Molecular Structure of Boc-Aib-Aib-Phe-Met-NH_2·DMSO. A Fragment of a Biologically Active Enkephalin Analogue

B. V. Venkataram Prasad, T. Subrahmanyam Sudha, and Padmanabhan Balaram *
Molecular Biophysics Unit and Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560 012, India

The tetrapeptide t-butyloxycarbonyl-α-aminoisobutyryl-α-aminoisobutyryl-L-phenylalanyl-L-methionyl amide crystallizes in the orthorhombic space group $P2_12_12_1$ with a = 9.096, b = 18.067, c = 21.701 Å and Z = 4. The crystals contain one molecule of dimethyl sulphoxide (DMSO) associated with each peptide. The structure has been solved by direct methods and refined to an R value of 0.103 for 2 672 observed reflections. The peptide adopts a distorted 3_{10} helical structure stabilized by two intramolecular 4 ⟶ 1 hydrogen bonds between the Boc CO and Aib(1) CO groups and the NH groups of Phe(3) and Met(4), respectively. A long hydrogen bond (N · · · O = 3.35 Å) is also observed between Aib(2) CO and one of the terminal amide hydrogens. The DMSO molecule is strongly hydrogen bonded to the Aib(1) NH group. The solid-state conformation agrees well with proposals made on the basis of n.m.r. studies in solution.

There has been tremendous effort in attempting to establish structure-activity relationships for enkephalins and synthetic analogues.[1-3] The necessity of considering both backbone conformation and side-chain orientation in acyclic pentapeptide sequences contributes significantly to the complexity of the problem. The presence of the Gly-Gly segment in the natural enkephalins [4,5] (Tyr-Gly-Gly-Phe-Met/Leu) results in considerable conformational flexibility of the peptide backbone. Attempts to elucidate conformation-function correlations have been limited by the difficulties in unambiguously demonstrating the occurrence of specific conformations for enkephalins, in solution.[6-15] Several attempts have therefore been made to synthesize conformationally constrained enkephalin analogues, which retain biological activity.[16-20] It has been established that replacement of Gly residues by α-aminoisobutyryl (Aib) residues, dramatically restricts backbone flexibility and abolishes conformational transitions detected by c.d. methods.[21] The pentapeptides Tyr-Aib-Gly-Phe-Met-NH_2 and Tyr-Aib-Aib-Phe-Met-NH_2 induce long lasting, 'enkephalin like' behavioural effects in mice following intracerebral administration.[20] ^{1}H N.m.r. studies establish that in the latter both Phe(4) and Met(5) NH groups are hydrogen bonded, suggesting a consecutive β-turn or 3_{10} helical conformation. Here we describe the crystal structure of the related peptide, Boc-Aib-Aib-Phe-Met-NH_2. The peptide folds into an incipient 3_{10} helical structure.

Experimental

The tetrapeptide was synthesized by conventional solution-phase procedures. Single crystals were obtained from chloroform–dimethyl sulphoxide mixtures in the space group $P2_12_12_1$ with a = 9.096 (7), b = 18.067 (5), c = 21.701 (6) Å and Z = 4. One molecule of dimethyl sulphoxide (DMSO) was found associated with each peptide molecule, after structure determination. The X-ray intensity data were collected on a CAD-4 diffractometer, using ω − 2θ scan up to a Bragg angle of 23° with Mo-K_α radiation. Of the 3 220 reflections collected, 2 672 reflections having $I > 2\sigma(I)$ were used in the refinement. The intensities were corrected for Lorentz and polarization factors but not for absorption.

The structure was determined using the direct methods program MULTAN,[22] and refined using standard procedures. Hydrogen atoms were fixed stereochemically [23] and refinement carried out by a block diagonal least-squares method with anisotropic and isotropic temperature factors for non-hydrogen and hydrogen atoms, respectively. Refinement converged to a final R value of 0.103. The scattering factors for the non-hydrogen and hydrogen atoms are from refs. 24 and 25, respectively. The final positional parameters and the equivalent isotropic temperature factors [26] of the non-hydrogen atoms are given in Table 1. The anisotropic thermal parameters of the non-hydrogen atoms and the observed and calculated structure factors are given in Supplementary Publication No. 23480 (22 pages).

