Peptide Structures

Hydrogen-Bond Lengths in Polypeptide Helices: No Evidence for Short Hydrogen Bonds**

Subrayashastry Aravinda, Saumen Datta, Narayanaswamy Shamala, and Padmanabhan Balaram**

The 3_{10} helix and α helix are closely related secondary structures observed in polypeptides. The 3_{10} helix is characterized by successive 4→1 (C_{10}) hydrogen bonds (C_{10} = 10 atom hydrogen-bonding ring) of the type $C{=}O_i\cdots HN_{i+3}$, while the α helix displays a hydrogen bond of the type $C{=}O_i\cdots HN_{i+4}$ (C_{13}). The α helix is widely distributed in proteins, while the occurrence of segments of 3_{10} helix is very much less frequent.[1] In polypeptides containing $C^{\alpha,\alpha}$ dialkylated residues, α-aminoisobutyric acid (Aib) being the prototype, both the 3_{10} and α helical structures are detected.[2] In homooligomers of Aib, 3_{10} helices are invariably found in crystals,[3] while in heteromeric sequences the precise helical type appears to depend on both Aib content and positioning.[2a,d,4,5] While distinctions between 3_{10} and α helices are possible in the crystalline state, such differentiation becomes difficult in solution.[6] Circular dichroism (CD) has been proposed for distinguishing between 3_{10} and α helix structures by using the ratio of CD bands at 222 nm and 207/208 nm.[7] However, the use of the $[\theta]_{222}/[\theta]_{208}$ ratio has been questioned, suggesting that the distinction between 3_{10} and α helix structures by chiroptical methods may not be readily possible.[8] The conventional interpretation of the CD spectra of helical polypeptides has been further called into question by the careful work of Kemp and co-workers, who have reported the observation of large values of $[\theta]_{222}$, which are inconsistent with those currently accepted for 100% helical structures.[9] Kemp et al. have noted that the 222 nm nπ* band has not "been modeled satisfactorily by theory". The widespread use of the CD band intensities at 208 and 222 nm, in estimating helicity values quantitatively and in making qualitative distinctions between helix subtypes, underscores the importance of relating CD spectral intensities to specific peptide structural features. In addressing this issue, Dang and Hirst[10] have used an improved theoretical method to calculate the 220 nm CD band intensities and have suggested that $[\theta]_{220}$ is extremely sensitive to main-chain hydrogen-bond length. They argue that "shortening from a conventional oxygen–nitrogen separation of about 3.0 Å to 2.8 Å or 2.7 Å is predicted to lead to a sizable enhancement of the intensity at 220 nm", with the effect being most pronounced for α helices and less dramatic for 3_{10} and π helices. These calculations also reveal a dependence of $[\theta]_{220nm}$ on N···O separations in the range 3.0–3.5 Å, a factor which may contribute to the variations in band intensities in model 3_{10} and α helical peptides. With the exception of the Dang and Hirst proposal no testable explanations have been advanced for the observed variation in the 220 nm CD band intensity.

Herein we examine the distribution of hydrogen-bond parameters in high-resolution crystal structures of Aib-containing peptides to assess whether *short* hydrogen bonds are indeed observed in experimental structures. Aib-containing peptides were chosen because of the availability of a large number of accurately determined structures and the occurrence of pure α, 3_{10}, and mixed 3_{10}/α helical structures in crystals.

