D. Phillips, *Biochem. Biophys. Res. Commun.*, **58**, 974 (1974).
(52) I. M. Kolthoff, E. J. Meehan, M. S. Tsao, and Q. W. Choi, *J. Phys. Chem.*, **66**, 1233 (1962); E. J. Meehan, I. M. Kolthoff, and H. Kakiuchi, *ibid.*, **66**, 1238 (1962); S. S. Panpalia, R. N. Mehrotra, and R. C. Kapoor, *Indian J. Chem.*, **12**, 1166 (1974), and references cited therein. Note, however, that in these (and other) cases ferricyanide oxidation of thiols may have been catalyzed by traces of Cu(II): G. J. Bridgart, M. W. Fuller, and I. R. Wilson, *J. Chem. Soc., Dalton Trans.*, 1274 (1973); G. J. Bridgart and I. R. Wilson, *ibid.*, 1281 (1973).
(53) R. S. Murray, *J. Chem. Soc., Dalton Trans.*, 2381 (1974).
(54) Observations concerning the oxidation of $[Fe_4S_4(SPh)_4]^{3-}$ in the absence of ferricyanide are not to be interpreted as demonstrating that Fd_{red} sites are first extruded as the analogue trianions, which are then largely oxidized. While this reaction sequence is possible, these experiments do not define the point at which the $Fe_4S_4^*$ core of Fd_{red} achieves the core oxidation level of $[Fe_4S_4(SPh)_4]^{2-}$. However, it is quite clear that if this trianion is liberated as such in the extrusion system, full oxidation can be achieved with ferricyanide. Although anaerobic techniques were used throughout, we cannot eliminate the possibility that in dilute protein solutions traces of dioxygen alone, PhSSPh, or HMPA peroxide formed in the presence of light[28b] may contribute to the oxidation of $[Fe_4S_4(SPh)_4]^{3-}$ prior to the addition of ferricyanide.
(55) The points raised here are readily appreciated on examination of the spectra of various Fe–S protein chromophores assembled by Hall et al.[5b]
(56) Owing to limited amounts of H_2ase, no detailed attempts have been made to examine certain potential variables of the extrusion reaction, including the minimal amounts of ferricyanide required to produce the final A_{458} values obtained by addition of the quantities stated in Table III.
(57) We have been unable to bind ligands simulating other types of peptide side-chain coordination to $Fe_4S_4^*$ cores. Consequently, it has not been established whether groups other than thiolate undergo ligand substitution with thiols.
(58) If, e.g., the true n values for solution A were 2.0 or 4.0, the stated protein concentration would have to be in error by ca. ±40%. The close agreement between initial calculated and independently determined final Fe molar concentrations makes an error of this size improbable.
(59) J. Cambray, R. W. Lane, A. G. Wedd, and R. H. Holm, unpublished observations.
(60) (a) Extents of reduction in aqueous solution were calculated from published spectral data: S. G. Mayhew, D. Petering, G. Palmer, and G. P. Foust, *J. Biol. Chem.*, **244**, 2830 (1969). These same data were used to estimate reduction extents in 80% HMPA solutions. (b) Slight precipitate formation was usually encountered in 80% solutions when ferricyanide was added.
(61) References pertinent to formation and kinetic stabilities of Fe(III)/thiol complexes are collected elsewhere;[2b] cf., e.g., C. M. Bell, E. D. McKenzie, and J. Orton, *Inorg. Chim. Acta*, **5**, 109 (1971); L. G. Stadtherr and R. B. Martin, *Inorg. Chem.*, **11**, 92 (1972); A. Tomita, H. Hirai, and S. Makishima, *ibid.*, 7, 760 (1968). None of these studies involved benzenethiol. Judging from the comparative spectra of Rd_{ox}/Rd_{red} and their analogues $[Fe(S_2\text{-}o\text{-}xyl)_2]^{-,2-}$,[3] any reaction product of an Fe(II) contaminant and benzenethiol is less likely to have strong charge-transfer bands in the visible region. The only such complex characterized is red-brown tetrahedral $[Fe(SPh)_4]^{2-}$, whose visible/UV spectrum has not been reported: D. G. Holah and D. Coucouvanis, *J. Am. Chem. Soc.*, **97**, 6917 (1975). Its Fe(III) analogue has not been isolated.
(62) M. Tanaka, M. Haniu, K. T. Yasunobu, and L. E. Mortenson in "Iron and Copper Proteins", K. T. Yasunobu, H. F. Mower, and O. Hayaishi, Ed., Plenum Press, New York, N.Y., 1976, pp 83–90.
(63) K. T. Yasunobu and M. Tanaka in "Iron–Sulfur Proteins", Vol. II, W. Lovenberg, Ed., Academic Press, New York, N.Y., 1973, Chapter 2.
(64) M. J. Murphy, L. M. Siegel, S. R. Tove, and H. Kamin, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 612 (1974); P. J. Aparicio, D. B. Knaff, and R. Malkin, *Arch. Biochem. Biophys.*, **169**, 102 (1975).
(65) T. Ohnishi, J. C. Salerno, D. B. Winter, J. Lim, C. A. Yu, and T. E. King, *J. Biol. Chem.*, **251**, 2094 (1976); T. Ohnishi, J. Lim, D. B. Winter, and T. E. King, *ibid.*, **251**, 2105 (1976).
(66) N. A. Stombaugh, J. E. Sundquist, R. H. Burris, and W. H. Orme-Johnson, *Biochemistry*, **15**, 2633 (1976).
(67) G. Nakos and L. E. Mortenson, *Biochemistry*, **10**, 2442 (1971); *Biochim. Biophys. Acta*, **227**, 576 (1971).
(68) P. H. Gitlitz and A. I. Krasna, *Biochemistry*, **14**, 2561 (1975).
(69) R. H. Haschke and L. L. Campbell, *J. Bacteriol.*, **105**, 249 (1971).
(70) J. Legall, D. V. Dervartanian, E. Spilker, J. P. Lee, and H. D. Peck, *Biochim. Biophys. Acta*, **234**, 525 (1971).
(71) T. Yagi, *J. Biochem. (Tokyo)*, **68**, 649 (1970).
(72) T. Yagi, K. Kimura, H. Daidoji, F. Sakai, S. Tamura, and H. Inokuchi, *J. Biochem. (Tokyo)*, **79**, 661 (1976).