Results and Discussion

Molecular Conformation.—The bond lengths and valence angles in the tetrapeptide are summarized in Figures 1 and 2. Perspective views of the molecule are shown in Figure 3. Backbone and side-chain torsional angles [27] are listed in Table 2, while inter- and intra-molecular hydrogen bond parameters are given in Table 3.

The peptide backbone folds into an incipient 3_{10} helical conformation,[28,29] stabilized by two good intramolecular hydrogen bonds (4⟶1) between Boc CO · · · HN Phe(3) and Aib(1) CO · · · HN Met(4). The φ,ψ values for Aib(1) and Aib(2) are very close to that expected for an ideal right-handed 3_{10} helix (φ = −60°, ψ = −30°).[30] For Phe(3) and Met(4) the φ,ψ values deviate significantly from the 3_{10} helical values. While an ideal 3_{10} helix is generated by a repetitive Type III β-turn structure,[31] the tetrapeptide conformation may be best described as a consecutive β-turn structure of the Type III–Type I category. There is also the possibility of a third β-turn in the backbone, with Phe(3)–Met(4) as the corner residues. A weak hydrogen bonding interaction appears feasible between one of the terminal amide NH hydrogens [H(6-1)] and the Aib(2) CO group. The observed N · · · O distance of 3.35 Å is rather long for a hydrogen bond, but the N–H and C=O groups are appropriately aligned. Long hydrogen bonds have also been noted in other peptide structures.[32,33]

The Phe side-chain adopts a conformation similar to that observed in the crystal structure of Leu-enkephalin.[34] The Met side-chain adopts a *gauche-trans-gauche* orientation, similar to that observed in the structure of D,L-Met (α-form) [35] and the C-terminal residue in L-Met-L-Met.[36]

* For details of the Supplementary publications scheme, see Notice to Authors No. 7, *J. Chem. Soc., Perkin Trans. 1*, 1981, Index issue.