The structure of *p*BrBz-$(Aib)_{10}$-O*t*Bu (peptide **2**; *p*BrBz = *para*-bromobenzoyl) provides an example of a complete 3_{10} helix.[3e] The structure of the peptide Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-Ala-Leu-Aib-OMe (peptide **1**; Boc = 1,1-dimethylethoxycarbonyl (= *tert*-butyloxycarbonyl)) determined herein provides an example of a complete α helix over the length of the decapeptide. Figure 1 illustrates the molecular conformations of the two peptides and the distribution of backbone dihedral angles. The parameters for potential hydrogen-bond interactions in the two structures are compared in Table 1. The backbone dihedral angles are given in Table 2. Peptide **1** (α helix) is stabilized by seven intramolecular 5→1 hydrogen bonds while Peptide **2** (3_{10} helix) is stabilized by eight intramolecular 4→1 hydrogen bonds. The data in Table 1 suggests that for the α helix the observed range of 1→5 N···O separation is 2.978 Å to 3.113 Å, while for 1→4 separations in the 3_{10} helix it is 2.907 Å to 3.211 Å. Clearly, none of the observed hydrogen bonds would correspond to the *short* hydrogen bond (2.7 to 2.8 Å) used by Dang and Hirst in their theoretical computation of CD band intensities.[10]

We have classified the hydrogen-bond patterns in the available crystal structures of Aib-containing helical peptides into 4→1 and 5→1 hydrogen bonds by using the following criteria: N···O separation $\leq$ 3.5 Å, H···O separation $\leq$ 2.6 Å, $\measuredangle$N–H···O $\geq$ 90°, $\measuredangle$C=O···H $\geq$ 90°, and $\measuredangle$C=O···N $\geq$ 90°.[11] Table 3 summarizes the limits and mean of the hydrogen-bond parameters for the two hydrogen-bond types. Figure 2 summarizes the distribution of the parameters N···O separation and the angles C=O···N and N–H···O. In 4→1 and 5→1 hydrogen bonds the mean hydrogen-bond lengths are 3.034 ± 0.109 Å and 3.034 ± 0.118 Å, respectively. This analysis suggests that short hydrogen bonds N···O < 2.8 Å are extremely infrequent in polypeptide structures. While our analysis does not exclude the possibility of solvent-induced shortening of

[*] Prof. P. Balaram
Molecular Biophysics Unit
Indian Institute of Science
Bangalore—560 012 (India)
Fax: (+91) 80-2360-0683 / 0535
E-mail: pb@mbu.iisc.ernet.in
S. Aravinda, Dr. S. Datta, Prof. N. Shamala
Department of Physics
Indian Institute of Science
Bangalore—560 012 (India)
Fax: (+91) 80-2360-2602 / 0683
E-mail: shamala@physics.iisc.ernet.in

[**] This research was supported by a grant from the Council of Scientific and Industrial Research and Program Support in the area of Molecular Diversity and Design, Department of Biotechnology, India. We thank the reviewers for extremely constructive comments.