Alkylation of a Tripeptide by a Carcinogen: The Crystal Structures of Sarcosylglycylglycine, 9-Methyl-10-chloromethylanthracene, and Their Reaction Product

Jenny Pickworth Glusker,* H. L. Carrell, Helen M. Berman, Barbara Gallen, and Richard M. Peck

Contribution from the Institute for Cancer Research, The Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111. *Received May 17, 1976*

Abstract: The effect of alkylation of a peptide by a polycyclic aromatic hydrocarbon has been investigated as a model for protein–carcinogen interactions. The tripeptide sarcosylglycylglycine (I) reacts with 9-methyl-10-chloromethylanthracene (II) with the elimination of HCl to form the alkylated tripeptide (III). The crystal data are (I) a = 20.847 (3), b = 10.252 (1), and c = 8.719 (1) Å, space group *Pbca*, Z = 8; (II) a = 14.379 (1), b = 16.359 (1), c = 5.0993 (4) Å, space group $P2_12_12_1$, Z = 4; and (III) a = 19.022 (4), b = 10.727 (2), c = 10.728 (2) Å, β = 94.24 (2)°, space group $P2_1/c$, Z = 4. The crystal structures were determined by direct methods using MULTAN and refined by least-squares techniques. The tripeptide and the alkylated tripeptide both occur as zwitterions. In the crystalline alkylated peptide, III, the hydrophobic polycyclic groups stack together, in a manner similar to that found in the simple alkylating agent, II. The structure of III consists of layers of aromatic side chains, peptide residues, and water of crystallization. In addition, one of the peptide groups in the alkylated tripeptide is nonplanar (torsion angle 159°), possibly a result of packing forces in the crystal. The simple peptide I does not crystallize with water of crystallization. The crystal structure involves extensive hydrogen bonding, although a pleated sheet structure is not formed. The zwitterion folds so that part of the structure of the peptide I resembles the conformation found in an α helix. This helicity is lost on alkylation.

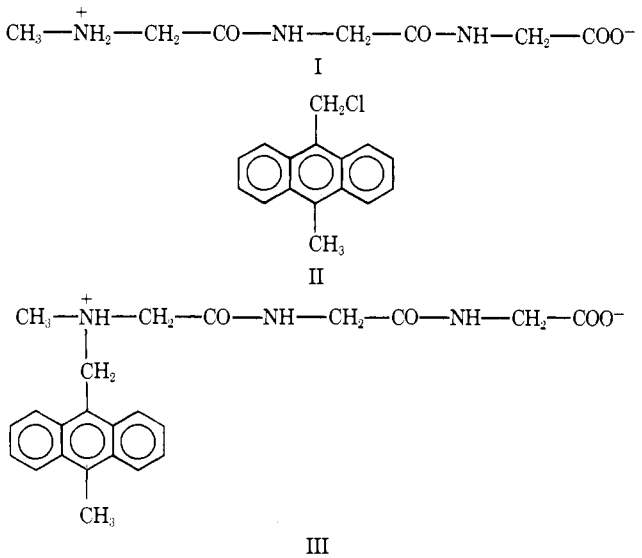
When polycyclic aromatic hydrocarbons interact with living matter they are generally metabolized, by enzyme systems, to compounds containing hydroxy or keto groups. These metabolites, now more soluble in water than the parent hydrocarbon, generally are then excreted. However some hydrocarbon molecules are found to become covalently bound to proteins and to nucleic acids, possibly through the interaction of these macromolecules with the metabolites of the polycyclic aromatic hydrocarbons.[1-5] Carcinogenesis may result from the interaction of certain (but not all) polycyclic aromatic hydrocarbons with living cells, but it is not yet clear whether the critical target is a protein or a nucleic acid.[6,7]

In order to investigate the molecular basis of chemical carcinogenesis by polycyclic aromatic hydrocarbons we are

studying, by x-ray crystallographic techniques, the molecular structures of covalent compounds formed between hydrocarbons and portions of biological macromolecules. A comparison with the structures of the individual components is then made, and the effect of alkylation can then be assessed.

We report here the crystal structure of a tripeptide, sarcosylglycylglycine, of an alkylating derivative of a polycyclic aromatic hydrocarbon that is carcinogenic,[8] and of the product of alkylation of the tripeptide by the hydrocarbon.[9] This represents a model for the interaction of a carcinogenic polycyclic aromatic hydrocarbon (or its metabolite) with a protein. A major interest of this study was to find the effect of the large hydrophobic group on the conformation of the tripeptide.

The formulas of the three compounds studied (I, II, and III, referred to as SGG, MACl, and MASGG, respectively) are as shown.



I

II

III

Experimental Section

Sarcosylglycylglycine was obtained from Nutritional Biochemicals Corporation, Cleveland, Ohio. The preparations of 9-methyl-10-chloromethylanthracene[10] and the alkylated peptide[9] have already been described.

Crystal Data. (i) **Tripeptide.** Sarcosylglycylglycine (SGG) (I) $C_7H_{13}N_3O_4$, orthorhombic, space group *Pbca*, a = 20.847 (3), b = 10.252 (1), c = 8.719 (1) Å, V = 1863.4 (4) Å^3, FW = 203.20, Z = 8, F (000) = 864, D_{calcd} = 1.45 g cm^{-3}, D_{obsd} = 1.45 g cm^{-3}. Crystals of I were grown from water and 2-methyl-2,4-pentanediol.

(ii) **Alkylating Agent.** 9-Methyl-10-chloromethylanthracene (MACl) (II) (ICR-433), $C_{16}H_{13}Cl$, orthorhombic, space group $P2_12_12_1$, a = 14.379 (1), b = 16.359 (1), c = 5.0993 (4) Å, V = 1199.5 (2) Å^3, FW = 240.73, Z = 4, F (000) = 504, D_{calcd} = 1.33 g cm^{-3}, D_{obsd} = 1.33 g cm^{-3} (KI solution). Crystals of II were grown from ethanol.

(iii) **Alkylated Peptide.** 9-Methyl-10-methylanthrylsarcosylglycylglycine (MASGG) (III) $C_{23}H_{25}N_3O_4 \cdot 2H_2O$, monoclinic, space group $P2_1/c$, a = 19.022 (4), b = 10.727 (2), c = 10.728 (2) Å, β = 94.24 (2)°, V = 2183.0 (8) Å^3, FW = 443.50, Z = 4, F (000) = 944, D_{calcd} = 1.35 g cm^{-3}, D_{obsd} = 1.34 g cm^{-3}. Crystals of III were grown from a mixture of 3:1 methanol and water.

Data Collection. Three-dimensional x-ray diffraction data were obtained with Cu Kα monochromatized radiation (λ 1.5418 Å) using an automated four circle diffractometer (Syntex $P\bar{1}$). The θ–2θ scan technique was used to a maximum value of sin θ/λ of 0.606 Å^{-1}. The threshold for observation was considered to be $2\sigma(I) = I_{obsd}$ for SGG, $2\sigma(I)$ for MACl, and $2.33\sigma(I)$ for MASGG, where $\sigma(I)$ was derived from counting statistics. Values of $\sigma(F)$ were determined as $\sigma(F) = (F/2)\{[\sigma^2(I)/I^2] + \delta^2\}^{1/2}$ where δ is a measured instrumental uncertainty (0.015 for SGG, 0.019 for MACl, and 0.015 for MASGG. The fall-off in intensities of the standard reflections was corrected for, with a maximum value of 1.112 for MASGG. There was no intensity fall-off for MACl or for SGG. The crystal sizes were 0.1 × 0.1 × 0.2 mm for SGG, 0.10 × 0.05 × 0.25 mm for MACl, and 0.1 × 0.25 × 0.25 mm for MASGG. The data were corrected for Lorentz and polarization factors. No absorption corrections were applied. The data were then placed on an absolute scale by a Wilson plot. For SGG, 1746 independent data were measured of which 854 were below threshold values. For MACl, 1331 independent data were measured of which 405 were below threshold values. For MASGG, 4186 independent data were measured for which 1384 were below threshold values.

