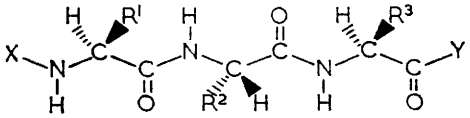
The Circular Dichroism of *N*-Thiobenzoyl-L-α-amino-acids. Part V.[1] Solvent-dependent Circular Dichroism of Terminal *N*-Thiobenzoyl Oligopeptides

By **G. C. Barrett**, Department of Chemistry, West Ham College of Technology, Romford Road, London E.15

The solvent-dependent c.d. behaviour of terminal *N*-thiobenzoyl derivatives of twenty peptides has been determined. The terminal *N*-thiobenzoyl derivative of a peptide carrying an *N*-terminal L-α-amino-acid residue shows positive $n \rightarrow \pi^*$ c.d. in aqueous solution; the derivative of a peptide carrying an *N*-terminal L-α-imino-acid residue (L-proline) shows negative $n \rightarrow \pi^*$ c.d. in water. The penultimate amino-acid residue in these derivatives also contributes to their c.d., directly by perturbation of the $n \rightarrow \pi^*$ thiobenzamide transition and indirectly by influencing conformational equilibria involving the substituted *N*-terminal amino-acid residue.

Optical rotatory dispersion and c.d. studies have yielded much information concerning the tertiary structures of native polypeptides and proteins, and of synthetic polypeptides, in solution.[3] The information has been derived from measurements of the optical rotatory power developed by asymmetric perturbation of electronic transitions within side-chain chromophores [R^1, R^2, and R^3 in (I)]; *e.g.* the disulphide grouping of cystine residues, and aromatic and heteroaromatic groupings, with $\lambda_{max.}$ *ca.* 230—300 nm., and of electronic transitions responsible for the absorption features at shorter wavelengths of the amide grouping repeated along the peptide backbone in these compounds (I; X and Y = extensions of the peptide chain).



(I)

The o.r.d. of short peptides, in aqueous solution at various pH values, has been thoroughly analysed, with similar structural objectives;[4] the side-chains of constituent amino-acid residues of the peptides studied in this way showed no absorption maximum in the near-ultraviolet region, permitting conclusions to be drawn concerning the contributions of individual chromophores ($n \rightarrow \pi^*$ transitions of amide and carboxy-groupings) to the o.r.d. of small peptides.

The present study exemplifies a different approach, in which structural assignments are based on o.r.d. or c.d. studies of 'chromophorically substituted'[5] peptides (*e.g.* I; X or Y = substituent with $\lambda_{max.}$ at wavelengths longer than *ca.* 270 nm.) designed to introduce Cotton effects in wavelength regions not only more readily accessible instrumentally, but also removed from regions in which amide and side-chain chromophores of peptides show absorption maxima. Blout,[6] in discussing the Cotton effects shown by dye molecules asymmetrically adsorbed on dissymmetric helical macromolecules, has suggested that local conformations at a particular site in a biological macromolecule might be determined by using the dye molecule as a 'molecular probe', since the finer details of the Cotton effect curves of such a system will reflect these structural features. However, extensive study of relationships between structure and o.r.d. or c.d. data for related systems must precede any attempt to assign conformations in this way; a viable first step appears to be the interpretation of features of the longest-wavelength Cotton effect of a 'chromophorically substituted' peptide in terms of the conformation adopted by the short peptide sequence adjacent to the chromophore, after due account is taken of the o.r.d. or c.d. data of a series of model compounds. The c.d. technique should be particularly suitable for such studies, since quantitative data can be derived more readily from c.d. curves associated with well separated isotropic absorption maxima than from corresponding o.r.d. curves.[5]

Representative derivatives of this type which have been subjected to o.r.d. and c.d. study are terminal *N*-phthaloyl-[7] and *N*-ethoxythiocarbonyl-peptides.[8] Assignment of absolute configuration to the *N*-terminal amino-acid residue of a polypeptide follows [7,8] from the sign of the Cotton effect of either derivative; no attempt was made to deduce additional structural information from the data. Crabbé and Halpern,[9] in an o.r.d. study of *N*-(3-oxo-5,5-dimethylcyclohex-1-enyl) derivatives obtained by the condensation of dimedone with representative amino-acids and peptides, found that the position of the o.r.d. extrema and the amplitude of the Cotton effect were dependent upon structural features at the asymmetric centre adjacent to the chromophore in the derivatives. The statement 'these specific properties might be of some help in structural and synthetic studies performed in the protein and peptide fields' was included in the preliminary communication [9]