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

  DOI: 10.1002/ange.200461127

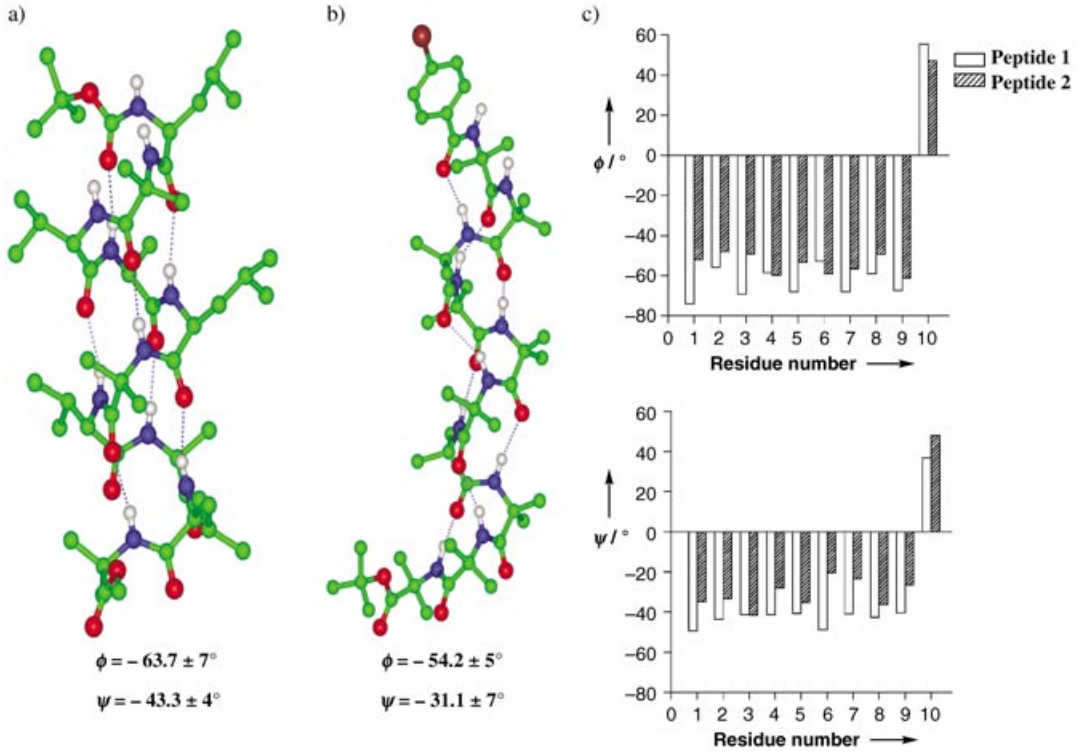

Figure 1. a) Molecular conformation of Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-Ala-Leu-Aib-OMe (peptide **1**) in the crystal structure. b) Molecular conformation of pBrBz-$(Aib)_{10}$-O*t*Bu (peptide **2**) in the crystal structure. All the intramolecular hydrogen bonds are shown as dotted lines; green C, red O, blue N. c) The distribution of the torsion angles in the peptide backbone ϕ,ψ observed in peptides **1** and **2**.

Table 1: Hydrogen bonds in peptides **1** and **2**.

Type	Donor	Acceptor	N···O [Å]	H···O [Å]	C=O···H [°]	C=O···N [°]	N–H···O [°]	N···O [Å]	H···O [Å]	C=O···H [°]	C=O···N [°]	N–H···O [°]
Intramolecular			Ideal α helix (peptide **1**)					Ideal 3_{10} helix (peptide **2**)				
4→1	N(3)	O(0)	3.559	3.104	96.7	108.5	115.4	2.945[a]	2.123	127.5	131.7	159.8
4→1	N(4)	O(1)	3.200	2.779	94.9	108.4	111.8	2.975[a]	2.182	131.8	136.4	153.4
4→1	N(5)	O(2)	3.401	2.920	91.7	103.5	117.3	2.979[a]	2.198	118.9	126.0	151.1
4→1	N(6)	O(3)	3.221	2.855	92.9	106.9	107.5	3.179[a]	2.382	119.9	125.3	154.4
4→1	N(7)	O(4)	3.393	2.923	90.3	102.4	116.3	3.020[a]	2.183	122.1	126.5	164.2
4→1	N(8)	O(5)	3.283	2.884	92.3	105.7	110.3	3.211[a]	2.355	129.3	130.9	173.4
4→1	N(9)	O(6)	3.390	2.932	93.4	105.6	115.3	3.057[a]	2.225	130.7	133.3	163.0
4→1	N(10)	O(7)	3.235	2.782	90.9	104.5	114.5	2.907[a]	2.087	122.1	128.0	159.2
5→1	N(4)	O(0)	2.978[a]	2.165	149.9	156.1	157.7	4.613	3.916	136.8	143.1	140.7
5→1	N(5)	O(1)	3.113[a]	2.277	157.0	159.8	163.9	4.337	3.665	135.5	143.2	137.3
5→1	N(6)	O(2)	3.071[a]	2.237	147.0	151.4	163.3	4.331	3.630	142.5	149.2	140.8
5→1	N(7)	O(3)	3.005[a]	2.174	155.2	159.1	162.1	4.600	4.012	129.9	138.3	128.8
5→1	N(8)	O(4)	3.014[a]	2.183	147.6	152.2	162.3	4.791	4.181	135.7	142.9	131.2
5→1	N(9)	O(5)	3.048[a]	2.209	154.2	157.5	165.2	5.200	4.547	125.1	131.5	136.0
5→1	N(10)	O(6)	2.986[a]	2.194	146.7	154.0	153.0	4.600	3.970	128.7	136.4	133.0

