# The Crystal and Molecular Structures of Two Isostructural
## N-(Haloacetyl)-L-phenylalanyl-L-phenylalanine Ethyl Esters, $C_{22}H_{25}XN_2O_4$*

By Chin Hsuan Wei, David G. Doherty and J. R. Einstein

*Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, U.S.A.*

The crystal and molecular structures of N-(bromoacetyl)-L-phenylalanyl-L-phenylalanine ethyl ester and the isostructural chloroacetyl derivative have been determined from three-dimensional X-ray diffractometer data. The bromo derivative crystallizes with two molecules in a monoclinic unit cell of symmetry $P2_1$ and parameters $a = 12\cdot951$ (9), $b = 17\cdot617$ (26), $c = 4\cdot951$ (3) Å, and $\beta = 103\cdot23$ (7)°. The chloro derivative has a unit cell with parameters $a = 12\cdot844$ (3), $b = 17\cdot677$ (4), $c = 4\cdot931$ (1) Å, and $\beta = 103\cdot30$ (2)°. The structure of the bromo derivative was first solved by the heavy-atom technique and refined by the full-matrix least-squares method to $R(F) = 0\cdot090$ based on 522 observed intensity data. The refinement of the chloro derivative, carried out with starting parameters obtained from the bromo derivative, gave a value of $R(F) = 0\cdot105$ based on 487 data. Because of a disorder problem, the ethoxy terminus was treated as a rigid group in the refinements for both derivatives. The molecular configuration is of the parallel-chain pleated-sheet type. In the chloro derivative, the lengths of the peptide units are 3·79 and 3·81 Å; the end-to-end length of the two units is 6·09 Å; the interchain hydrogen bond lengths are 2·78 and 2·90 Å; the backbone conformation angles are $\varphi_1 = -102°$, $\varphi_2 = -122°$, $\psi_1 = 106°$, $\psi_T^1 = -133°$, $\psi_T^2 = 53°$; and the principal side-chain conformation angles are $\chi_1^1 = -175°$ ($\gamma$-carbon position II) and $\chi_2^1 = -71°$ ($\gamma$-carbon position III).

## Introduction

Relatively few three-dimensional structure determinations have been accomplished for peptides, and very few for peptides containing more than one large side chain (Kennard & Watson, 1970). We report here the crystal and molecular structure of N-(haloacetyl)-L-phenylalanyl-L-phenylalanine ethyl ester, a substrate of pepsin and a molecule containing two aromatic groups. Although crystalline disorder prevented a very precise determination of individual interatomic distances and bond angles, the molecular conformation has been clearly established and may be of use in comparative and theoretical studies of peptide conformation.

## Experimental

*Crystal specimens*

The chloroacetyl and bromoacetyl dipeptide esters were each prepared in good yield and purity by coupling the haloacetyl halide to L-phenylalanyl-L-phenylalanine ethyl ester hydrobromide and by coupling the haloacetyl-L-phenylalanine to L-phenylalanine ethyl ester hydrochloride. Full details of the syntheses will be given elsewhere (Doherty, 1971). The crystalline products, obtained by evaporation from ethyl acetate–ligroin mixtures, were quite pure, as shown by the following analyses. Calculated for $C_{22}H_{25}ClN_2O_4$ (416·89):

C, 63·38; H, 6·04; N, 6·72. Found: C, 63·21; H, 6·22; N, 6·74. Calculated for $C_{22}H_{25}BrN_2O_4$ (461·35): C, 57·27; H, 5·46; N, 6·07. Found: C, 57·40; H, 5·56; N, 6·04. The melting points were 155–157° for the chloro derivative, 168–170° for the bromo derivative. The dipeptide esters synthesized by either method were 95% hydrolyzed in four hours by pepsin, indicating that little if any racemization had occurred in the preparations.

Samples of both derivatives were recrystallized by slow evaporation from ethanol solutions in an attempt to obtain single crystals large enough for X-ray work. Since the needle-shaped crystals were very fragile and difficult to handle, specimens were mounted in thin-walled glass capillaries. The crystal dimensions were $0\cdot39 \times 0\cdot05 \times 0\cdot03$ mm for the chloro derivative and $0\cdot26 \times 0\cdot07 \times 0\cdot05$ mm for the bromo derivative, with the longest dimension (corresponding to the c axis) in each case nearly parallel to the capillary axis.