[1] Part IV, G. C. Barrett and A. R. Khokhar, preceding paper.
[2] Preliminary communication, G. C. Barrett, *Chem. Comm.*, 1968, 40.
[3] (*a*) S. Beychok, *Science*, 1966, **154**, 1288; S. Beychok, *Proc. Nat. Acad. Sci. U.S.A.*, 1965, **53**, 999; E. Schechter and E. R. Blout, *ibid.*, 1964, **51**, 695; chapters by S. Beychok, J. T. Yang, in 'Poly-α-amino-acids; Protein Models for Conformational Studies,' ed. G. D. Fasman, Edward Arnold, London, and Marcel Dekker, New York, 1967, pp. 239, 293.
[4] I. Schechter and A. Berger, *Biochemistry*, 1966, **5**, 3362; J. Beacham, V. T. Ivanov, P. M. Scopes, and D. R. Sparrow, *J. Chem. Soc.* (*C*), 1966, 1449; D. R. Dunstan and P. M. Scopes, *J. Chem. Soc.* (*C*), 1968, 1585.
[5] C. Djerassi, *Proc. Chem. Soc.*, 1964, 314; G. Snatzke, *Angew. Chem. Internat. Edn.*, 1968, **7**, 14.
[6] E. R. Blout, *Biopolymers*, 1964, Symposium No. 1, 397.
[7] C. Djerassi, E. Lund, E. Bunnenberg, and J. C. Sheehan, *J. Org. Chem.*, 1961, **26**, 4509.
[8] R. C. Sheppard, *Coll. Czech. Chem. Comm.*, 1962, **27**, 2251.
[9] P. Crabbé and B. Halpern, *Chem. and Ind.*, 1965, 346.

describing this work, but has not been amplified further in the full paper.[10] The dimedone condensation product with glycyl-L-leucine ethyl ester was reported [9] to exhibit no Cotton effect (near 280 nm.), implying that the o.r.d. or c.d. of a 'chromophorically substituted' peptide might be expected to be derived only from the terminal amino-acid residue; however, this statement has now [10] been corrected, in agreement with results of the present study.[2]

Yamaoka, Idelson, and Blout have reported [6] the preparation and o.r.d. study of a poly-(γ-benzyl-L-glutamate) with a dye molecule attached covalently at one end of the polypeptide chain (*e.g.* I; X = polypeptide chain, Y = dye molecule). The wording of the description [6] implies that the observed o.r.d. is found to be compounded of a contribution from the helical polypeptide as well as from the perturbation by adjacent asymmetric centres of electronic transitions within the chromophore; so far, detailed description of this work remains unpublished.

Against this background, and with extensive data on the solvent-dependent o.r.d. and c.d. of *N*-thiobenzoyl-L-α-amino-acids already available,[11–16] it seemed that a detailed c.d. study of terminal *N*-thiobenzoyl oligopeptides (I; X = PhCS, Y = peptide residue) would be useful, for the following reasons: first, in confirming whether reliable stereochemical assignments can be made to *N*-terminal amino-acid residues in peptides by c.d. measurements with 'chromophorically substituted' peptides, in the light of the recently-discovered solvent-dependence of the c.d. of 'chromophoric derivatives',[12,13,15–17] and of the relatively few examples [7,8] which have been reported in support of empirical relationships; second, whether the characteristic [15] solvent-dependent c.d. of an *N*-thiobenzoyl-L-α-amino-acid was also exhibited by a corresponding terminal *N*-thiobenzoyl peptide, and could therefore be usefully exploited in a novel *N*-terminal analysis technique in the polypeptide and protein field; and third, whether the c.d. data of a large number of terminal *N*-thiobenzoyl oligopeptides could be interpreted to yield information concerning conformational preferences in small peptides.

