Partial 400 MHz ^{1}H NMR spectra in $CDCl_3$ showing the effect of TEMPO on the line width of NH resonance in peptide **1**. Resonance assignments are indicated. TEMPO concentrations (M) are marked on the spectra. Peptide concentration was 10.4 mM. [X = Dpg].

Table 5. NMR Parameters for Peptide **1**

	NH			
residue	$CDCl_3$	$(CD_3)_2SO$	$\Delta\delta^a$ (ppm)	$-d\delta/dt^b$ (ppb/K)
Gly (1)	5.60	7.15	0.89	8.2
Dpg(2)	6.68	7.73	1.15	4.9
Gly(3)	7.53	8.22	0.51	6.1
Gly(4)	7.92	8.03	0.20	4.4
Dpg(5)	6.96	7.73	0.30	3.8
Gly(6)	7.13	8.03	0.31	6.1
NH(7)Me	7.24	7.49	0.12	3.7

[a] $\Delta\delta$ is the chemical shift difference for NH protons in $CDCl_3$ and 21.05% $(CD_3)_2SO/CDCl_3$. [b] $d\delta/dt$ (ppb/K) values in $(CD_3)_2SO$.

measured using 2D ROESY methods.[23] The Gly, Dpg, and methylamide NH protons can be readily differentiated on the basis of their multiplet patterns. In both solvents, the methylamide NH proton was unambiguously assigned to the only quartet resonance, while the Gly(1)NH was easily identified by virtue of its high field position in both solvents. The chemical shifts of the backbone NH protons are summarized in Table 5. The number of intramolecularly hydrogen bonded NH groups were delineated using radical induced line broadening[24,25] and solvent induced perturbation of chemical shifts[26] in $CDCl_3$-$(CD_3)_2SO$ mixtures.[27] Figure 3 shows the effect of the free radical TEMPO on the line widths of the various NH resonances. It is clear that while Gly(1), Dpg(2), and Gly(3) NH groups are exposed to the solvent, the remaining four NH groups are inaccessible. All NH groups show downfield shifts on addition of small amounts of the hydrogen bonding solvent $(CD_3)_2SO$ to $CDCl_3$ solutions. The magnitude of chemical shift changes ($\Delta\delta$) on going from $CDCl_3$ to 21.05% $(CD_3)_2SO-CDCl_3$ are again diagnostic of peptide NH accessibility. From the data summarized in Table 5, it is evident that the first three NH groups also have larger $\Delta\delta$ values confirming their solvent accessible nature. It is thus clear that the peptide favors a folded conformation involving the NH groups of residues 4−6 and the terminal NHMe in intramolecular hydrogen bonding. Figure 4 shows the NH-NH region of the ROESY spectrum of **1** in $CDCl_3$, while Figure 5 illustrates observed $C^{\alpha}H \leftrightarrow NH$ NOEs. Sequential $N_iH \leftrightarrow N_{i+1}H$ (d_{NN}) NOEs are observed over the entire length of the peptide strongly indicating a continuous stretch of backbone ϕ, ψ values in the helical regions of conformational space. It may be noted that in short peptides NOEs are usually observed only up to distances of ≤3.5 Å.

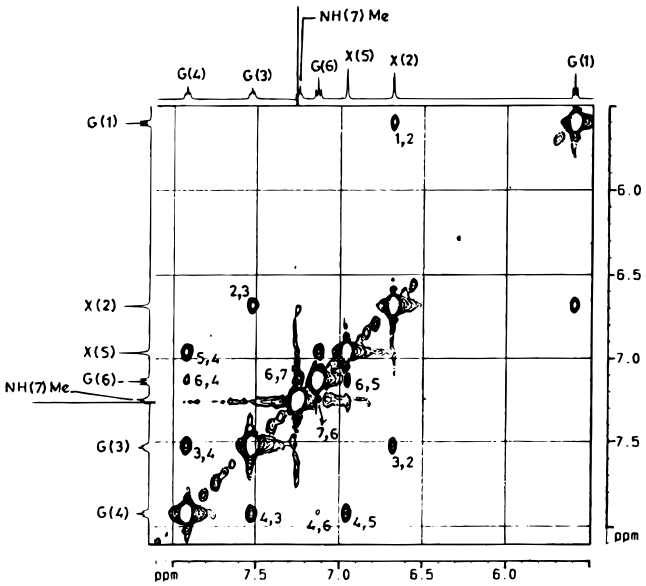

Figure 4. Partial 400 MHz ROESY spectrum of peptide **1** in $CDCl_3$ showing $N_iH \leftrightarrow N_{i+1}H$ NOEs. [X = Dpg].