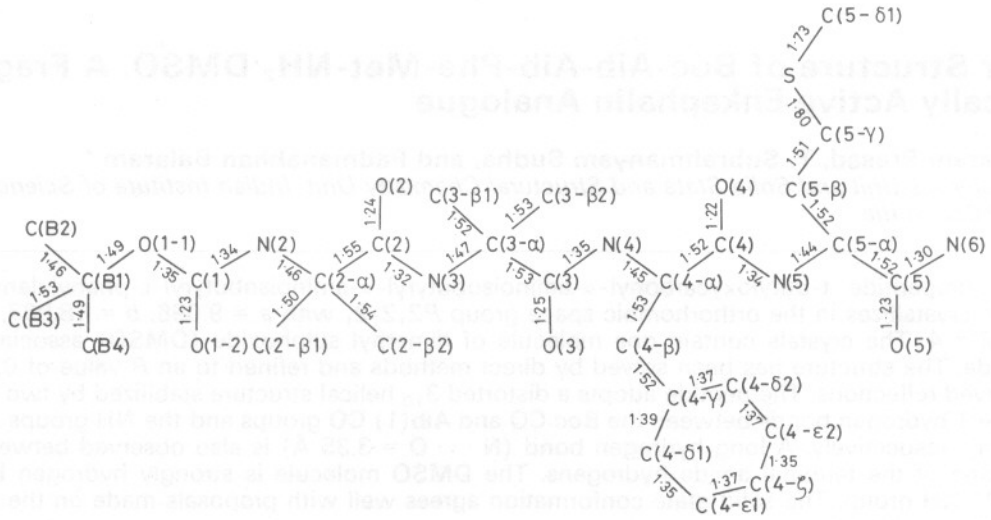


Figure 1. Bond lengths in Boc-Aib-Aib-Phe-Met-NH_2

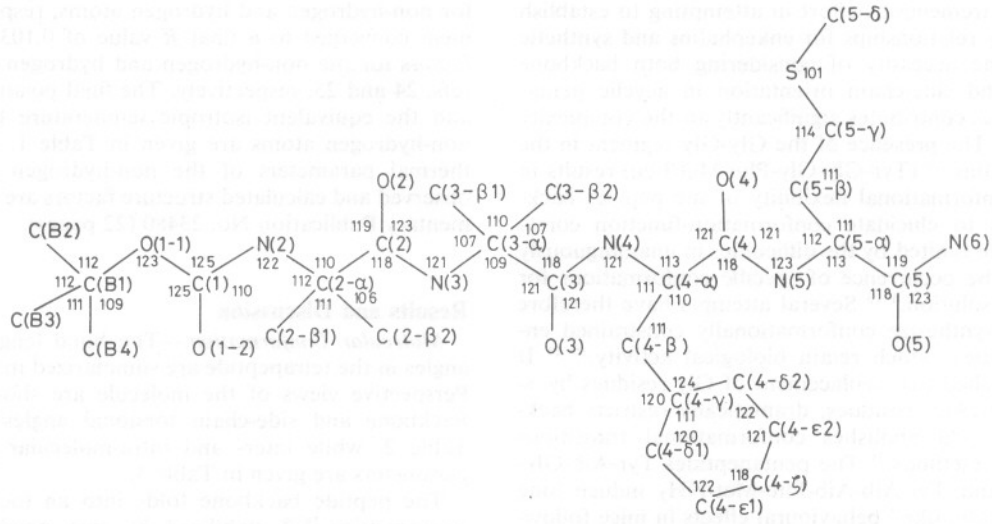


Figure 2. Bond angles in Boc-Aib-Aib-Phe-Met-NH_2

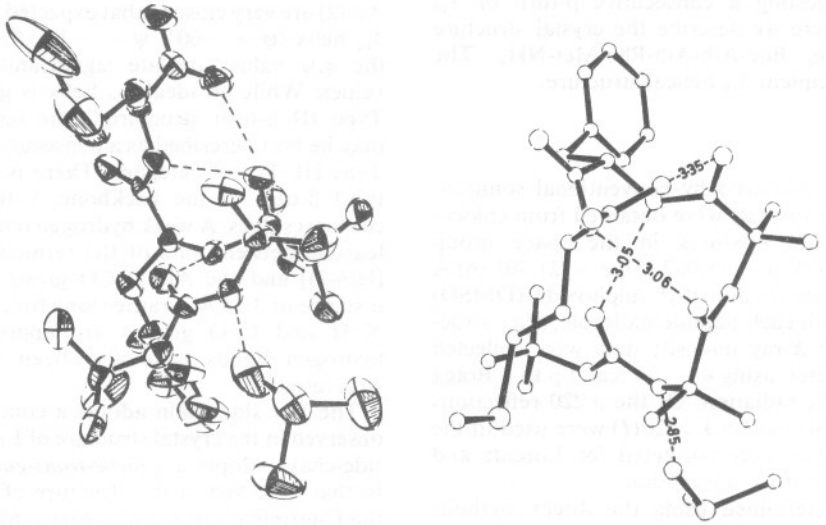


Figure 3. Perspective views of the molecular structure of Boc-Aib-Aib-Phe-Met-NH_2. (Left) View down the *c* axis. Non-hydrogen atoms are represented by thermal ellipsoids defined by the principal axes of thermal vibration and scaled to include 50% probability. Note the large thermal motion of the C(5-δ) atom. (Right) A view down the *b* axis, which is approximately parallel to the helix axis. Intramolecular hydrogen bonds are indicated by dashed lines

Dimethyl Sulphoxide Structure.—DMSO has been found only infrequently as a solvent of crystallization in organic structures. Table 4 compares the molecular parameters observed for DMSO in the present study, with others reported earlier.[34–40] In the structure of the 15 residue cyclic peptide, the DMSO molecules were poorly determined due to disorder in the crystal.[37] In all the other structures the O–S–C angles are close to tetrahedral values, whereas the C–S–C angle is substantially smaller. This is presumably due to repulsions between the sulphoxide oxygen and the lone pair of electrons on sulphur, resulting in a compression of the C–S–C angle.