Structure Refinement. The structures of all three compounds were determined by the program MULTAN.[11]

(I) SGG. All 14 atoms were located in the map calculated from phases from MULTAN (R = 0.213). Isotropic refinement gave R = 0.12, and a difference map was calculated which showed peaks at the positions of all hydrogen atoms. All atoms were then refined by a full-matrix least-squares procedure, heavier atoms anisotropic, hydrogen atoms isotropic to R = 0.061 (R_w = 0.049), after an extinction correction ($\alpha = 1.2 \times 10^{-6}$)[12] had been applied to the data.

(II) MACl. The structure, determined from MULTAN and checked against the Patterson map, refined anisotropically only to R = 0.195. Data on a second crystal refined to R = 0.135. It was thought that the space group might be in error, and therefore the hypothesis that the space group was noncentric was verified by a physical test which involves the use of a second harmonic analyzer. The experiment was performed by Drs. J. P. Dougherty and S. K. Kurtz of Philips Laboratories, Briarcliff Manor, N.Y.,[13] who demonstrated a strong optical second harmonic generation in a sample of MACl. However, on refinement with partial disorder between a chlorine and a hydrogen atom, i.e., disorder between the methyl and chloromethyl groups, the structure refined anisotropically (hydrogen atoms isotropic) to R = 0.046 (R_w = 0.042). This probably represents a situation in which molecules are flipped over in domains in the crystal since a random

Table I. Final Parameters for SGG[a]

	x	y	z	Av B
C(1)	−0.2610 (2)	−0.1885 (5)	0.1475 (7)	3.2 (2)
N(2)	−0.2230 (2)	−0.0667 (4)	0.1594 (4)	2.0 (2)
C(3)	−0.1683 (2)	−0.0787 (5)	0.2661 (5)	2.1 (2)
C(4)	−0.1300 (2)	0.0473 (5)	0.2628 (5)	2.0 (2)
O(5)	−0.1318 (2)	0.1200 (3)	0.1511 (4)	2.5 (1)
N(6)	−0.0932 (2)	0.0687 (4)	0.3862 (4)	2.2 (2)
C(7)	−0.0565 (2)	0.1879 (5)	0.3952 (6)	2.6 (2)
C(8)	0.0009 (2)	0.1974 (6)	0.2891 (5)	2.0 (2)
O(9)	0.0253 (2)	0.3048 (3)	0.2653 (4)	2.8 (1)
N(10)	0.0227 (2)	0.0868 (4)	0.2288 (4)	2.0 (2)
C(11)	0.0799 (2)	0.0841 (5)	0.1330 (6)	2.2 (2)
C(12)	0.1425 (2)	0.0733 (5)	0.2222 (6)	2.0 (2)
O(13)	0.1930 (1)	0.0659 (3)	0.1445 (4)	2.8 (2)
O(14)	0.1418 (2)	0.0733 (3)	0.3654 (3)	3.2 (2)
	x	y	z	B
H(11)	−0.279 (3)	−0.209 (6)	0.245 (6)	5 (2)
H(12)	−0.227 (3)	−0.260 (6)	0.109 (6)	5 (2)
H(13)	−0.300 (3)	−0.177 (5)	0.078 (6)	2 (2)
H(21)	−0.207 (3)	−0.050 (6)	0.061 (7)	6 (2)
H(22)	−0.249 (4)	−0.011 (7)	0.196 (8)	7 (2)
H(31)	−0.186 (2)	−0.096 (4)	0.376 (5)	3 (1)
H(32)	−0.143 (2)	−0.150 (4)	0.231 (5)	1 (1)
H(6)	−0.097 (3)	0.023 (6)	0.468 (7)	4 (2)
H(71)	−0.083 (2)	0.261 (5)	0.370 (6)	3 (1)
H(72)	−0.040 (2)	0.198 (5)	0.496 (6)	3 (1)
H(10)	0.006 (3)	0.013 (5)	0.243 (6)	3 (1)
H(111)	0.076 (2)	0.014 (5)	0.065 (5)	2 (1)
H(112)	0.079 (2)	0.170 (5)	0.068 (5)	3 (1)

[a] Positional parameters are given as fractions of cell edges. Anisotropic factors (av B) are expressed as the average of B_{11}, B_{22}, and B_{33}. Values of B_{ij} are listed in the microfilm edition of this journal. Isotropic temperature factors as exp($-B \sin^2\theta/\lambda^2$) with all B and av B values given in Å^2. The standard deviations for each parameter, determined from the inverted full matrix, are given in parentheses and apply to the last specified digits.

Table II. Final Parameters for MACl[a]

	x	y	z	Av B
Cl(17)	0.6124 (1)	0.2074 (1)	0.2710 (4)	6.1 (1)
Cl(18)	0.3874 (4)	−0.1999 (3)	−0.2449 (15)	8.5 (3)
C(1)	0.3917 (3)	0.1508 (2)	0.0290 (9)	4.9 (2)
C(2)	0.3194 (3)	0.1624 (3)	−0.1381 (10)	5.6 (2)
C(3)	0.2945 (3)	0.1008 (3)	−0.3137 (11)	6.1 (2)
C(4)	0.3412 (3)	0.0295 (3)	−0.3200 (9)	5.4 (2)
C(5)	0.6018 (3)	−0.1437 (2)	−0.0095 (9)	5.2 (2)
C(6)	0.6760 (3)	−0.1535 (3)	0.1527 (12)	6.1 (2)
C(7)	0.7019 (3)	−0.0930 (3)	0.3300 (13)	6.3 (2)
C(8)	0.6521 (3)	−0.0222 (3)	0.3458 (11)	5.4 (2)
C(9)	0.5208 (2)	0.0647 (2)	0.1999 (8)	4.3 (2)
C(10)	0.4703 (3)	−0.0598 (2)	−0.1658 (9)	4.4 (2)
C(11)	0.4454 (3)	0.0761 (2)	0.0286 (8)	3.9 (2)
C(12)	0.4193 (3)	0.0140 (2)	−0.1537 (8)	4.1 (2)
C(13)	0.5726 (3)	−0.0082 (2)	0.1849 (10)	4.1 (2)
C(14)	0.5473 (3)	−0.0707 (3)	0.0004 (9)	4.2 (2)
C(15)	0.4424 (4)	−0.1263 (3)	−0.3508 (11)	6.6 (3)
C(16)	0.5439 (3)	0.1281 (3)	0.4021 (11)	5.4 (2)
Occupancies: Cl(17) 0.737(3); Cl(18) 0.241(3)				
	x	y	z	B
H(1)	0.414 (2)	0.195 (2)	0.147 (8)	5 (1)
H(2)	0.281 (3)	0.214 (3)	−0.106 (10)	7 (1)
H(3)	0.240 (4)	0.112 (3)	−0.403 (12)	7 (2)
H(4)	0.325 (3)	−0.012 (2)	−0.428 (10)	5 (1)
H(5)	0.581 (3)	−0.187 (3)	−0.165 (11)	9 (2)
H(6)	0.713 (3)	−0.205 (3)	0.157 (9)	5 (1)
H(7)	0.760 (3)	−0.097 (2)	0.463 (11)	6 (1)
H(8)	0.672 (3)	0.017 (3)	0.503 (11)	5 (2)
H(15)	0.481 (4)	−0.158 (3)	−0.397 (12)	8 (2)
H(15′)	0.410 (4)	−0.109 (3)	−0.501 (12)	8 (2)
H(16)	0.485 (3)	0.160 (3)	0.447 (10)	6 (1)
H(16′)	0.584 (4)	0.112 (3)	0.554 (13)	10 (2)

[a] See footnote *a*, Table I.

packing of methyl and chloromethyl groups in the crystal would not fit (the chlorine atoms would be too close to each other). Attempts to dissect out a crystal without such disorder were unsuccessful.