Observation of sequential d_{NN} connectivities thus provides a strong constraint on the structure. Together with the fact that the four C-terminus NH groups are hydrogen bonded, the NOE results could, in principle, be interpreted in terms of a continuous α-helical conformation over the entire peptide segment as the predominant structure in $CDCl_3$ solution. This conclusion is however at variance with the observed crystal structure. Close examination of the solid state conformation suggests that residues Dpg(2), Gly(3), Gly(4), Dpg(5), and Gly(6) do indeed have ϕ, ψ values lying in the 3_{10}/α-helical regions of conformational space (Table 2) with the two C-terminal residues forming a chiral reversal resulting in a change of the helix handedness. Thus sequential d_{NN} connectivities over the 2−6 segment are in fact consistent with this conformation. Indeed the only "nonhelical" residue in the crystal state structure is Gly(1) with a Gly(NH)···Dpg(2)NH distance of 4.6 Å. (All other $N_iH \leftrightarrow N_{i+1}H$ distances lie between 2.4 and 2.9 Å.) The observation of a d_{NN} NOE of moderate intensity between Gly(1) and Dpg(2) could be indicative of conformational averaging between Type II/II′ and Type III/III′ conformation at the Gly(1)-Dpg(2) segment. Such a process requires a flip of the

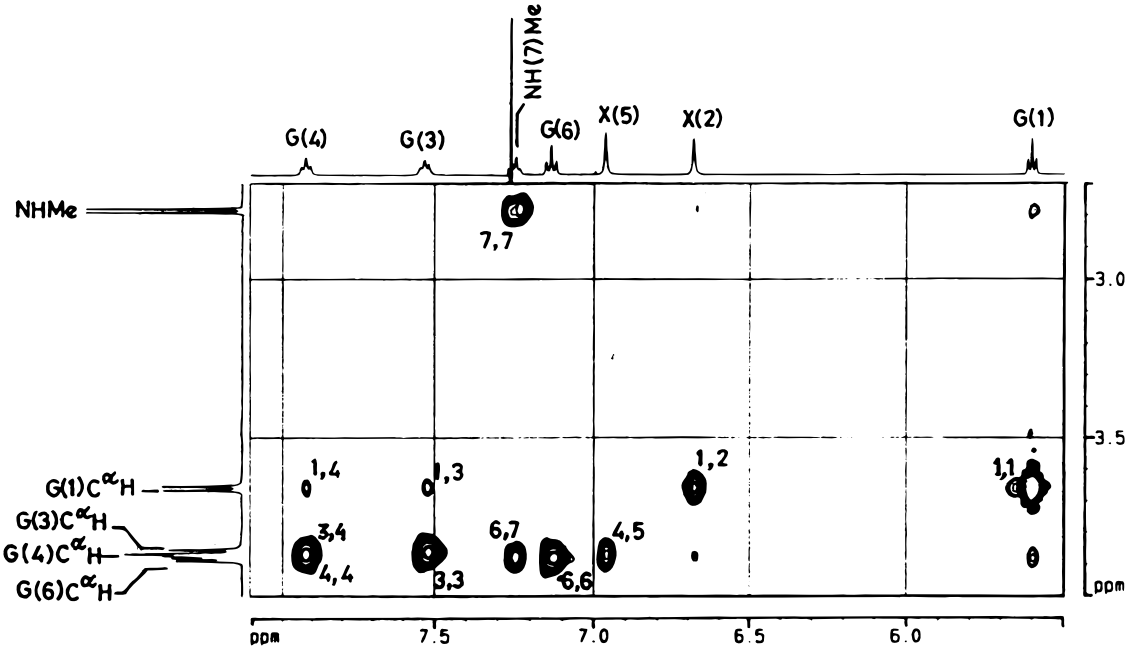

Figure 5. Partial 400 MHz ROESY spectrum of peptide **1** in $CDCl_3$ showing $C^{\alpha}H \leftrightarrow N_{i+1}H$ NOEs. [X = Dpg].

peptide bond between the residues 1 and 2, which can be achieved by correlated rotation about ψ (Gly1) and ϕ (Dpg 2). It should also be noted that if the water invasion of the peptide backbone is ignored, then intramolecular hydrogen bonds of the $4 \rightarrow 1/6 \rightarrow 1$ type will result in solvent shielding of all but the two N-terminal NH groups. Indeed in the crystal the potential type II/II′ β-turn at the Gly(1)-Dpg(2) segment is distorted by hydration, as is the Schellman motif[29−32] at the C-terminus.