Molecular Packing.—Figure 4 illustrates the packing of molecules observed in the crystals of Boc-Aib-Aib-Phe-Met-NH_2. Two intermolecular hydrogen bonds link each peptide to its neighbours. The DMSO molecule is strongly hydrogen bonded to the Aib(1) NH group.

Comparison with Solution Conformation.—The molecular conformation of Boc-Aib-Aib-Phe-Met-NH_2, determined in the solid state, agrees very well with proposals based on ^{1}H n.m.r. studies in $(CD_3)_2SO$. The Phe and Met NH groups show very low temperature dependence of chemical shifts, suggesting that they are involved in intramolecular hydrogen bonds in solution. The presence of two Aib residues appears to impart sufficient conformational rigidity to the peptide backbone so that a specific folded conformation predominates, in solution.[21a] The n.m.r. data does not provide clear evidence for involvement of the C-terminal amide group in intramolecular hydrogen bonding. This is because temperature dependence studies in $(CD_3)_2SO$ are complicated by exchange processes, due to rotation about the C–N bond, leading to line broadening and coalescence of the primary amide NH resonances at 343 K.

A comparison of the n.m.r. parameters of the NH groups in the tetrapeptide and the active enkephalin analogue, Tyr-Aib-Aib-Phe-Met-NH_2, suggests that the latter adopts a similar folded conformation.[41] It appears that folded, helical backbone conformations in enkephalins permit proper orientation of the Tyr, Phe and Met/Leu side-chains at the appropriate receptor sites.

Table 1. Final positional co-ordinates and equivalent isotropic temperature factors of non-hydrogen atoms. ESDs are given in parentheses

Atom	*x*	*y*	*z*	*B*
C(B1)	0.839 3(7)	−0.428 2(4)	0.337 2(4)	4.7
C(B2)	0.829 7(10)	−0.405 5(5)	0.272 9(3)	8.2
C(B3)	0.857 9(7)	−0.361 7(4)	0.380 2(3)	5.6
C(B4)	0.953 9(8)	−0.485 2(5)	0.349 3(4)	7.1
O(1—1)	0.695 3(5)	−0.456 4(2)	0.360 5(2)	4.3
C(1)	0.635 4(7)	−0.520 5(2)	0.341 0(3)	3.6
C(1—2)	0.682 4(4)	−0.557 8(2)	0.297 9(2)	4.3
N(2)	0.515 9(5)	−0.536 5(3)	0.374 7(2)	3.5
C(2-α)	0.447 6(6)	−0.609 4(3)	0.373 2(3)	3.5
C(2-β1)	0.542 6(9)	−0.665 9(4)	0.404 4(3)	5.9
C(2-β2)	0.298 1(8)	−0.602 0(4)	0.406 3(3)	5.9
C(2)	0.410 8(6)	−0.603 8(3)	0.305 6(3)	3.5
O(2)	0.438 5(4)	−0.694 6(2)	0.287 3(2)	4.2
N(3)	0.346 0(5)	−0.581 0(2)	0.270 6(2)	3.5
C(3-α)	0.284 5(6)	−0.600 5(3)	0.210 0(3)	3.7
C(3-β1)	0.231 2(8)	−0.529 1(4)	0.180 1(3)	5.4
C(3-β2)	0.157 1(7)	−0.655 0(4)	0.216 4(3)	5.2
C(3)	0.406 8(7)	−0.630 8(3)	0.169 1(3)	3.4
O(3)	0.381 3(5)	−0.679 3(2)	0.129 6(2)	5.2
N(4)	0.542 1(5)	−0.601 0(2)	0.175 4(2)	3.2
C(4-α)	0.600 5(7)	−0.619 9(3)	0.133 9(3)	3.5
C(4-β)	0.768 8(7)	−0.556 0(3)	0.127 1(3)	3.9
C(4-γ)	0.698 3(7)	−0.489 8(3)	0.095 3(3)	4.1
C(4-δ1)	0.715 7(9)	−0.480 6(4)	0.032 2(3)	5.7
C(4-δ2)	0.612 0(9)	−0.438 9(4)	0.125 1(3)	6.1
C(4-ε1)	0.651 1(9)	−0.423 0(4)	0.002 9(3)	7.0
C(4-ε2)	0.544 3(10)	−0.382 3(4)	0.094 4(4)	8.0
C(4-ξ)	0.561 4(10)	−0.374 4(5)	0.033 0(4)	8.8
C(4)	0.740 6(6)	−0.690 5(3)	0.151 7(3)	3.2
O(4)	0.831 0(5)	−0.718 3(2)	0.117 2(2)	5.0
N(5)	0.706 6(5)	−0.722 6(2)	0.205 5(2)	3.6
C(5-α)	0.758 3(6)	−0.794 9(3)	0.222 7(3)	3.1
C(5)	0.666 3(7)	−0.856 8(3)	0.194 9(3)	4.2
O(5)	0.708 5(5)	−0.920 9(2)	0.201 4(2)	6.1
C(5-β)	0.769 6(7)	−0.803 6(3)	0.292 5(3)	3.8
C(5-γ)	0.883 9(8)	−0.752 5(4)	0.319 1(3)	5.6
S(5)	0.915 9(3)	−0.765 3(2)	0.400 4(1)	9.2
C(5-δ)	1.011 6(17)	−0.848 3(7)	0.400 6(5)	17.1
N(6)	0.645 5(6)	−0.839 7(3)	0.165 6(3)	7.4
C(D1)	0.192 5(9)	−0.395 6(6)	0.483 8(5)	5.7
C(D2)	0.400 9(11)	−0.298 7(4)	0.509 1(4)	9.8
S(D)	0.373 5(2)	−0.393 9(1)	0.516 5(1)	9.0
O(D)	0.466 8(6)	−0.430 5(3)	0.469 8(2)	7.7