EXPERIMENTAL

Peptides.—Many of the dipeptides were obtained from commercial sources; new peptides, and protected derivatives, prepared for the present study, were *L-seryl-L-prolylglycine monohydrate*, m.p. 174—176° (from aq. EtOH) (Found: C, 43·45; H, 7·5; N, 14·85. $C_{10}H_{17}N_3O_5,H_2O$ requires C, 43·3; H, 6·9; N, 15·15%), obtained by hydrogenation (1 atmos.) of the corresponding *N*-benzyloxycarbonyl tripeptide (cyclohexylammonium salt,[18] m.p. 120—122°) in 80% aq. acetic acid over 10% palladium–charcoal; *L-[N(1)-benzylhistidyl]glycine*, m.p. 192—193° (decomp.) (from 70% ethanol) (Found: C, 59·2; H, 5·7; N, 17·75. $C_{15}H_{18}N_4O_3$ requires C, 59·6; H, 6·0; N, 18·55%), obtained by hydrogenation of the corresponding *N*-benzyloxycarbonyl dipeptide benzyl ester [19] as described before. Hydrogenation of *N-benzyloxycarbonyl-L-valyl-D-leucine benzyl ester*, m.p. 132° (from chloroform–ether) (Found: C, 69·0; H, 7·05; N, 6·3. $C_{26}H_{34}N_2O_5$ requires C, 68·7; H, 7·55; N, 6·15%) gave the *dipeptide*, m.p. 280—285° (decomp.), $[\alpha]_D^{20}$ +72·1° (*c* 1·9 in H_2O) (Found: N, 11·25. $C_{11}H_{22}N_2O_3,H_2O$ requires N, 11·3%), as an amorphous solid from benzene–ether. L-Valyl-L-phenylalanine (from Z-Val-Phe-OEt [20] by successive saponification and hydrogenation), L-valylglycine (from Z-Val-Gly-OBZL, m.p. 129—130°; DL-isomer,[21] m.p. 135°), and L-prolylglycine (from Z-Pro-Gly-OBZL [22]) were prepared similarly; peptide couplings (dicyclohexylcarbodi-imide in acetonitrile at room temperature) and saponification [N-sodium hydroxide (1 equiv.) in acetone at room temp., 1 hr.] were accomplished by standard procedures.[23]

Terminal N-Thiobenzoyl peptides.—A solution of the peptide and carboxymethyl dithiobenzoate [24] (each 1 equiv.) in N-sodium hydroxide (2 equiv.) was kept at room temp. for 4—12 hr.; the shorter period was appropriate in most cases for completion of the reaction, as indicated by the colour change from red towards yellow. The solution was acidified and extracted with ether; evaporation of the dried ($MgSO_4$) extracts, and crystallisation of the residue, gave the crystalline terminal *N*-thiobenzoyl peptides listed in Table 1. *N*-Thiobenzoyl-L-valyl-D-leucine could not be induced to crystallise, and was purified by chromatography over silica gel.[12]

N-Thiobenzoyl derivatives of glycyl-L-valylglycine benzyl ester, DL-leucylglycyl-D-leucine benzyl ester, and L-valyl-L-leucylglycine benzyl ester were prepared [25] by peptide couplings (dicyclohexylcarbodi-imide in acetonitrile) of *N*-thiobenzoyl dipeptides with amino-acid benzyl esters; the last-named derivative was obtained as an oil.

U.V. Spectra and C.D. of Terminal N-Thiobenzoyl Peptides.—U.v. and c.d. measurements were made with a Unicam SP 700 spectrophotometer and a Roussel-Jouan Dichrographe, respectively. U.v. spectral data of the derivatives were similar to those of *N*-thiobenzoylamino-acids,[11–16] *e.g.* $\lambda_{max.}$ (Et_2O) *ca.* 240, 288, and 395 nm. (log ε *ca.* 4·1, 3·9, and 2·4); the location of the long-wavelength absorption maximum showed the same marked solvent-dependence.[12,16] C.d. data for seventeen terminal *N*-thiobenzoyl peptides are listed in Table 2; *N*-thiobenzoyl-DL-leucylglycyl-D-leucine benzyl ester showed no c.d. near 380 nm. in the seven solvents mentioned in Table 2, while *N*-thiobenzoylglycyl-L-valylglycine benzyl ester and *N*-thiobenzoyl-

[10] P. Crabbé, B. Halpern, and E. Santos, *Tetrahedron*, 1968, **24**, 4315.
[11] B. Sjöberg, B. Karlen, and R. Dahlbom, *Acta Chem. Scand.*, 1962, **16**, 1071.
[12] G. C. Barrett, *J. Chem. Soc.*, 1965, 2825.
[13] G. C. Barrett, *J. Chem. Soc.* (*C*), 1966, 1071.
[14] G. C. Barrett, *J. Chem. Soc.* (*C*), 1967, 1.
[15] G. C. Barrett in 'Some Newer Physical Methods in Structural Chemistry,' ed. R. Bonnett and J. G. Davis, United Trade Press, London, 1967, p. 201.
[16] E. Bach, A. Kjaer, R. Dahlbom, T. Walle, B. Sjöberg, E. Bunnenberg, C. Djerassi, and R. Records, *Acta Chem. Scand.*, 1966, **20**, 2781.
[17] W. S. Briggs and C. Djerassi, *Tetrahedron*, 1965, **21**, 3455.
[18] G. C. Barrett and H. N. Rydon, unpublished work.
[19] D. Theodoropoulos and G. Fölsch, *Acta Chem. Scand.*, 1958, **12**, 1955, 2043.
[20] J. M. Johnson and J. A. Stock, *J. Chem. Soc.*, 1962, 3806.
[21] J. Heintges, H. Jung, and J. Heideberg, *Z. Naturforsch.*, 1962, **176**, 143.
[22] L. Grimm and W. Grassman, *Z. physiol. Chem.*, 1964, **337**, 161.
[23] J. P. Greenstein and M. Winitz, 'Chemistry of the Amino-acids,' Wiley, New York, 1961, vol. II, ch. 10.
[24] F. Kurzer and A. Lawson, *Org. Synth.*, 1962, **42**, 100.
[25] G. C. Barrett, unpublished work.