The NMR data in $CDCl_3$ solution is thus broadly consistent with a significant population of the multiple turn conformation seen in crystals. Further support for this conclusion is obtained from the medium range NOEs Gly(4)NH↔Gly(6)NH (w) (Figure 4), Gly(1)$C^{\alpha}H$↔Gly(3)NH (w), and Gly(1)$C^{\alpha}H$↔Gly-(4)NH (w) (Figure 5). The corresponding distances in the crystal structure are 4.21, 3.77, and 4.21 Å, respectively. An unusual, strong NOE is also observed between the Boc CH_3 resonance (1.45 δ) and the cluster of $C^{\alpha}H$ resonance of Gly(3), Gly(4), and Gly(6) centred at (3.86−3.88δ) (data not shown). Interestingly, the crystal structure places one of the Boc CH_3 groups relatively close to the Gly(4) $C^{\alpha}H_2$ group. (The shortest Boc CH_3···Gly(4)$C^{\alpha}H$ distance is 2.52 Å.)

The temperature coefficients of NH chemical shifts in $(CD_3)_2$-SO solution are summarized in Table 5. While Gly(1), Dpg-(2), and Gly(3) have high dδ/dt values (>4.5 ppb/K), a very high value of 6.1 ppb/K is also observed for Gly(6)NH. this clearly suggests that in a strongly solvating medium, solvent invasion destabilizes intramolecularly hydrogen bonded backbone conformations. The dδ/dt values observed for Gly(4)-NH, Dpg(5)NH, and NHMe groups lie between 3.7−4.4 ppb/K, which are suggestive of only moderately solvent shielded protons. Figure 6 shows the NH-NH region of the ROESY spectrum of **1** in $(CD_3)_2SO$. Although all sequential d_{NN} connectivities can be identified, the NOE magnitudes appear to be significantly smaller than in $CDCl_3$. This is particularly true for the ROESY cross peaks for the segment residues 2−6. Together with the temperature coefficients, the NOE data suggest a diminished population of folded compact conformations and support unfolding by solvation. These conclusions are reinforced by the observation of very strong exchange cross peaks between residual water protons in $(CD_3)_2SO$ and all the

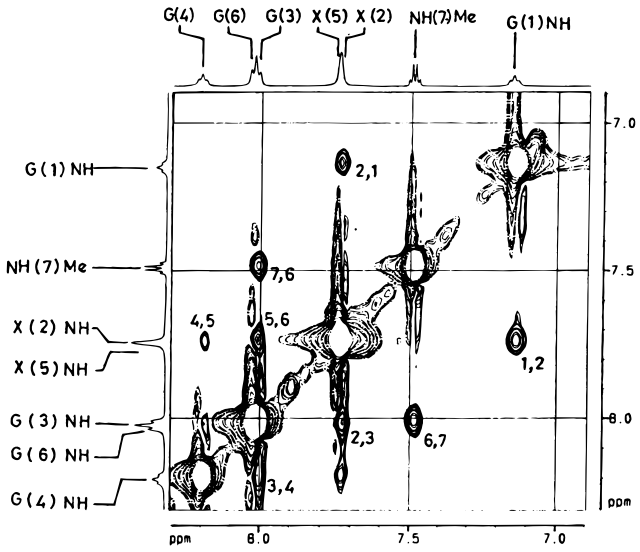

Figure 6. Partial 400 MHz ROESY spectrum of peptide **1** in $(CD_3)_2$-SO showing $N_iH \leftrightarrow N_{i+1}H$ NOEs. [X = Dpg].

backbone NH groups in peptide **1** (Figure 7). The NMR results thus favour a largely multiple turn conformation in $CDCl_3$ which is unfolded and solvated in $(CD_3)_2SO$. In $CDCl_3$ both Dpg residues adopt conformations lying in the helical regions of ϕ, ψ space.

Conclusions

A novel multiple turn conformation is established in single crystals for the peptide Boc-Gly-Dpg-Gly-gly-Dpg-Gly-NHMe. The observed structure provides examples of hydrated TypeII/II′ turns, Type III turns, and a chiral reversal to establish a $6 \rightarrow 1$ hydrogen bonded π-turn, in a hydrated Schellman motif. While five out of six residues lie in right or left handed helical regions of ϕ, ψ space, the molecule as a whole forms a relatively compact folded structure. NMR studies support substantial retention of the crystal state conformation in the apolar solvent $CDCl_3$ with key NOEs and solvent delineation of shielded NH groups favoring the multiple turn structure. A particularly interesting feature of the crystal structure is the adoption of