Acknowledgements

This research was supported by the University Grants Commission and the Department of Science and Technology. T. S. S. and B. V. V. P. were supported by fellowships from the I.C.M.R. and C.S.I.R., respectively. P. B. is the recipient of a U.G.C. Career Award.

Table 2. Backbone and side chain conformational angles (°) [a]

Residue	φ	ψ	ω	χ^1	$\chi^{2,1}$	$\chi^{2,2}$	χ^3
Aib	−54	−46	−169				
Aib	−60	−35	−173				
Phe	−84	−7	170	−66	82	−96	
Met	−82	−8		−65	−175		73

[a] Nomenclature recommended by the IUPAC–IUB Commission on Biochemical Nomenclature (ref. 27).

Table 3. Hydrogen bond parameters

Type	Donor D–H	Acceptor A	A···D(Å)	A···H(Å)	H–$\hat{D}$···A(°)	D–$\hat{H}$···A(°)
Intrapeptide	N(4)–H(4)	O(2-1)	3.05	2.24	28.6	139.3
	N(5)–H(5)	O(2)	3.06	2.14	12.2	162.4
	N(6)–H_6^1 ~ H(6-1)	O(3)	3.35	2.34	10.2	165.2
Peptide–DMSO	N(1)–H(1)	O(D)	2.85	1.90	11.7	162.3
Interpeptide	N(6)–H_6^{2*} ~ H(6-2*)	O(1-1)	3.09	2.21	20.7	150.4
	N(2)–H(2)	O(5-*)	2.98	2.09	4.0	174.2

* Corresponds to molecule at $1 - x, y - \frac{1}{2}, \frac{1}{2} - z$.