*Unit cells and space groups*

From the preliminary Weissenberg and precession X-ray photographs with nickel-filtered Cu K$\alpha$ radiation ($\lambda = 1\cdot5418$ Å), approximate cell parameters for both compounds were obtained; and space groups for both crystals were found from systematic absences to be $P2_1$ ($C_2^2$, No. 4) (see *International Tables for X-ray Crystallography*, 1965). In each case, the crystal was transferred from the camera to an Oak Ridge automatic computer-controlled diffractometer (Busing, Ellison, Levy, King & Roseberry, 1968). The cell parameters, together with the orientational parameters of the crystal, were refined by the method of least

---

squares from angle data of six reflections for the chloro derivative and nine reflections for the bromo derivative. The cell parameters thus determined, together with other crystal data, are tabulated in Table 1.

Table 1. *Unit-cell parameters of* $C_{22}H_{25}XN_2O_4$, *with estimated standard errors*

|   | Chloro derivative | Bromo derivative |
|---|---|---|
| *a* | 12·844 (3) Å | 12·951 (9) Å |
| *b* | 17·677 (4) | 17·617 (26) |
| *c* | 4·931 (1) | 4·951 (3) |
| *β* | 103·30 (2)° | 103·23 (7)° |
| *Z* | 2 | 2 |
| *V* | 1089·5 Å³ | 1099·6 Å³ |
| $D_o$* | 1·27 (1) g.cm⁻³ | 1·39 (1) g.cm⁻³ |
| $D_c$ | 1·27 g.cm⁻³ | 1·39 g.cm⁻³ |
| *F*(000) | 440 | 476 |
| Space group | $P2_1$ | $P2_1$ |

* The observed densities of the compounds were obtained by the flotation method.

*X-ray intensity data*

(*a*) *Chloro derivative.* The three-dimensional data were measured with the automatic diffractometer using Cu *K*α radiation. Reflections with $2\theta \le 60°$ were measured by an ω step-scan technique (Levy, 1966; Brown, 1969) in which three scans were run for each reflection: a main scan with the detector set at the computed $2\theta$, and two background scans with the detector offset lower and higher in $2\theta$. The detector-slit width and the $2\theta$ offset were adjusted for different ranges of $2\theta$ in a manner similar to that described by Brown (1969).

Reflections in the range 60–90° $2\theta$, selected by rapid screening, were measured by $\theta$–$2\theta$ step scans. Because the reflections were in general very weak, no filter was used, and two reflections (900 and 10,00) significantly affected by *K*β overlap were removed from the data set. The intensity of a reference reflection, measured periodically throughout the data collection (overall decrease 8%), was used to normalize the data for the effects of radiation damage. Lorentz–polarization corrections, and an empirical absorption correction depending only on *φ* and based on azimuthal scans of three reflections, were applied.

Variances $\sigma^2(F_o^2)$ for use in least-squares weighting were estimated according to the expression:

$$\sigma^2(F_o^2) = K^2[G + (t_G/t_B)^2 B]/(Lp)^2,$$

in which *G* is the gross intensity count, *B* is the background count, $t_G/t_B$ is the ratio of the counting times of gross intensity and background, *K* is the appropriate scaling constant, and Lp is the Lorentz–polarization factor. Where several measurements were made, the average $F_o^2$ value and its variance were computed, using as weights the reciprocals of the variances for the individual observations. Finally, the quantity (0·03 $F_o^2)^2$ was added to the variance of the average $F_o^2$, in order to make allowance for possible instability in the instrument, errors in absorption corrections, deficiencies of the structure model, *etc.* (Peterson & Levy, 1957).

After the preliminary processing of the intensity data, 566 independent reflections with intensities greater than zero were obtained, of which 487 had $F_o^2 \ge 3\sigma(F_o^2)$.

(*b*) *Bromo derivative* Measurement and initial processing of the data paralleled that for the chloro derivative, except that $\theta$–$2\theta$ step scans with no filter were used for the entire $2\theta$ range of 0–90°, and that no absorption corrections were applied. Of the 864 independent reflections having intensities greater than zero, 522 had $F_o^2 \ge 3\sigma(F_o^2)$.