(III) MASGG. 29 atoms from the map calculated from phases from MULTAN (R = 0.37) were input to a Fourier map from which the remaining three atoms were located (R = 0.20). Isotropic and then anisotropic full-matrix least-squares refinement gave R = 0.12 at which stage all hydrogen atoms were located from a difference map. All atoms were then refined, carbon, nitrogen, and oxygen atoms anisotropically and hydrogen atoms isotropically for all the data. The R values were 0.060 for observed data (R_w = 0.078) and R = 0.090 for all data. This refinement was done in segments because of the large number of parameters being varied.

The X-ray 72 System of Programs[14] and the Brookhaven National Laboratory System of Programs[15] were used to calculate maps, and a full-matrix least-squares method was used for the refinement.[16] The weights, w, of the reflections during refinement were $1/(\sigma^2(F))$ with zero weight for those reflections below the threshold value except for MASGG. In this case a weighting scheme was derived from an analysis of the agreement between calculated and observed data. The quantity minimized in the least-squares calculations was $\Sigma w\{||F_o| - |F_c||\}^2$. The atomic scattering factors for chlorine, oxygen, and carbon are those of Cromer and Mann[17] and for hydrogen those of Stewart, Davidson, and Simpson.[18] The real component ($\Delta f'$) of the anomalous corrections for the chlorine atom (+0.348) is listed by Cromer and Liberman.[19]

Final positional and average thermal parameters are listed in Table I for SGG, Table II for MACl, and Table III for MASGG. Individual thermal parameters are listed in the microfilm edition of this journal (Table 1). A table of observed and calculated structure factors may be obtained from the authors on request.

Results and Discussion

The interatomic distances and interbond angles found for SGG and MASGG are shown in Figures 1 and 2, respectively. Values for MACl, less accurate because of disorder, are listed in the microfilm edition (Table 2). From the location of hydrogen atoms, it can be seen that both SGG and MASGG exist

Table III. Final Parameters for MASGG[a]

	x	y	z	Av B
C(1)	0.6754 (2)	−0.0990 (3)	0.1607 (3)	3.5 (1)
C(2)	0.6447 (2)	−0.2119 (3)	0.1348 (4)	4.5 (2)
C(3)	0.5954 (2)	−0.2257 (3)	0.0310 (4)	4.9 (2)
C(4)	0.5784 (2)	−0.1266 (4)	−0.0446 (4)	4.4 (2)
C(5)	0.5862 (2)	0.3286 (4)	−0.1205 (3)	3.9 (2)
C(6)	0.6097 (2)	0.4426 (4)	−0.0870 (3)	4.2 (2)
C(7)	0.6631 (2)	0.4573 (3)	0.0106 (3)	4.0 (2)
C(8)	0.6887 (2)	0.3554 (3)	0.0747 (3)	3.3 (1)
C(9)	0.6874 (1)	0.1267 (3)	0.1151 (2)	2.7 (1)
C(10)	0.5872 (2)	0.0989 (3)	−0.0945 (3)	3.6 (1)
C(11)	0.6579 (1)	0.0082 (3)	0.0867 (3)	2.9 (1)
C(12)	0.6072 (2)	−0.0054 (3)	−0.0202 (3)	3.4 (1)
C(13)	0.6116 (2)	0.2179 (3)	−0.0584 (3)	3.2 (1)
C(14)	0.6638 (1)	0.2323 (3)	0.0464 (3)	2.8 (1)
C(15)	0.7439 (1)	0.1399 (3)	0.2214 (2)	2.5 (1)
N(16)	0.8158 (1)	0.0971 (2)	0.1862 (2)	2.1 (1)
C(17)	0.8389 (2)	0.1643 (3)	0.0730 (3)	2.9 (1)
C(18)	0.8700 (1)	0.1118 (3)	0.2925 (2)	2.7 (1)
C(19)	0.8582 (1)	0.0151 (3)	0.3926 (2)	2.6 (1)
O(20)	0.8172 (1)	−0.0716 (2)	0.3713 (2)	3.7 (1)
N(21)	0.8971 (1)	0.0296 (2)	0.4996 (2)	2.7 (1)
C(22)	0.8978 (2)	−0.0660 (3)	0.5955 (3)	3.1 (1)
C(23)	0.8478 (1)	−0.0411 (3)	0.6962 (2)	2.5 (1)
O(24)	0.8374 (1)	0.0652 (2)	0.7351 (2)	3.4 (1)
N(25)	0.8196 (1)	−0.1425 (2)	0.7439 (2)	3.0 (1)
C(26)	0.7908 (1)	−0.1386 (3)	0.8648 (3)	3.0 (1)
C(27)	0.8475 (1)	−0.1273 (2)	0.9726 (2)	2.6 (1)
O(28)	0.9105 (1)	−0.1170 (2)	0.9485 (2)	3.5 (1)
O(29)	0.8261 (1)	−0.1284 (2)	1.0804 (2)	3.6 (1)
C(30)	0.5384 (2)	0.0820 (4)	−0.2104 (3)	5.1 (2)
O(31)	0.9935 (1)	0.3001 (2)	0.1191 (2)	5.4 (1)
O(32)	1.0034 (2)	0.4308 (4)	0.3402 (3)	9.1 (2)
	x	y	z	B
H(1)	0.709 (2)	−0.088 (3)	0.233 (3)	2 (1)
H(2)	0.654 (2)	−0.288 (3)	0.196 (3)	5 (1)
H(3)	0.574 (2)	−0.309 (3)	0.011 (3)	4 (1)
H(4)	0.546 (2)	−0.136 (3)	−0.130 (3)	6 (1)
H(5)	0.547 (2)	0.306 (3)	−0.188 (3)	3 (1)
H(6)	0.591 (2)	0.521 (3)	−0.132 (3)	3 (1)
H(7)	0.677 (2)	0.538 (3)	0.038 (3)	4 (1)
H(8)	0.728 (2)	0.374 (3)	0.146 (3)	2 (1)
H(151)	0.745 (2)	0.231 (3)	0.241 (3)	1 (1)
H(152)	0.732 (2)	0.089 (3)	0.299 (3)	1 (1)
H(16)	0.814 (2)	0.009 (3)	0.162 (3)	1 (1)
H(171)	0.840 (2)	0.250 (3)	0.088 (3)	2 (1)
H(172)	0.889 (2)	0.137 (3)	0.054 (3)	2 (1)
H(173)	0.802 (2)	0.145 (3)	0.000 (3)	3 (1)
H(181)	0.921 (2)	0.097 (3)	0.259 (3)	1 (1)
H(182)	0.871 (2)	0.196 (3)	0.323 (3)	1 (1)
H(21)	0.930 (2)	0.091 (3)	0.512 (3)	3 (1)
H(221)	0.946 (2)	−0.068 (3)	0.640 (3)	2 (1)
H(222)	0.886 (2)	−0.150 (3)	0.553 (3)	3 (1)
H(25)	0.828 (2)	−0.217 (3)	0.710 (3)	3 (1)
H(261)	0.758 (2)	−0.062 (3)	0.875 (3)	2 (1)
H(262)	0.763 (2)	−0.216 (3)	0.880 (3)	2 (1)
H(301)	0.545 (2)	0.158 (3)	−0.281 (3)	6 (1)
H(302)	0.553 (2)	0.011 (3)	−0.271 (3)	6 (1)
H(303)	0.489 (2)	0.084 (3)	−0.208 (3)	6 (1)
H(311)	1.026 (2)	0.243 (3)	0.110 (3)	5 (1)
H(312)	1.007 (2)	0.340 (3)	0.185 (3)	6 (1)
H(321)	0.962 (2)	0.480 (3)	0.381 (3)	6 (1)
H(322)	1.033 (2)	0.398 (3)	0.414 (3)	6 (1)