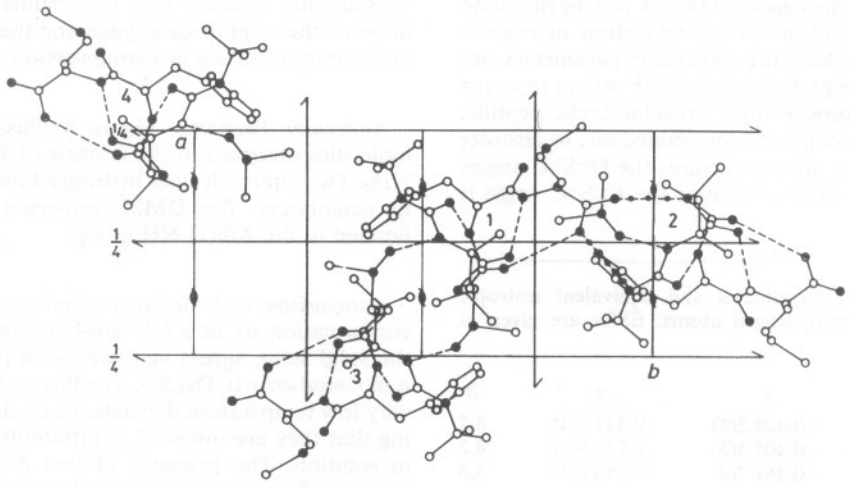
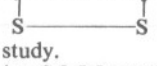


Figure 4. Crystal packing viewed down the *c* axis. Only four molecules (peptide + DMSO) numbered 1—4 are shown. Molecule 1 is at *x*, *y*, *z*; 2 at $1-x$, $\frac{1}{2}+y$, $\frac{1}{2}-z$, 3 at $\frac{1}{2}+x$, $\frac{1}{2}-y$, $-z$ and 4 at $\frac{1}{2}-x$, $1-y$, $\frac{1}{2}+z$. Intramolecular (– – – –) and intermolecular (— · — · —) hydrogen bonds are indicated

Table 4. DMSO geometries in crystal structures

	1	2	3	4	5 a	5 b
S(D)–O(D)	1.48	1.51	1.49	1.48	1.51 (1.57)	1.50 (1.52)
S(D)–C(D1)	2.00	1.78	1.78	1.79	1.75 (1.79)	1.74 (1.59)
S(D)–C(D2)	1.76	1.77	1.76	1.75	1.85 (1.65)	1.73 (1.59)
C(D1)–S(D)–O(D)	105	106	103	104	108 (103)	106 (113)
C(D2)–S(D)–O(D)	108	106	107	107	103 (110)	109 (117)
C(D1)–S(D)–C(D2)	105	98	92	96	98 (105)	100 (114)

1. c(L-Val-L-Pro-Gly-L-Val-Gly)$_3$·DMSO.[37]
2. 1 :1 Solvate 2-(bromotelluro)benzamide-DMSO.[38]
3. Boc-Cys-Pro-Aib-Cys—NHMe DMSO.[39]



4. This study.
5. 4-Amino-2,2,5,5-tetrakis(trifluoromethyl)-3-imidazoline·DMSO.[40]

Values under a and b are for two independent molecules in the asymmetric unit and those in parentheses are due to positional disorder at sulphur.