TABLE 1

Terminal *N*-thiobenzoyl peptides

N-Thiobenzoyl derivative of	M.p.	Cryst. from	Found (%) C	H	N	S	Formula	Calc. (%) C	H	N	S
L-Alanylglycine	56—58°	Et_2O	54·1	5·85	10·35	11·8	$C_{12}H_{14}N_2O_3S$	54·1	5·3	10·5	12·05
Glycyl-L-alanine	101	Et_2O	54·0	5·35	10·3	12·1	$C_{12}H_{14}N_2O_3S$	54·1	5·3	10·5	12·05
L-Valylglycine	79—86	C_6H_6	64·55	6·65	7·9	8·6	$C_{14}H_{18}N_2O_3S,C_6H_6$	64·5	6·5	7·55	8·6
Glycyl-L-valine	157	Et_2O	57·45	5·8	9·6	10·55	$C_{14}H_{18}N_2O_3S$	57·1	6·15	9·5	10·9
L-Valyl-L-leucine	159	Et_2O	61·9	7·55	8·4	9·05	$C_{18}H_{26}N_2O_3S$	61·7	7·5	8·0	9·15
L-Valyl-L-phenylalanine	97	Et_2O–petroleum	64·95	6·6	7·05	7·9	$C_{21}H_{24}N_2O_3S$	65·6	6·3	7·3	8·35
L-Leucylglycine	76—80	C_6H_6	65·2	6·75	7·3	8·5	$C_{15}H_{20}N_2O_3S,C_6H_6$	65·25	6·8	7·25	8·3
Glycyl-L-leucine	176—178	Et_2O	58·65	6·5	9·1	10·2	$C_{15}H_{20}N_2O_3S$	58·4	6·55	9·1	10·4
L-Leucyl-L-alanine	81	Et_2O–petroleum	60·2	7·25	8·15	9·15	$C_{16}H_{22}N_2O_3S$	59·6	6·9	8·7	9·95
L-Prolylglycine	174—176 (decomp.)	Et_2O	56·85	5·45	9·5		$C_{14}H_{16}N_2O_3S$	57·5	5·5	9·6	
Glycyl-L-proline	173—174	Et_2O	57·5	5·55	9·6	11·15	$C_{14}H_{16}N_2O_3S$	57·5	5·5	9·6	10·95
Glycyl-L-phenylalanine	144—145	Et_2O	63·1	5·0	8·1	9·1	$C_{18}H_{18}N_2O_3S$	63·15	5·3	8·2	9·35
Glycylglycylglycine	167—168	H_2O	50·3	4·7	13·8	10·35	$C_{13}H_{15}N_3O_4S$	50·45	4·9	13·6	10·35
L-[N(1)-Benzylhistidyl]glycine	204—205 (decomp.)	Pr^iOH	62·85	5·6	13·35	7·85	$C_{22}H_{22}N_4O_3S$	62·55	5·25	13·25	7·6
L-Seryl-L-prolylglycine	113	H_2O	51·45	5·9	10·5	7·95	$C_{17}H_{21}N_3O_5S,H_2O$	51·35	5·85	10·55	8·05
Glycyl-L-serine	122	Et_2O	51·1	5·1	10·25	11·3	$C_{12}H_{14}N_2O_4S$	51·05	5·0	9·95	11·35
Glycyl-L-methionine	150	Et_2O	51·4	5·65	8·85	19·5	$C_{14}H_{18}N_2O_3S_2$	51·5	5·55	8·6	19·65
Glycyl-L-valylglycine benzyl ester	152—154	$CHCl_3$–Et_2O	62·4	6·25	9·8	7·4	$C_{23}H_{27}N_3O_4S$	62·55	6·2	9·5	7·25
DL-Leucylglycyl-D-leucine benzyl ester	109—111	Et_2O	66·3	7·1	8·45	5·85	$C_{28}H_{37}N_3O_4S$	65·7	7·3	8·2	6·25