## Solution and refinement of the structures

The structure of the bromo derivative was solved by conventional heavy-atom methods, with the use of a three-dimensional Patterson function computed with the Blount (1965) program, full-matrix least-squares refinements (Busing, Martin & Levy, 1962), and successive Fourier syntheses. After 29 independent non-hydrogen atoms had been located, the *R*(*F*) value, defined as $\sum||F_o| - |F_c||/\sum|F_o|$, was 0·12, with the thermal motion of the bromine atom treated anisotropically, and the thermal motion of the other atoms treated isotropically. Refinement of the chloro derivative data, with starting parameters obtained from the bromo derivative, yielded an *R*(*F*) value of 0·10 in a similar refinement. Although the structure appeared to be essentially correct, the bond lengths and bond angles involving the ethoxy group [in the bromo derivative, O(26)–C(27) = 1·80 Å; C(27)–C(28) = 1·23 Å; O(26)–C(27)–C(28) = 74°] did not conform to a physically meaningful arrangement. For both derivatives, difference Fourier syntheses phased on all atoms except the ethyl carbons revealed two very broad peaks of electron density in the region of the ethyl group. Since refinements with both derivatives yielded essentially the same abnormal arrangement of the ethyl group, we have concluded that these results were not due to error, but that the ethyl group was disordered in a similar
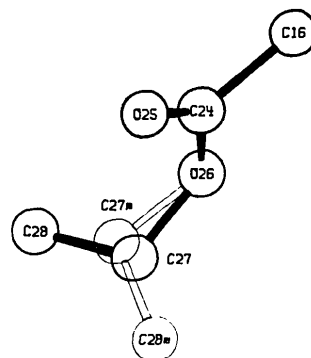


Fig. 1. Two configurations for the ethyl group found by a model-building computer program to be consistent with packing requirements. The atoms in bold outline lie close to the positions obtained by refinement with a single ethoxy group.

manner in both crystals, and that the refined position of each ethyl carbon atom represented the average of two or more sites.

With the aid of a model-building computer program (assuming standard bond distances and angles and varying torsion angles) it was determined that two general configurations for the ethoxy group were consistent with packing restrictions, i.e., implied no interatomic distances 0·2 Å less than the normal van der Waals distances. The pairs of positions, roughly related by mirror symmetry across the plane of the carboxyl group, are illustrated in Fig. 1.

A number of attempts were made to refine, by the method of least-squares, a randomly disordered model with two alternative configurations for the ethoxy group. With the use of a rigid-group refinement program, kindly supplied by Dr L. F. Dahl of the University of Wisconsin, constraints were applied in several

Table 2. *Positional, thermal, and group parameters for* $C_{22}H_{25}XN_2O_4$*

| Atom | | $x$ | $y$ | $z$ | $B$ |
|---|---|---|---|---|---|
| X | Cl | 0·4666 (8) | $\frac{1}{4}$ | 0·4949 (25) | …† |
| | Br | 0·4655 (4) | $\frac{1}{4}$ | 0·5062 (10) | …† |
| C(1) | Cl | 0·374 (3) | 0·284 (2) | 0·678 (7) | 6·1 (9) |
| | Br | 0·371 (3) | 0·302 (2) | 0·676 (7) | 6·3 (10) |
| C(2) | Cl | 0·273 (2) | 0·307 (2) | 0·463 (7) | 4·8 (7) |
| | Br | 0·268 (2) | 0·323 (2) | 0·479 (7) | 3·9 (7) |
| O(3) | Cl | 0·259 (2) | 0·308 (2) | 0·202 (5) | 5·6 (5) |
| | Br | 0·256 (2) | 0·326 (1) | 0·222 (5) | 5·1 (5) |
| N(4) | Cl | 0·197 (2) | 0·323 (2) | 0·601 (5) | 4·5 (6) |
| | Br | 0·192 (2) | 0·339 (1) | 0·610 (5) | 3·7 (6) |
| C(5) | Cl | 0·086 (2) | 0·349 (2) | 0·452 (5) | 3·6 (6) |
| | Br | 0·087 (2) | 0·365 (2) | 0·468 (5) | 2·8 (6) |
| C(6) | Cl | 0·002 (3) | 0·307 (2) | 0·565 (6) | 5·5 (8) |
| | Br | 0·004 (2) | 0·318 (2) | 0·582 (6) | 4·6 (8) |
| C(13) | Cl | 0·082 (2) | 0·434 (2) | 0·494 (6) | 4·0 (6) |
| | Br | 0·084 (2) | 0·448 (2) | 0·502 (7) | 3·9 (7) |
| O(14) | Cl | 0·085 (2) | 0·458 (2) | 0·736 (5) | 6·6 (6) |
| | Br | 0·084 (2) | 0·476 (1) | 0·748 (5) | 4·7 (5) |
| N(15) | Cl | 0·084 (2) | 0·483 (2) | 0·293 (5) | 4·0 (5) |
| | Br | 0·081 (2) | 0·494 (1) | 0·295 (4) | 2·5 (5) |
| C(16) | Cl | 0·095 (2) | 0·562 (2) | 0·336 (8) | 6·2 (8) |
| | Br | 0·095 (3) | 0·578 (2) | 0·320 (7) | 5·9 (9) |
| C(17) | Cl | 0·190 (3) | 0·597 (2) | 0·229 (7) | 6·5 (9) |
| | Br | 0·185 (3) | 0·608 (2) | 0·247 (7) | 5·1 (8) |
| C(24) | Cl | −0·011 (4) | 0·598 (3) | 0·130 (10) | 9·1 (12) |
| | Br | −0·010 (4) | 0·613 (3) | 0·152 (10) | 8·3 (12) |
| O(25) | Cl | −0·008 (2) | 0·651 (2) | −0·023 (6) | 9·5 (8) |
| | Br | −0·010 (2) | 0·659 (2) | −0·031 (6) | 9·0 (8) |