[a] See footnote *a*, Table I.

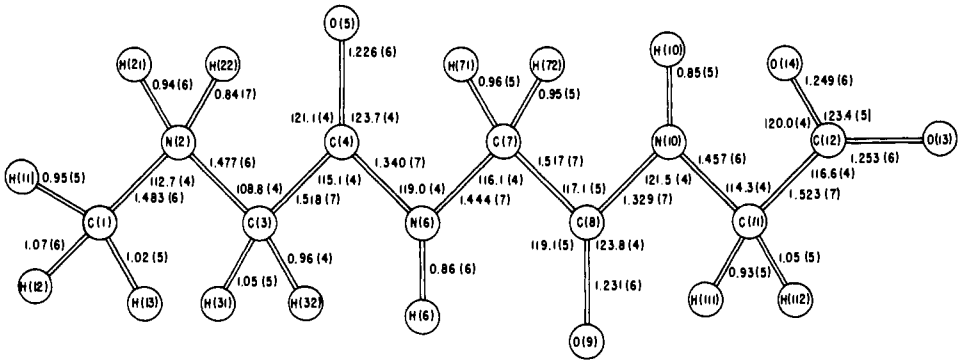


Figure 1. Interatomic distances and interbond angles for SGG. Estimated standard deviations from the least-squares refinement are given in parentheses with respect to the last digit quoted for any given dimension.

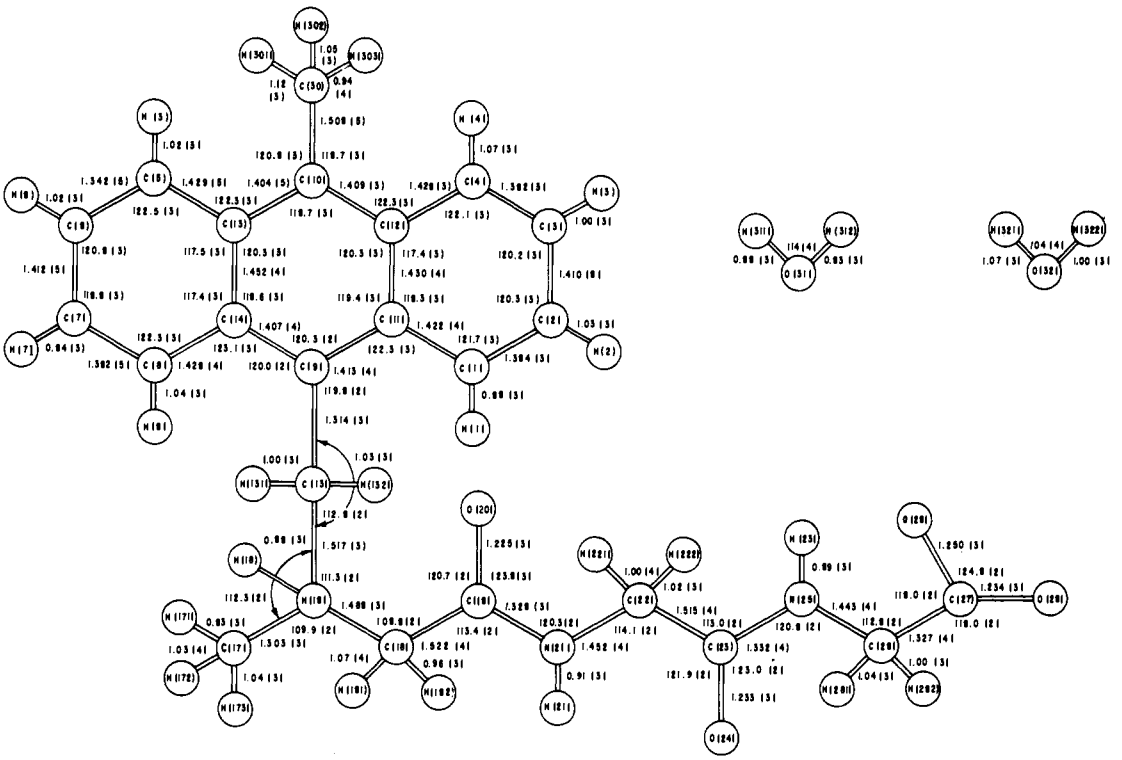


Figure 2. Interatomic distances and interbond angles for MASGG. Estimated standard deviations from the least-squares refinement are given in parentheses with respect to the last digit quoted for any given dimension.

as zwitterions, the C-terminal carboxyl groups being ionized and the N-terminal nitrogen atoms bearing positive charges. Since MACl is disordered in the approximate ratio 3:1, the molecular dimensions are less accurate (esd values from the least-squares refinement are probably underestimated). There are no significant changes in dimensions on alkylation of the tripeptide compared with both the simple peptide and the alkylating agent except at the point of alkylation. The bond lengths and angles in SGG and MASGG all approximate those listed by Marsh and Donohue.[20]