[a] Acceptable hydrogen bonds.

hydrogen bonds as a factor influencing observed CD spectra, there is at present no evidence for such a phenomenon in solution. It may, therefore, be useful to reexamine the factors that contribute to the intensities of the n→π* band in CD spectra of polypeptides. The extensive application of CD methods in estimating helicity in peptides and proteins emphasizes the importance of a rigorous theoretical underpinning for the analysis of peptide CD spectra. In particular n→π* band intensities may be influenced by a number of factors. The results herein suggest that short hydrogen bonds may not always be the sole contributing factor. Chiroptical distinctions between 3_{10} and α helical structures must also be tempered by the fact that many sequences can exhibits mixed hydrogen-bonding patterns and can exist in solution as equilibrium mixtures of closely related conformations.

Table 2: Torsion angles [°][a] of peptides **1** and **2**.

	Peptide **1**				Peptide **2**		
Residue	ϕ	ψ	ω	Residue	ϕ	ψ	ω
Leu(1)	−74.1[b]	−49.5	−178.0	Aib(1)	−51.9	−34.9	−178.2
Aib(2)	−55.9	−43.7	−175.7	Aib(2)	−48.2	−33.5	−177.3
Val(3)	−69.3	−41.3	176.6	Aib(3)	−49.0	−41.7	−173.2
Ala(4)	−58.5	−41.4	179.2	Aib(4)	−59.8	−28.1	−177.3
Leu(5)	−68.1	−40.7	175.4	Aib(5)	−53.2	−35.4	−173.7
Aib(6)	−52.7	−49.0	−174.9	Aib(6)	−59.0	−20.6	−178.0
Val(7)	−68.0	−40.8	177.6	Aib(7)	−56.5	−23.3	177.6
Ala(8)	−59.1	−42.7	179.1	Aib(8)	−49.3	−36.3	−177.6
Leu(9)	−67.2	−40.5	−176.9	Aib(9)	−61.2	−26.5	−172.1
Aib(10)	55.4	36.8[c]	179.1[d]	Aib(10)	47.3	48.0[f]	176.4
Mean(9)[e]	−63.7 ± 7.2	−43.3 ± 4			−54.2 ± 5.0	−31.1 ± 6.9	

[a] The torsion angles for rotation about bonds of the peptide backbone (ϕ, ψ, and ω) and about bonds of the amino acid side chains (χ^1, χ^2) as suggested by the IUPAC-IUB Commission on Biochemical Nomenclature.[15] Estimated standard deviation ≈0.5°. [b] C′(0)-N(1)-C$^\alpha$(1)-C′(1). [c] N(10)-C$^\alpha$(10)-C′(10)-O(OMe). [d] C$^\alpha$(10)-C′(10)-O(OMe)-C(OMe). [e] Mean value of the first nine residues. [f] The ψ value of −136.4° in ref [3e] corresponds to the torsion angle N(10)-C$^\alpha$(10)-C′(10)-O(10).

Table 3: Summary of 4→1 and 5→1 hydrogen bonds in peptide helices.

	4→1 hydrogen bonds				5→1 hydrogen bonds			
	min	max	mean	Std	min	max	mean	Std
N···O [Å]	2.761	3.480	3.034	0.109	2.795	3.460	3.034	0.118
H···O [Å]	1.915	2.600	2.267	0.142	1.951	2.600	2.215	0.130
C=O···H [°]	90.7	156.0	119.7	10.5	127.5	174.3	151.0	7.6
C=O···N [°]	102.5	164.7	125.8	8.5	136.8	175.3	155.9	6.4
N−H···O [°]	114.1	173.4	150.4	14.3	121.3	178.1	158.2	9.2