TABLE 2

$n \rightarrow \pi^*$ C.d. of terminal *N*-thiobenzoyl peptides and related derivatives [a]

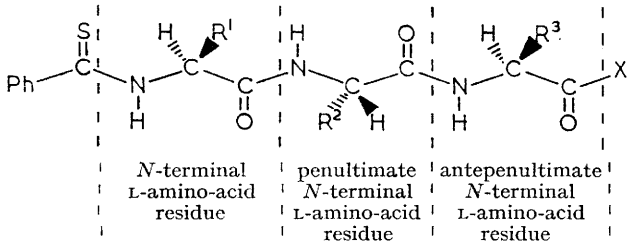
N-Thiobenzoyl derivative of	Δε Values at c.d. maxima in different solvents ($\lambda_{max.}$ in parentheses) Water	MeOH	EtOH	Pr^iOH	Bu^tOH	Dioxan	Ether
L-Valine	+1·25 (358)	+0·42 (383) [b,c]	+0·05 (420); −0·07 (372)	−0·20 (384)	−0·39 (386)	−0·24 (383) [c]	−1·44 (397) [b]
L-Valylglycine [d]	+1·68 (358)	+0·97 (382)	+0·68 (387)	+0·43 (395)	+0·03 (425); −0·25 (366)	−0·23 (386)	−1·00 (394)
L-Valyl-L-leucine	+1·75 (358)	+1·51 (378)	+1·24 (383)	+0·98 (385)	+0·36 (396)	+0·70 (384)	−0·62 (400)
L-Valyl-D-leucine [e]	+1·20 (356)	+0·88 (383)	+0·67 (390)	+0·46 (388)	+0·11 (415); −0·15 (350)	−0·50 (383)	−0·60 (390)
L-Valyl-L-phenylalanine	+1·62 (358)	+1·13(379)	+0·91 (378)	+0·64 (380)	+0·12 (414)	+0·10 (415)	−0·39 (396)
L-Leucine	+3·12 (355)	+1·26 (375) [f]	+1·13 (379)	+0·73 (386)	+0·44 (400)	+1·47 (393)	−0·83 (389) [f]
L-Leucinamide	Insoluble	+1·92 (379) [c]	+1·69 (384)	+1·28 (386)	+0·61 (390)	+0·70 (386) [c]	−0·23 (396)
L-Leucylglycine [d]	+3·30 (360)	+1·76 (380)	+1·43 (386)	+0·95 (385)	+0·39 (402)	+0·69 (390)	−0·63 (389)
L-Leucyl-L-alanine	+2·62 (355)	+1·61 (376)	+1·43 (383)	+1·08 (385)	+0·42 (391)	+0·79 (387)	−0·08 (396)
L-Alanine	+0·94 (357)	+0·09 (424) [b,f]; −0·14 (367)	−0·42 (374)	−0·47 (382)	−0·65 (382)	−0·47 (386)	−1·32 (391) [f]
L-Alanylglycine	+1·35 (359)	+0·12 (390)	−0·28 (375)	−0·60 (378)	−1·05 (379)	−1·21 (388)	−1·42 (394)
L-[N(1)-Benzyl]histidine	+0·62 (370)	+0·07 (414) [b]; −0·15 (364)	+0·41 (388)	+0·62 (392)	+1·16 (383)	+0·38 (403)	Insoluble
L-[N(1)-Benzylhistidyl]glycine	+0·79 (370)	+0·47 (386)	+0·79 (388)	+0·94 (388)	+1·11 (389)	+0·54 (398)	Insoluble
L-Serine	+0·46 (364)	−0·14 (382)	−0·60 (384)	−0·69 (384)	−0·52 (389)	−1·64 (390)	−1·75 (396) [f,g]
L-Seryl-L-prolylglycine	+0·66 (358)	+0·36 (378)	+0·15 (405)	+0·13 (410); −0·04 (360)	−0·44 (375)	−0·96 (392)	Insoluble
L-Proline	−0·16 (367); +0·16 (330)	−0·29 (390) [c,f]; +0·10 (345)	−0·36 (396); +0·24 (345)	−0·36 (395); +0·24 (347)	−0·46 (391); +0·05 (348)	−0·56 (396) [e]; +0·11 (356)	−1·03 (402) [f,g]; +0·14 (356)
L-Prolylglycine	−0·44 (370)	−0·40 (396); +0·20 (345)	−0·39 (397); +0·20 (350)	−0·37 (398); +0·20 (355)	−0·37 (400); +0·19 (360)	−0·39 (400); +0·19 (360)	−0·40 (404)
Glycyl-L-methionine	−0·14 (375)	+0·15 (367)	+0·14 (370)	+0·11 (376)	+0·14 (385)	+0·22 (386)	+0·12 (386)
Glycyl-L-valine	+0·11 (355)	+0·19 (370)	+0·06 (376)	−0·03 (408); +0·07 (355)	−0·01 (410); +0·06 (360)	+0·17 (388)	+0·15 (383)
Glycyl-L-alanine	−0·04 (390)	+0·13 (370)	+0·10 (376)	+0·11 (378)	+0·10 (390)	−0·06 (368)	+0·07 (388)
Glycyl-L-leucine	−0·13 (365)	+0·10 (363)	+0·05 (370)	+0·05 (360)	0·00	+0·16 (388)	+0·12 (387)
Glycyl-L-phenylalanine	+0·13 (370)	−0·04 (408); +0·02 (360)	−0·05 (425); +0·05 (365)	−0·05 (410); +0·05 (360)	−0·07 (390)	+0·23 (385)	+0·05 (390)
Glycyl-L-serine	+0·13 (345)	+0·10 (378)	+0·12 (384)	+0·09 (388)	+0·12 (388)	−0·09 (382)	+0·05 (392)
Glycyl-L-proline	−0·06 (420)	+0·13 (376)	+0·14 (385)	−0·06 (420)	+0·05 (395)	+0·06 (405); −0·02 (365)	−0·12 (390)