The conformations of the two tripeptides differ, as shown in Figure 3, and, surprisingly, one peptide group in MASGG (III) is nonplanar (torsion angle 159°), as shown in Table IV. It is of particular interest that this nonplanar peptide conformation occurs in the peptide group farthest from the site of alkylation. No reason for this nonplanarity is immediately obvious. It does not seem to be due to intramolecular repulsions but may possibly be a result of packing forces (see later). The major difference in conformation (for SGG on alkylation) occurs at the C(7)-C(8) bond. It should be noted that, since each peptide occurs as a racemate in the crystals studied, the signs of all torsion angles (in Table IV) are changed for the mirror image of each molecule. In Table IV a comparison with selected torsion angles for glycylphenylalanylglycine (GPG)[21] is given. In SGG there is a bend in the molecule so that N(6) lies only 2.78 Å from N(10) in a cis configuration. In MASGG this N-C-C-N conformation is trans and N(21) ··· O(24) is 2.86 Å. For a fully extended peptide chain torsion angles would be $\phi = \psi = \omega = +180°$.[22] For a right-hand α helix the values are $\phi = -57°, \psi = -47°, \omega = +180°$.[23] The torsion angles for MASGG and SGG have been drawn as a Ramachandran plot[24] (in the microfilm edition, Figure I). This map compares ϕ and ψ values for α helices,[23] pleated sheet structures,[25,26] and polyglycine II.[26] Values for MASGG lie near those for pleated sheets and for polyglycine, but for SGG, values (the crystal is a racemate) lie near those for α helices. These points on the plot correspond to the torsion angles involving the atom C(7) (see Figure 3 and Table IV), about which the two peptide groups hinge. The conformation in this area (C(22)) of MASGG is different. Thus the partial helical character of SGG has been

Table IV. Torsion Angles for Tripeptides

		MSAGG					SGG					GPG
C-N-C-C	ϕ_1	17	16	18	19	−163.3	1	2	3	4	−176.4	
	ϕ_2	19	21	22	23	−96.7	4	6	7	8	72.1	−126.3
	ϕ_3	23	25	26	27	−71.1	8	10	11	12	−85.4	84.3
N-C-C-O		16	18	19	20	11.4	2	3	4	5	24.1	
		21	22	23	24	−36.7	6	7	8	9	−163.7	
		25	26	27	28	3.1	10	11	12	14	3.1	
		25	26	27	29	−177.3	10	11	12	13	−177.5	
N-C-C-N	ψ_1	16	18	19	21	−171.0	2	3	4	6	−158.5	127.6
	ψ_2	21	22	23	25	147.0	6	7	8	10	17.4	132.0
O-C-N-C		20	19	21	22	6.0	5	4	6	7	−4.4	
		24	23	25	26	−17.0	9	8	10	11	−2.6	
C-C-N-C	ω_1	18	19	21	22	−171.4	3	4	6	7	178.3	−175.4
	ω_2	22	23	25	26	159.3[b]	7	8	10	11	176.2	179.7

[a] Note: Each peptide exists in the crystalline state as a racemate. Therefore the signs of torsion angles may be reversed. MASGG, SGG: this work; GPG = glycylphenylalanylglycine: ref 21. [b] Nonplanar peptide bond group.

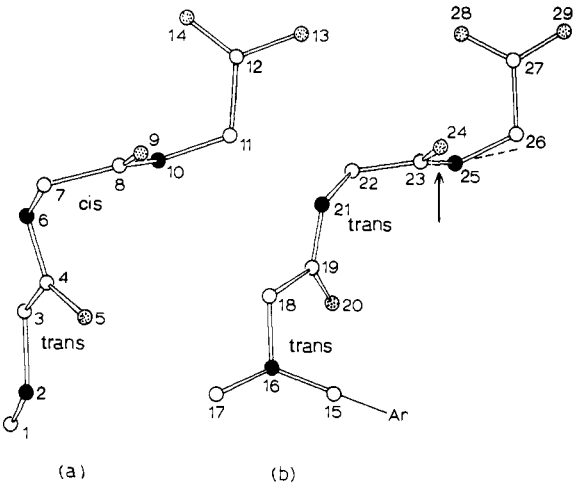


Figure 3. Views of the two tripeptides showing differences in conformation: (a) SGG, (b) MASGG. The nonplanar peptide group in MASGG is indicated by an arrow, as is the large aromatic ring system. Nitrogen atoms are black, carbon atoms white, and oxygen atoms stippled.

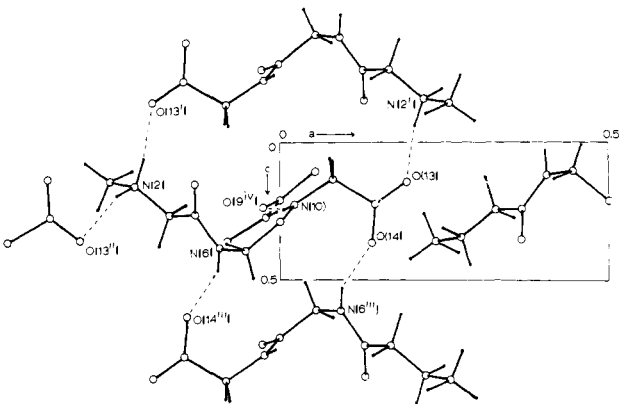


Figure 4. Packing in SGG. *h0l* projection. For clarity only a few molecules are shown (and only part of some). Hydrogen bonds are shown by broken lines. Superscripts are defined in Table V.

altered on alkylation to give an extended polypeptide chain.

The ring system is not precisely planar but shaped like a butterfly with an angle of 8.5° between the outer rings for MASGG but only 2.2° for MACl. Details are given in the microfilm edition (Table 3).

The hydrogen bond systems for SGG and MASGG are described in Tables V and VI. There is no water of crystallization present in SGG and the packing (Figure 4) may be described as a network of hydrogen bonds in all directions. However, MASGG, with a bulky hydrophobic substituent, has two molecules of water of crystallization per peptide molecule and, as shown in Figure 5, the packing is much simpler since molecules aggregate to give a layered structure of aromatic groups, peptide chains, and water.

In SGG the terminal carboxyl goup, O(13), and terminal nitrogen atoms, N(2) (i.e., the areas of charge), hydrogen bond across centers of symmetry to give dimers. However, there is no other hydrogen bonding between these two molecules, i.e., a pleated sheet does not form. The hydrogen bond network is extended by similar dimer formation between N(6) and O(14) (of the terminal carboxyl group). This network lies in a plane perpendicular to the *b* axis. There is also hydrogen bonding between N(10) and O(9^{iv}) in a direction parallel to the *b* axis. One carbonyl oxygen atom, O(5), does not take part in hydrogen bonding but the other, O(9), does.

The packing in MASGG is dominated by that of the MA portion of the molecule. This may be part of the cause of the nonplanarity of one peptide group since, if the group were planar, as shown in Figure 5 by dashed lines, the hydrophilic carboxyl group (O(28) and O(29)) would penetrate into the hydrophobic layer. The energy required to make a peptide bond nonplanar is small. From the equation[28] $V(\omega) = (K_\omega/2)(1 - \cos 2\omega)$ (where ω is the torsion angle, K_ω is the barrier to internal rotation about the bond and is approximately 20 kcal mol^{-1}, and $V(\omega)$ is the potential function), the torsion angle, 159°, corresponds to 2.5 kcal per mole.[29,30] This energy loss may be balanced by efficient N-H···O hydrogen bonding, by C-H···O interactions between the anthracene group and carbonyl groups (see later), and by the repulsion of hydrophobic groups as mentioned above.