Experimental Section

Peptide **1** was synthesized by conventional solution-phase procedures using a segment condensation strategy. Boc groups are used as N-terminal protecting groups and methyl ester as the C-terminal protecting groups. Peptide couplings were mediated by *N*,*N*′- dicyclohexylcarbodiimide (DCC) and 1-hydroxybenzotriazole.[12] The final step involved a 3+7 coupling. Purification of the peptide was by reverse-phase medium-pressure liquid chromatography (C_{18}, 40–60μ), using methanol–water gradients. The peptide was characterized by 500 MHz ^{1}H NMR spectroscopy and MALDI mass spectrometry ($[M+Na^+]_{obs}$ = 1089.7; M_{calcd} = 1067.37) Crystals of were grown by slow evaporation of a methanol–water mixture of peptide **1**. X-ray data were collected at room temperature from a dry crystal, on a CAD-4 diffractometer, using $Cu_{K\alpha}$ radiation (λ = 1.5418 Å). ω-2θ scan type was used, with 2θ = 150°, for a total of 7381 independent reflections. The crystal size was 0.2 × 1.0 × 0.2 mm. Space group $P2_12_12_1$, a = 10.791(3), b = 16.632(2), c = 36.068(9) Å, V = 6473(3) Å^3, Z = 4 for chemical formula $C_{52}H_{94}N_{10}O_{13}$, with one molecule per asymmetric unit. ρ_{calcd} = 1.095 g cm^{-3}, μ = 0.064 mm^{-1}, $F(000)$ = 2320. The structure was obtained by direct methods using SHELXS-86.[13a] Refinement was carried out against F^2, with full-matrix least-squares methods using SHELXL-93.[13b] All non-hydrogen atoms were refined isotropically. The R value at the end of isotropic refinement was 0.10. The R value dropped to 0.06 after anisotropic refinement. The hydrogen atoms were fixed geometrically in idealized positions and refined in the final cycle of refinement as riding over the atoms to which they are bonded. The final R value was 0.0418 (wR_2 =

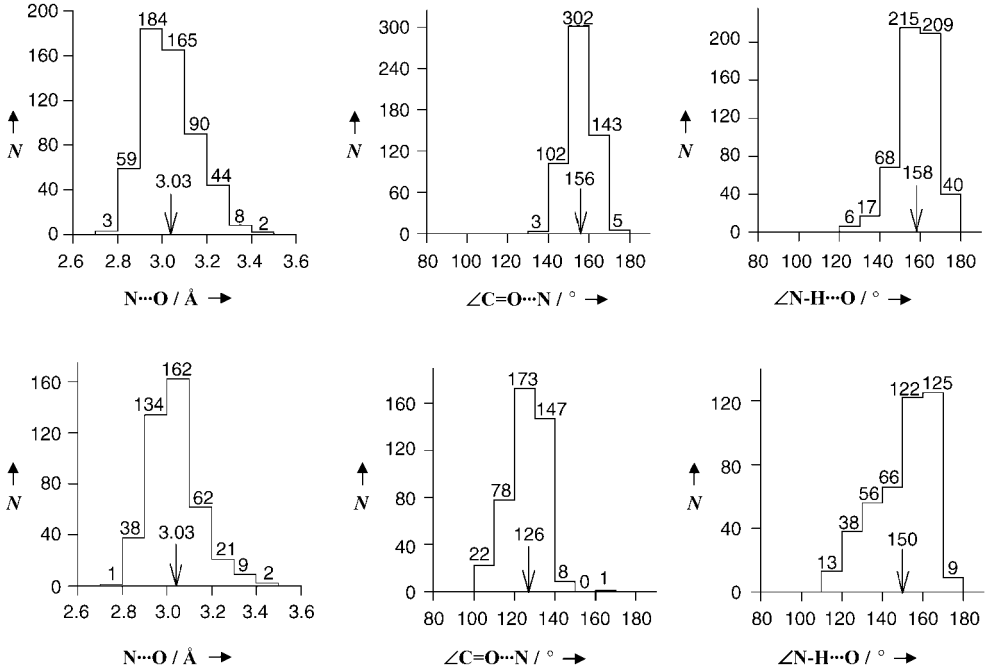

Figure 2. Distribution of N···O separation, C=O···N angles, and N-H···O angles observed in 5→1 hydrogen bonds (top) and 4→1 hydrogen bonds (bottom). *N* is the number of occurrences in the dataset.