[a] Some of these data were reported in ref. 2. [b] From ref. 13. [c] *Cf.* ref. 16. [d] Data for benzene solvate; identical with corresponding data for the non-solvated, non-crystalline, derivative. [e] Oil. [f] From ref. 12. [g] Corrected figure.

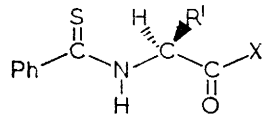
L-valyl-L-leucylglycine benzyl ester exhibited solvent-dependent c.d. behaviour similar to that of the *N*-thiobenzoyl derivatives of glycyl-L-valine and L-valyl-L-leucine (Table 2), respectively. Points of interest in the data in Table 2 (' double-humped ' c.d. curves,[15] and wavelength locations of c.d. maxima) of little relevance to the following discussion will be referred to in later parts of the present series.

DISCUSSION

The solvent-dependent c.d. behaviour of seventeen terminal *N*-thiobenzoyl peptides is summarised in Table 2 (c.d. data for an additional three peptides are included in the Experimental section), together with corresponding data for reference compounds relevant to the present discussion. The c.d. behaviour in various solvents of the terminal *N*-thiobenzoyl derivative (II) of a peptide is generally similar to that of the *N*-thiobenzoyl derivative (III; X = OH) of the L-α-amino-acid corresponding to the *N*-terminal amino-acid residue of the peptide.



(II)



(III)

Further, the sign of the $n \rightarrow \pi^*$ c.d. of a terminal *N*-thiobenzoyl peptide is diagnostic of the absolute configuration of the *N*-terminal amino-acid residue only in specified solvents. The results show that in aqueous solution, *N*-thiobenzoyl derivatives of peptides with an *N*-terminal L-prolyl residue must be expected to show negative c.d. near 370 nm. while derivatives of peptides with other *N*-terminal L-amino-acid residues show strong positive c.d. near this wavelength. Although there appears to be a completely satisfactory generalisation for ether solutions too, namely, that *N*-thiobenzoyl derivatives of peptides with an L-amino-acid residue at the *N*-terminus show negative c.d. near 395 nm., comparison of data for the derivatives of L-leucylglycine and L-leucyl-L-alanine shows that the penultimate amino-acid residue can exert a considerable influence on the c.d. developed at the *N*-terminus in these compounds, reducing the c.d. almost to zero in the latter case, for measurements with ether solutions. Also, the fact that *N*-thiobenzoyl derivatives of a number of L-α-amino-acids show almost no c.d.,[12] or even positive c.d.,[13] in ethereal solution, makes it likely that certain combinations of L-amino-acids in the *N*-terminal and penultimate residue positions in a peptide will yield *N*-thiobenzoyl derivatives showing positive c.d. in ether; an empirical relationship linking negative $n \rightarrow \pi^*$ c.d. in ether solutions of the present derivatives with the L-configuration of the *N*-terminal amino-acid residue of a peptide cannot be presumed from the present results. Although methanol appears to be as satisfactory as water as solvent for the present purpose, *i.e.* only L-proline falls out of line in a c.d.–absolute configuration correlation, the c.d. data for *N*-thiobenzoyl-L-alanylglycine show that the solvent change from methanol to ethanol is sufficient to cause inversion of the sign of the c.d. of this compound. There is a strong possibility, therefore, that the sign of the c.d., in methanol, of the *N*-thiobenzoyl derivative of a peptide carrying an *N*-terminal L-alanyl residue, may depend upon the structure (and absolute configuration) of the amino-acid occupying the penultimate residue position; there are already indications elsewhere in the data in Table 2 of such an influence of the penultimate residue. In summary, reliable assignment of absolute configuration to the *N*-terminal amino-acid residue of a peptide follows from the c.d. of its terminal *N*-thiobenzoyl derivative, though only for measurements with aqueous solutions and with terminal L-proline-containing peptides as exceptions to this correlation.