No distortion to tetrahedral symmetry, which might have been expected as a result of the distortion of a peptide group in MASGG, can be detected around N(25). The sum of the angles around N(25) is 359°. The hydrogen bonded contact O(29^{iii}) lies 0.81 Å from the plane through the atoms C(22), C(23), N(25), and O(24). For comparison O(31^{ii}) lies 0.13 Å from the plane of C(18), C(19), N(21), and O(20) (a planar peptide group). It seems that some torque is imposed on the peptide group in order to satisfy the N(25)-H···O(29^{iii}) hydrogen bond. Since O(29^{iii}) is below the plane of the peptide bond, C(26) is forced above this plane as a hydrogen bond via N(25) is formed. Cross sections of the packing in the water, peptide, and aromatic group layers are given in the microfilm edition of this journal (Figure II).

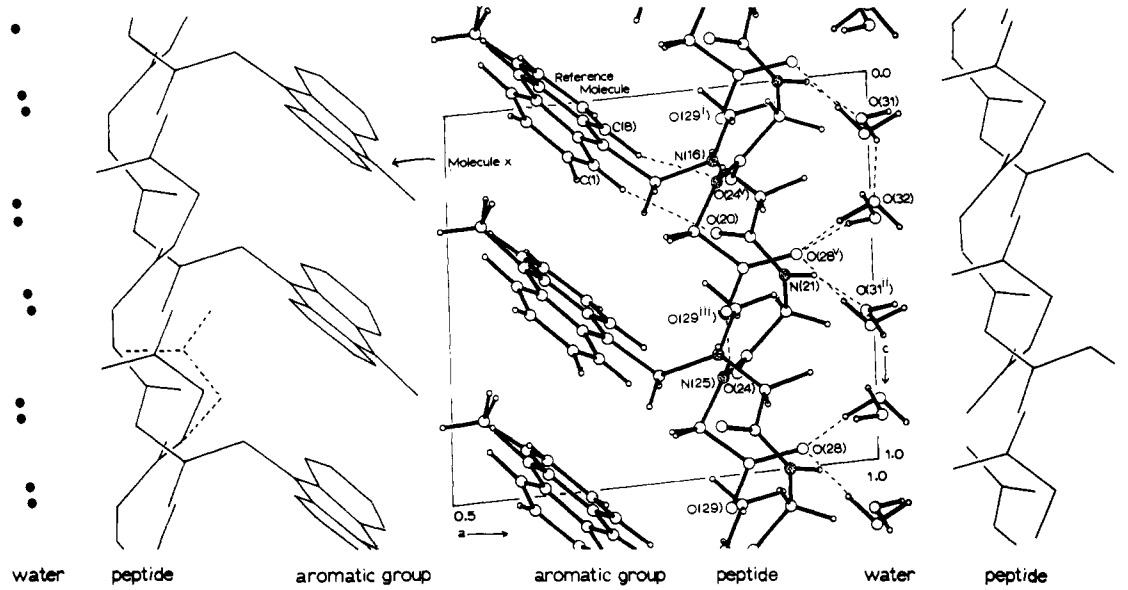


Figure 5. Packing in MASGG. *h0l* projection. This view, partially diagrammatic, shows the segregation into layers of aromatic groups, peptides, and water. The location of the one molecule, if the peptide groups were both planar, is indicated by broken lines on the left hand side of the diagram. Nitrogen atoms are stippled on the right hand side of the diagram. Molecule *x* at $1 - x, -y, -z$.

Table V. Hydrogen Bond Distances and Angles in SGG[a]

D–H- - -A	A- - -D, Å	D- -H, Å	H- - -A, Å	∠D–H- - -A, deg	∠H–D- - -A, deg
N(2)–H(21)- - -O(13^{i})	2.722 (6)	0.94 (6)	1.82 (6)	160 (6)	13 (4)
N(2)–H(22)- - -O(13^{ii})	2.801 (5)	0.84 (7)	2.01 (7)	155 (7)	17 (5)
N(6)–H(6)- - -O(14^{iii})	2.799 (6)	0.86 (6)	1.98 (6)	158 (6)	16 (4)
N(10)–H(10)- - -O(9^{iv})	3.059 (5)	0.85 (5)	2.23 (5)	167 (5)	9 (4)

[a] Code: (i) $-x, -y, -z$; (ii) $-\frac{1}{2} + x, y, \frac{1}{2} - z$; (iii) $-x, -y, 1 - z$; (iv) $-x, -\frac{1}{2} + y, \frac{1}{2} - z$.

Table VI. Distances and Angles in the Hydrogen Bond System in MASGG[a]

D–H- - -A	A- - -D, Å	D- -H, Å	H- - -A, Å	∠D–H- - -A, deg	∠H–D- - -A, deg
N(16)–H(16)- - - -O(29^{i})	2.685 (3)	0.98 (3)	1.74 (3)	161 (3)	12 (2)
N(21)–H(21)- - - -O(31^{ii})	2.826 (3)	0.91 (3)	1.98 (3)	153 (3)	18 (2)
N(25)–H(25)- - - -O(29^{iii})	3.028 (3)	0.90 (3)	2.16 (3)	162 (3)	13 (2)
O(31)–H(311)- - - -O(28^{iv})	2.814 (3)	0.88 (3)	1.95 (3)	167 (3)	9 (2)
O(31)–H(312)- - - -O(32)	2.749 (4)	0.85 (3)	1.93 (3)	161 (3)	13 (2)
O(32)–H(321)- - - -O(28^{v})	2.959 (4)	1.07 (3)	1.93 (3)	159 (3)	14 (2)
O(32)–H(322)- - - -O(28^{vi})	2.745 (3)	1.00 (3)	1.77 (3)	164 (3)	10 (2)
		Other interactions			
C(1)–H(1)- - - - - - -O (20)	3.401 (4)	0.98 (3)	2.45 (3)	164 (3)	11 (2)
C(8)–H(8)- - - - - - -O(24^{v})	3.312 (4)	1.04 (3)	2.33 (3)	157 (3)	16 (2)

[a] Code: (i) $x, y, z - 1$; (ii) $x, \frac{1}{2} - y, \frac{1}{2} + z$; (iii) $x, -\frac{1}{2} - y, z - \frac{1}{2}$; (iv) $2 - x, -y, 1 - z$; (v) $x, \frac{1}{2} - y, z - \frac{1}{2}$; (vi) $2 - x, \frac{1}{2} + y, 1\frac{1}{2} - z$.

The water molecules and O(31) and O(32) hold two layers of peptide groups together (along the *a* axis) via hydrogen bonding. O(32) forms hydrogen bonds to O(28) of different peptides, and O(31) interacts with N(21) of one peptide and O(28) of another. The carbonyl groups (O(20) and O(24)) serve as mediators between the water and hydrophobic layers. O(24) packs against C(8^{ii}) and O(32) (a water molecule) while O(20) packs against C(1) and O(29) (terminal carboxyl group). The C–H···O interactions are linear, possibly indicative of some acidic character in carbon atoms 1 and 8 (equivalent by symmetry). It is interesting that MASGG has almost tetragonal cell dimensions and some of this symmetry is evident in the water layer.