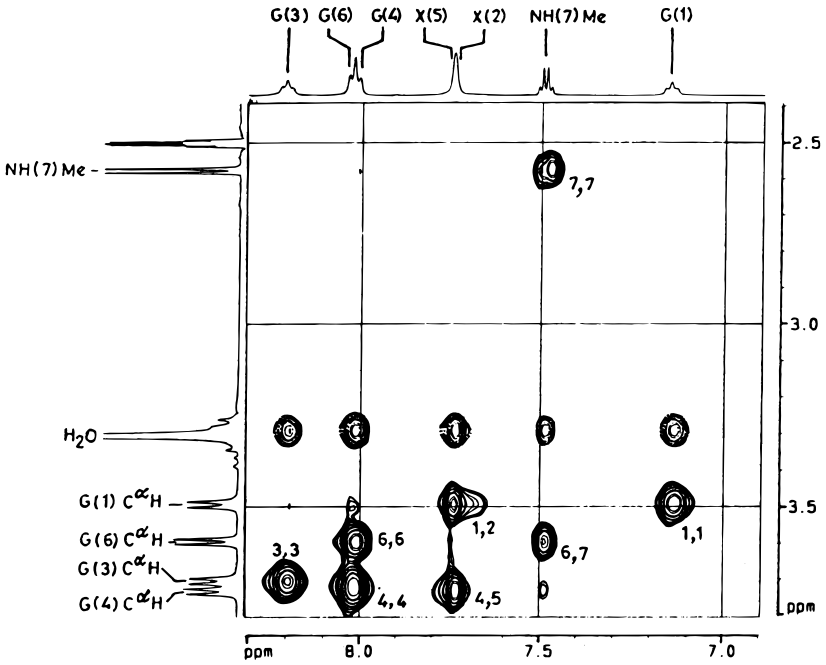

Figure 7. Partial 400 MHz ROESY spectrum of peptide **1** showing $C^{\alpha}H \leftrightarrow N_{i+1}H$ connectivities. Strong exchange cross peaks between residual water protons in $(CD_3)_2SO$ and all the backbone NH groups are indicated. [X = Dpg].

helical ϕ, ψ values for the two Dpg residues even in the context of a very poor helix promoting segments containing as many as four Gly residues, including two central contiguous glycines. Folded conformations at higher dialkylglycine residues must be considered in interpreting structure−activity relationships for biologically active peptides. Indeed, studies of chemotactic tripeptide analogs have shown high biological activity when the central Leu residue is replaced by dialkylglycine containing both cyclic and acyclic side chains.[37,38] In the case of 1-aminocycloalkane-1-carboxylic acids, helical conformations have been exclusively reported, whereas for linear dialkyl residues both helical and extended conformations have been crystallographically characterized.[21] The present results suggest that the *local helical* conformations are likely to be significantly more important even for the higher dialkylglycines. Indeed, the segment Gly-Dpg-Gly has been crystallographically characterized in 3, 6, 10, and 14 residues peptides, with the sole example of an extended conformation being observed for the Dpg residue in the tripeptide Boc-Gly-Dpg-Gly-OH.[39] The conditions under which fully extended conformations can be preferentially stabilized remain to be firmly established. A clear understanding of the factors that determine conformational state of the higher dialkylglycines will be of great value in the control of backbone stereochemistry in peptide design.[13,14]

Acknowledgment. This research was supported in part by National Institute of Health Grant GM-30902, the Office of Naval Research, and the Department of Science and Technology, India. R.K. acknowledges award of a Research Associateship from the Department of Biotechnology, India.

Supporting Information Available: Tables of atomic coordinates, bond lengths, bond angles, anisotropic temperature factors, and H atom coordinates for **1** (7 pages). See any current masthead page for ordering and Internet access information.

JA970596Z

(37) Dentino, A, R.; Raj, P. A.; Bhandary, K. K.; Wilson, M. E.; Levine, M. J. *J. Biol. Chem.* **1991**, *266*, 18460−18468.
(38) Prasad, S.; Rao, R. B.; Bergstrand, H.; Lundquist, B.; Becker, E. L.; Balaram, P. *Int. J. Pept. Protein Res.* **1996**, *48*, 312−318.
(39) Datta, S.; Kaul, R.; Rao, R. B.; Shamala, N.; Balaram, P. *J. Chem. Soc., Perkin Trans. 2* 1997, in press.
(40) Crisma, M.; Valle, G.; Toniolo, C.; Rao, R. B.; Prasad, S.; Balaram, P. *Biopolymers* **1995**, *35*, 1−9.