In comparison with data for *N*-thiobenzoyl-L-α-amino-acids (III; X = OH), trends in the c.d. data for corresponding amino-acid amide derivatives[16] (III; X = NH_2) imply that the structural change from carboxylic acid to carboxamide at the asymmetric centre in this series is accompanied by a shift towards more positive $\Delta\varepsilon$ values; we have confirmed reported data[16] for *N*-thiobenzoyl-L-leucinamide, and measured the c.d. of this compound in a number of other solvents (Table 2). The trend towards more positive c.d. values is clear; however, this trend is partly reversed in the further structural change involved in reaching the dipeptide (III; $R^1 = Bu^i$; X = OH → NH_2 → $NH{\cdot}CH_2{\cdot}CO_2H$), *N*-thiobenzoyl-L-leucylglycine. Comparison of c.d. data (Table 2) of an *N*-thiobenzoyl-L-α-amino-acid [PhCS–X–OH; X = Ala, Val, Pro, or N(1)-BZL-His] with those of its corresponding dipeptide with *C*-terminal glycine (PhCS–X–Gly–OH) reveals only a small shift, generally towards more positive $\Delta\varepsilon$ values, particularly for solutions in the more polar solvents. Therefore, it appears that the characteristic family of c.d. curves shown[15] by an *N*-thiobenzoyl-L-α-amino-acid in different solvents may be similar in general appearance to that of a terminal *N*-thiobenzoyl peptide carrying the same L-α-amino-acid as *N*-terminal amino-acid residue. In certain cases (*N*-thiobenzoyl-L-proline compared with *N*-thiobenzoyl-L-prolylglycine) the two sets of curves are practically identical, though in other cases (*N*-thiobenzoyl-L-alanine and *N*-thio-

benzoyl-L-alanylglycine), differences are more pronounced.

Replacement of glycine as penultimate *N*-terminal amino-acid residue by a D- or L-amino-acid in these dipeptides (*e.g.* PhCS-L-Val-Gly-OH to PhCS-L-Val-L-Leu-OH) has a marked effect on the resulting c.d. behaviour of the *N*-thiobenzoyl derivative. This is partly due to a c.d. contribution derived from the interaction of the asymmetric centre [at R^2 in (II)] of the penultimate amino-acid residue with the chromophore, since *N*-thiobenzoylglycyl-L-amino-acids (III; $R^1 = H$, X = L-amino-acid residue) exhibit significant c.d. (Table 2). In a given solvent, $\Delta\varepsilon$ values for *N*-thiobenzoylglycyl-L-amino-acids are usually small compared with those of corresponding *N*-thiobenzoyl-L-α-amino-acids, as would be expected from the greater separation of chromophore and asymmetric centre, and also from the greater conformational mobility, in the former series. No c.d. was detectable for solutions of *N*-thiobenzoyl-DL-leucylglycyl-D-leucine benzyl ester, implying that the antepenultimate amino-acid residue might be assumed not to contribute directly to the c.d. of a terminal *N*-thiobenzoyl peptide.

The sign of the c.d. near 390 nm. of an *N*-thiobenzoylglycyl-L-amino-acid is generally opposite to that of the corresponding *N*-thiobenzoyl-L-amino-acid; at least, this is so for ethereal solutions (a similar observation has been made for corresponding *N*-dimedonyl compounds [10]). This might be expected if the fully extended conformation [*e.g.* (II)] characteristic [26] of small peptides in solution is also adopted by the peptide residue of (II; X = extension of peptide chain). From the point of view of the chromophore, the asymmetric centre of an *N*-thiobenzoylglycyl-L-amino-acid [*cf.* (II; $R^1 = H$)] in an extended conformation is inverted relative to that in the corresponding *N*-thiobenzoyl-L-amino-acid, thus leading to oppositely signed c.d. curves for these two series.