The packing in MACl is very similar to that in 9,10-dimethylanthracene[31] and 9,10-bis(chloromethyl)anthracene,[32] particularly down the short axis, which varies from 4.5 to 5.3 Å for the three compounds. A diagram is provided in the microfilm edition of this journal (Figure III).

The overlap of rings, i.e., their stacking in planes 3.5–3.6 Å apart, is very similar (although not absolutely identical) in MACl and MASGG (see microfilm edition for diagram, Figure IV). There is a strong tendency for methyl groups to overlap the outer rings as shown. Thus the hydrophobic layer may be well structured by these specific interactions, so that water molecules are needed in the crystal structure to complete the necessary packing of the peptide chains. Water molecules

have also been found to crystallize with glycylphenylalanylglycine[21] and with alanylalanylalanine,[33] but not with glycylglycylglycine.[34]

This study has shown the effect of alkylation by a bulky hydrophobic group on the conformation of a peptide in the crystalline state. The tendency to segregate the hydrophobic from hydrophilic regions of the molecule in packing and the formation of strong hydrogen bonds may have contributed to the experimentally observed distortion of the peptide conformation on alkylation.

Acknowledgment. We thank Drs. S. K. Kurtz and J. P. Dougherty of Philips Laboratories, for their help in demonstrating the noncentrosymmetry in MACl.[13] This research was supported by USPHS Grants CA-10925, CA-06927, and RR-05539 from the National Institutes of Health, AG 370 from the National Science Foundation, and by an appropriation from the Commonwealth of Pennsylvania.

Supplementary Material Available: The following tables and figures are also provided: Table 1, anisotropic temperature factors for SGG, MACl, and MASGG; Table 2, interatomic distances and interbond angles for MACl; Table 3, angles between planes; Figure 1, Ramachandran plot of torsion angles; Figure II, packing in MASGG. View down the *a* axis; Figure III, comparison of packing in MACl with that in 9,10-dimethylanthracene[31] and 9,10-bis(chloromethyl)anthracene;[32] Figure IV, stacking in ring systems 3.5 Å apart (11 pages). Ordering information is given on any current masthead page.

References and Notes

(1) J. A. Miller, *Cancer Res.*, **30**, 559 (1972).
(2) P. Sims, P. L. Grover, A. Swaisland, K. Pal, and A. Hewer, *Nature (London)*, **252**, 326 (1974).
(3) J. Booth and P. Sims, *FEBS Lett.*, **47**, 30 (1974).
(4) A. J. Swaisland, A. Hewer, K. Pal, G. R. Keysell, J. Booth, P. J. Grover, and P. Sims, *FEBS Lett.*, **47**, 34 (1974).
(5) M. S. Newman and D. R. Olson, *J. Am. Chem. Soc.*, **96**, 6207 (1974).
(6) S. Sorof, *Jerusalem Symp. Quantum Chem. Biochem.*, **1**, 208 (1969).
(7) P. Brookes and P. D. Lawley, *Nature (London)*, **202**, 781 (1964).
(8) R. M. Peck, T. K. Tan, and E. B. Peck, *Cancer Res.*, **36**, 2423 (1976).
(9) R. M. Peck, T. K. Tan, and E. B. Peck, *J. Med. Chem.*, **19**, 1422 (1976).
(10) G. M. Badger and R. S. Pearce, *J. Chem. Soc.*, 2317 (1950); J. L. Adelfang and G. H. Daub, *J. Am. Chem. Soc.*, **80**, 1405 (1958).
(11) G. Germain, P. Main, and M. M. Woolfson, *Acta Crystallogr., Sect. A*, **27**, 368 (1971).
(12) W. H. Zachariasen, *Acta Crystallogr.*, **16**, 1139 (1963).
(13) J. P. Dougherty and S. K. Kurtz, *J. Appl. Crystallogr.*, **9**, 145 (1976).
(14) J. M. Stewart, "The X-Ray System—version of 1972", Technical Report TR-192 of the Computer Science Center, University of Maryland, June 1972.
(15) H. J. Bernstein, L. C. Andrews, H. M. Berman, F. C. Bernstein, G. H. Campbell, H. L. Carrell, H. B. Chiang, W. C. Hamilton, D. D. Jones, D. Klunk, T. F. Koetzle, E. F. Meyer, C. N. Morimoto, S. S. Sevian, R. K. Stodola, M. M. Strongson, and T. V. Willoughby, Second Annual AEC Scientific Computer Information Exchange Meetings, Proceeding of the Technical Program, 1974, p 148.
(16) P. K. Gantzel, R. A. Sparks, R. E. Long, and K. N. Trueblood, UCLALS4 Program in Fortran IV (modified by H. L. Carrell) (1969).
(17) D. T. Cromer and J. Mann, *Acta Crystallogr., Sect. A*, **24**, 321 (1968).
(18) R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, **42**, 3175 (1965).
(19) D. T. Cromer and D. Liberman, *J. Chem. Phys.*, **53**, 1891 (1970).
(20) R. E. Marsh and J. D. Donohue, *Adv. Protein Chem.*, **22**, 234 (1967).
(21) R. E. Marsh and J. P. Glusker, *Acta Crystallogr.*, **14**, 1110 (1961).
(22) IUPAC–IUB Commission on Biochemical Nomenclature; *J. Mol. Biol.*, **52**, 1 (1970).
(23) S. Arnott and S. D. Dover, *J. Mol. Biol.*, **30**, 209 (1967).
(24) G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, *J. Mol. Biol.*, **95**, 7, (1963).
(25) J. A. Schellman and C. Schellman in "The Proteins", Vol. 2, H. Neurath, Ed., Academic Press, New York, N.Y., 1964, p 1.
(26) S. Arnott, S. D. Dover, and A. Elliott, *J. Mol. Biol.*, **30**, 201 (1967).
(27) G. N. Ramachandran, V. Sasisekharan, and C. Ramakrishnan, *Biochim. Biophys. Acta*, **112**, 168 (1966).
(28) G. N. Ramachandran, A. V. Lakshminarayanan, and A. S. Kolaskar, *Biochim. Biophys. Acta*, **303**, 8 (1973).
(29) G. N. Ramachandran, *Biopolymers*, **6**, 1494 (1968).
(30) F. K. Winkler and J. D. Dunitz, *J. Mol. Biol.*, **59**, 169 (1971).
(31) J. Iball and J. N. Low, *Acta Crystallogr., Sect. B*, **30**, 2203 (1974).
(32) E. J. Gabe and J. P. Glusker, *Acta Crystallogr., Sect. B*, **27**, 1925 (1971).
(33) J. K. Fawcett, N. Camerman, and A. Camerman, *Acta Crystallogr., Sect. B*, **31**, 658 (1975).
(34) R. Nathan, Ph.D. Thesis, California Institute of Technology, Pasadena, California, 1956.