References

1 J. S. Morley, *Annu. Rev. Pharm. Toxicol.*, 1980, **20**, 81.
2 R. J. Miller, K. J. Chang, P. Cuatrecasas, S. Wilkinson, L. Lowe, C. Beddell and R. Follenfant, in ' Centrally Acting Peptides,' ed. J. Hughes, Macmillan, London, 1978, p. 195.
3 D. H. Coy and A. J. Kastin, *Pharm. Therap.*, 1980, **10**, 657.
4 J. Hughes, *Brain Res.*, 1975, **88**, 295.
5 J. Hughes, T. W. Smith, H. W. Kosterlitz, L. A. Fothergill, B. A. Morgan, and H. R. Morris, *Nature*, 1975, **253**, 577.
6 F. A. Gorin, T. M. Balasubramanian, C. D. Barry, and G. R. Marshall, *J. Supramol. Struct.*, 1978, **9**, 27.
7 B. P. Roques, C. Garbay-Jaureguiberry, R. Oberlin, M. Anteunis, and A. K. Lala, *Nature*, 1976, **262**, 778.
8 C. R. Jones, W. A. Gibbons, and V. Garsky, *Nature*, 1976, **262**, 779.
9 M. A. Khaled, M. M. Long, W. D. Thompson, R. J. Bradley, G. B. Brown, and D. W. Urry, *Biochem. Biophys. Res. Commun.*, 1977, **76**, 224.
10 P. W. Schiller and J. St-Hilaire, *J. Med. Chem.*, 1980, **23**, 290.
11 E. R. Stimson, Y. C. Meinwald, and H. A. Scheraga, *Biochemistry*, 1979, **18**, 1661.
12 A. J. Fischman, M. W. Riemen, and D. Cowburn, *FEBS Lett.*, 1978, **94**, 236.
13 P. Tancrede, R. Deslauriers, W. H. McGregor, E. Ralston, D. Sarantakis, R. L. Somorjai, and I. C. P. Smith, *Biochemistry*, 1978, **17**, 2905.
14 I. Z. Siemion, Z. Szewczuk, Z. S. Herman, and Z. Stachura, *Mol. Cell. Biochem.*, 1981, **34**, 23.
15 B. A. Morgan, in 'Amino Acids, Peptides and Proteins,' ed. R. C. Sheppard, Spec. Period. Rep., Chem. Soc. London, 1978, p. 481.
16 F. A. Gorin, T. M. Balasubramanian, T. J. Cicero, J. Schwietzer, and G. R. Marshall, *J. Med. Chem.*, 1980, **23**, 1113.
17 P. W. Schiller, B. Egginmann, J. DiMaio, C. Lemineux, and T. M. D. Nauyen, *Biochem. Biophys. Res. Commun.*, 1981, **101**, 337.
18 J. DiMaio and P. W. Schiller, *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 7162.
19 R. Nagaraj and P. Balaram, *FEBS Lett.*, 1978, **96**, 273.
20 R. Nagaraj, T. S. Sudha, S. Shivaji, and P. Balaram, *FEBS Lett.*, 1979, **106**, 271.
21 T. S. Sudha and P. Balaram, *FEBS Lett.*, 1981, **134**, 32.
22 G. Germain, P. Main, and M. M. Woolfson, *Acta Crystallogr.*, 1971, **A27**, 368.
23 B. V. V. Prasad, N. Shamala, R. Nagaraj, R. Chandrasekaran, and P. Balaram, *Biopolymers*, 1979, **18**, 1635.
24 D. T. Cromer and J. T. Waber, *Acta Crystallogr.*, 1965, **18**, 104.
25 R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, 1965, **42**, 3175.
26 W. C. Hamilton, *Acta Crystallogr.*, 1959, **12**, 609.
27 IUPAC–IUB Commission on Biochemical Nomenclature, *Biochemistry*, 1970, **9**, 3471.
28 R. Nagaraj, N. Shamala, and P. Balaram, *J. Am. Chem. Soc.*, 1979, **101**, 16.
29 R. Nagaraj and P. Balaram, *Acc. Chem. Res.*, 1981, **14**, 356.
30 G. N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.*, 1968, **23**, 283.
31 C. M. Venkatachalam, *Biopolymers*, 1968, **6**, 1425.
32 A. B. Mauger, O. A. Stuart, R. J. Highet, and J. V. Silverton, *J. Am. Chem. Soc.*, 1982, **104**, 174.
33 J. N. Brown and R. G. Teller, *J. Am. Chem. Soc.*, 1976, **98**, 7565.
34 G. D. Smith and J. F. Griffin, *Science*, 1978, **199**, 1214.
35 A. Mathieson, *Acta Crystallogr.*, 1952, **5**, 332.

36 R. E. Stenkamp and L. H. Jensen, *Acta Crystallogr.*, 1975, **B31**, 857.
37 W. K. Cook, H. Einspahr, T. L. Trapane, D. W. Urry, and C. E. Bugg, *J. Am. Chem. Soc.*, 1980, **102**, 5502.
38 L. Dupont, O. Dideberg, J. Lamotte, J. I. Pielte, *Acta Crystallogr.*, 1979, **B35**, 849.
39 A. Ravi, B. V. V. Prasad, and P. Balaram, *J. Am. Chem. Soc.*, in the press.
40 S. K. Arora, *Acta Crystallogr.*, 1981, **B37**, 2052.
41 T. S. Sudha and P. Balaram, *Int. J. Pept. Prot. Res.*, 1983, in the press.

Received 2*nd July* 1982; *Paper* 2/1109