Taken as a whole, c.d. data for *N*-thiobenzoylglycyl-L-amino-acids show large relative variations in magnitude; also, the generalisation that these compounds show oppositely signed c.d. relative to corresponding *N*-thiobenzoyl-L-amino-acids does not hold for every case studied, in every solvent used in the present work (Table 2). Comment has already been made on the influence of the penultimate amino-acid residue in undermining a correlation of sign of c.d. with the absolute configuration of the *N*-terminal amino-acid residue of a terminal *N*-thiobenzoyl peptide. The penultimate amino-acid residue must also be expected to distort the appearance of the family of c.d. curves given by the terminal *N*-thiobenzoyl derivative of a peptide (PhCS–X–Y–OH) in different solvents, in comparison with corresponding c.d. curves for the *N*-thiobenzoyl derivative of the *N*-terminal amino-acid itself (PhCS–X–OH), or with those of the *N*-thiobenzoyl derivative of the analogous peptide with glycine occupying the penultimate residue position (PhCS–X–Gly–···). The degree of distortion will depend upon the structure (and absolute configuration) of the penultimate amino-acid residue; the distortion resulting from successive replacements by different L-amino-acids at this position (*e.g.* PhCS-L-Val-L-Phe-··· to PhCS-L-Val-L-Leu-··· *etc.*) might be expected to reflect the c.d. behaviour of corresponding *N*-thiobenzoylglycyl-L-amino-acids (PhCS-Gly-L-Phe-···, PhCS-Gly-L-Leu-. . ., *etc.*, respectively), as a first approximation. As a consequence of this influence of the adjacent amino-acid residue, the *N*-terminal amino-acid residue of a peptide cannot be identified with any certainty by matching the c.d. behaviour of the terminal *N*-thiobenzoyl derivative of the peptide against corresponding c.d. curves for *N*-thiobenzoyl derivatives of the 'natural' amino-acids. Although the family of c.d. curves for an *N*-thiobenzoyl-L-α-amino-acid appears adequately distinctive,[15] implying that an 'unknown' D- or L-amino-acid may be identified from the solvent-dependent c.d. of its *N*-thiobenzoyl derivative (the technique is applicable in principle even to partly racemised amino-acids), a reliable *N*-terminal analysis of a peptide on the same basis would require prior knowledge of the c.d. contribution of the penultimate amino-acid residue of the 'unknown' peptide. This novel *N*-terminal analysis method is therefore of limited value; however, *N*-terminal amino-acid residues yielding *N*-thiobenzoyl derivatives with unmistakable c.d. behaviour, *e.g.* L-proline, might be identified in this way.

The c.d. data can be interpreted further to illustrate conformational mobility in terminal *N*-thiobenzoyl peptides, though little progress towards a rigorous quantitative assessment (ultimately requiring conversion of the c.d. data into rotatory strengths [17,27]) seems possible within the relatively limited variety of structure represented in the present study. Substitution of L-leucine for the glycine residue in *N*-thiobenzoyl-L-valylglycine shifts $\Delta\varepsilon$ values considerably towards more positive figures, although the data for *N*-thiobenzoylglycyl-L-leucine indicate that, aqueous solutions excepted, only a rather small positive c.d. contribution might be expected from an L-leucyl residue adjacent to the *N*-terminus in this series. Thus, summation of the c.d. contributions of separate asymmetric centres provides an unsatisfactory basis for the prediction of the c.d. of a representative *N*-thiobenzoyl dipeptide, and a general failure of this simple additive relationship is apparent from similar treatment of the c.d. data for the derivatives of L-valyl-D-leucine, L-valyl-L-phenylalanine, L-leucyl-L-alanine, and L-seryl-L-prolylglycine. Realistically, flexible molecules such as these are expected to exist as equilibrium mixtures of conformers in solution, the positions of conformational equilibria in a given solvent being determined for each derivative by the steric requirements of side-chains [R^1, R^2, *etc.* in (II)]. The differences between observed c.d. behaviour and

[26] J. Beacham, J. M. Halcrow, G. W. Kenner, N. H. Rogers, and R. C. Sheppard, Chemical Society Anniversary Meeting, Exeter, 1967, Abstracts, A.12, and references cited therein.
[27] S. F. Mason, *Quart. Rev.*, 1963, **17**, 20.

that calculated by summation of contributions from separate amino-acid residues for each derivative generally reflect the relative distortions of the fully extended conformation (II) which might be expected qualitatively on the basis of structural requirements in this series; for example, the large influence on conformation of a non-terminal L-prolyl residue [28] is demonstrated. The c.d. data for solutions in the less polar solvents show the greatest divergences from an additive relationship, a fact which is consistent with the influence of solvent polarity in moderating non-bonded interactions and consequently in determining conformer populations.

[8/1666 *Received, November 14th*, 1968]

[28] J. T. Edsall, *J. Polymer Sci.*, 1954, **12**, 253.