0.113) for observed reflections 5170 with $F_o \geq 4\sigma(|F_o|)$ and 673 variables, where the data to parameter ratio is 7.7:1.0 and $S = 1.038$. The largest difference peak was 0.17 e Å^{-3} and the largest difference hole was -0.25 e Å^{-3}. The standard deviations in bond lengths are approximately 0.004 Å and those of bond angles are approximately 0.3°. CCDC 243105 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/conts/retrieving.html (or from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: (+44) 1223-336-033; or deposit@ccdc.cam.ac.uk).

Data base analysis: Out of 407 sequences containing at least one Aib residue retrieved from the Cambridge Structural Database (CSD) (November 2003 release),[14] 393 were linear and 14 were cyclic peptides. For the analysis we have used peptide sequences of length six amino acids and more. 128 sequences were used for hydrogen-bond calculation. All the N–H separations were fixed at 0.86 Å. All the 4→1 and 5→1 hydrogen bonds were calculated. There are totally 555 (α helical) 5→1 hydrogen bonds and 429 (3_{10} helical) 4→1 hydrogen bonds in 128 sequences.

For a list of peptides, CCDC codes and references, used in the hydrogen-bond analysis see the Supporting Information.

Received: June 30, 2004
Revised: August 10, 2004

Keywords: circular dichroism · helical structures · hydrogen bonds · peptides · structure elucidation

[1] a) D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1988**, *201*, 601 – 619; b) K. A. Bolin, G. L. Millhauser, *Acc. Chem. Res.* **1999**, *32*, 1027 – 1033.

[2] a) I. L. Karle, P. Balaram, *Biochemistry* **1990**, *29*, 6747 – 6756; b) C. Toniolo, E. Benedetti, *Trends Biochem. Sci.* **1991**, *16*, 350 – 353; c) C. Toniolo, E. Benedetti, *Macromolecules* **1991**, *24*, 4004 – 4009; d) I. L. Karle, *Acta Crystallogr. Sect. B* **1992**, *48*, 341 – 356; e) E. Benedetti, *Biopolymers (Peptide Science)* **1996**, *40*, 3 – 44; f) I. L. Karle, *Biopolymers (Peptide Science)* **1996**, *40*, 157 – 180; g) R. Kaul, P. Balaram, *Bioorg. Med. Chem.* **1999**, *7*, 105 – 117.

[3] For a representative set of examples, see: a) N. Shamala, R. Nagaraj, P. Balaram, *J. Chem. Soc. Chem. Commun.* **1978**, 996 – 997; b) E. Benedetti, A. Bavoso, B. Di Blasio, V. Pavone, C. Pedone, M. Crisma, G. M. Bonora, C. Toniolo, *J. Am. Chem. Soc*, **1982**, *104*, 2437 – 2444; c) A. Bavoso, E. Benedetti, B. Di Blasio, V. Pavone, C. Pedone, C. Toniolo, G. M. Bonora, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 1988 – 1992; d) R. Gessmann, H. Brückner, M. Kokkinidis, *Acta Crystallogr. Sect. B* **1998**, *54*, 300 – 307; e) C. Toniolo, M. Crisma, G. M. Bonora, E. Benedetti, B. Di Blasio, V. Pavone, C. Pedone, A. Santini, *Biopolymers* **1991**, *31*, 129 – 138; f) R. Gessman, H. Brückner, K. Petratos, *J. Pept. Sci.* **2003**, *9*, 753 – 762.

[4] a) I. L. Karle, J. L. Flippen-Anderson, R. Gurunath, P. Balaram, *Protein Sci.* **1994**, *3*, 1547 – 1555; b) V. Pavone, E. Benedetti, B. Di Blasio, C. Pedone, A. Santini, A. Bavoso, C. Toniolo, M. Crisma, L. Sartore, *J. Biomol. Struct. Dyn.* **1990**, *7*, 1321 – 1331; c) G. R. Marshall, E. E. Hodgkin, D. A. Langs, D. G. Smith, J. Zabrocki, M. T. Leplawy, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 487 – 491.

[5] a) G. Basu, K. Bagchi, A. Kuki, *Biopolymers* **1991**, *31*, 1763 – 1774; b) G. Basu, A. Kuki, *Biopolymers* **1992**, *32*, 61 – 71; c) K. Otoda, Y. Kitagawa, S. Kimura, Y. Imanishi, *Biopolymers* **1993**, *33*, 1337 – 1345.

[6] a) G. L. Millhauser, *Biochemistry* **1995**, *34*, 3873 – 3877; b) S. Mammi, M. Rainaldi, M. Bellanda, E. Schievano, E. Peggion, Q. B. Broxterman, F. Formaggio, M. Crisma, C. Toniolo, *J. Am. Chem. Soc.* **2000**, *122*, 11735 – 11736; c) R. A. G. D. Silva, S. C. Yasui, J. Kubelka, F. Formaggio, M. Crisma, C. Toniolo, T. A. Keiderling, *Biopolymers* **2002**, *65*, 229 – 243.

[7] a) C. Toniolo, A. Polese, F. Formaggio, J. Kamphuis, *J. Am. Chem. Soc.* **1996**, *118*, 2744 – 2745; b) G. Yoder, A. Polese, R. A. G. D. Silva, F. Formaggio, M. Crisma, Q. B. Broxterman, J. Kamphuis, C. Toniolo, T. A. Keiderling, *J. Am. Chem. Soc.* **1997**, *119*, 10278 – 10285.

[8] a) T. S. Sudha, E. K. S. Vijayakumar, P. Balaram, *Int. J. Pept. Protein Res.* **1983**, *22*, 464 – 468; b) M. C. Manning, R. W. Woody, *Biopolymers* **1991**, *31*, 569 – 586; c) N. H. Andersen, Z. Liu, K. S. Prickett, *FEBS. Lett.* **1996**, *399*, 47 – 52.

[9] a) P. Wallimann, R. J. Kennedy, D. S. Kemp, *Angew. Chem.* **1999**, *111*, 1377 – 1379; *Angew. Chem. Int. Ed.* **1999**, *38*, 1290 – 1292; b) P. Wallimann, R. J. Kennedy, J. S. Miller, W. Shalongo, D. S. Kemp, *J. Am. Chem. Soc.* **2003**, *125*, 1203 – 1220.

[10] Z. Dang, J. D. Hirst, *Angew. Chem.* **2001**, *113*, 3731 – 3733; *Angew. Chem. Int. Ed.* **2001**, *40*, 3619 – 3621.

[11] a) E. N. Baker, R. E. Hubbard, *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97 – 179; b) S. Datta, N. Shamala, A. Banerjee, P. Balaram, *J. Pept. Res.* **1997**, *49*, 604 – 611.

[12] H. Balaram, M. Sukumar, P. Balaram, *Biopolymers* **1986**, *25*, 2209 – 2223.

[13] a) G. M. Sheldrick *SHELXS-86*, Program for the Crystal Structure Solution, Universität Göttingen, Germany, **1986**; b) G. M. Sheldrick, *SHELXL-93*, Program for the Refinement of Crystal Structures, Universität Göttingen, Germany, **1993**.

[14] F. H. Allen, *Acta Crystallogr. Sect. B* **2002**, *58*, 380 – 388.

[15] IUPAC–IUB Commission on Biochemical Nomenclature *Biochemistry* **1970**, *9*, 3471 – 3479