| Group‡ | | $x_c$ | $y_c$ | $z_c$ | $\varphi$ | $\theta$ | $\varrho$ |
|---|---|---|---|---|---|---|---|
| PH(1)§ | Cl | −0·214 (1) | 0·340 (1) | 0·267 (4) | −2·86 (2) | 2·45 (2) | −0·70 (2) |
| | Br | −0·213 (1) | 0·354 (1) | 0·280 (4) | −2·89 (2) | 2·50 (2) | −0·66 (2) |
| PH(2)§ | Cl | 0·388 (2) | 0·545 (1) | 0·619 (4) | 0·22 (2) | −2·51 (2) | 0·84 (2) |
| | Br | 0·381 (2) | 0·557 (1) | 0·590 (4) | 0·19 (2) | −2·45 (2) | 0·81 (2) |
| ETO | Cl | −0·200 (4) | 0·604 (3) | 0·036 (10) | 0·92 (7) | 3·00 (4) | −1·99 (6) |
| | Br | −0·201 (3) | 0·618 (3) | 0·046 (10) | 0·92 (7) | 3·05 (5) | −2·03 (9) |

* Standard deviations in units of the last significant figures are given in parentheses here and in Table 3.
† For the halogen atoms anisotropic temperature factors of the form: $\exp[-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+2\beta_{12}hk+2\beta_{13}hl+2\beta_{23}kl)]$ were used; the resulting thermal coefficients $\times 10^4$ are as follows:

| | | $\beta_{11}$ | $\beta_{22}$ | $\beta_{33}$ | $\beta_{12}$ | $\beta_{13}$ | $\beta_{23}$ |
|---|---|---|---|---|---|---|---|
| X | Cl | 86 (7) | 191 (12) | 805 (60) | 58 (9) | 20 (17) | −95 (26) |
| | Br | 118 (4) | 132 (4) | 795 (28) | 61 (4) | 18 (8) | −109 (11) |

‡ $x_c$, $y_c$, and $z_c$ are the fractional coordinates of the origins for the rigid phenyl-ring groups, PH(1) and PH(2), and for the ethoxy group, ETO. The rigid-body positional coordinates are defined relative to right-handed orthogonal axes $x'$, $y'$, and $z'$ with origins at the centers of the phenyl rings and at C(27) of the ETO group, and with each group in the $x'y'$ plane. In PH(1), C(10) lies on the $+x'$ axis, and the $+y'$ axis intersects the midpoint of C(8)–C(9). In PH(2), C(21) lies on the $+x'$ axis, and the $+y'$ axis intersects the midpoint of C(22)–C(23). In the ETO group C(28) lies on the $+x'$ axis, and the $+z'$ direction is the vector product of $+x'$ and the vector C(27)→O(26). For each group, the three angular parameters $\varphi$, $\theta$, and $\varrho$ (in radians) define the orientation of the rigid-body axial system with respect to an orthogonal set of crystal axes $a_0, b_0, c_0$, where $a_0 = a, b_0 = c_0 \times a_0$, and $c_0 = a \times b$. Three successive rotations bring the crystal axes into parallel coincidence with the rigid-body axes as follows. A positive rotation $\varphi$ about $c_0$ brings $a_0$ to $a_1$ and $b_0$ to $b_1$; next, a positive rotation $\theta$ about $a_1$ brings $b_1$ to $b'$ and $c_0$ to $c_1$; lastly, a positive rotation $\varrho$ about $b'$ brings $a_1$ to $a'$ and $c_1$ to $c'$.

§ PH(1) and PH(2) are bonded to C(6) and C(17), respectively.

different ways to the bond distances, bond angles, oc-
cupancies, and thermal parameters. However, none of
these attempts with either the chloro derivative or the
bromo derivative data led to convergence. Refinement
did converge when the structure was assumed to con-
tain a single ethoxy group, treated as a rigid group
[O(26)–C(27) = 1·47 Å;  C(27)–C(28) = 1·54 Å;  O(26)–
C(27)–C(28) = 109·5°] of occupancy factor 1·0. The
refined positions of the group atoms then corresponded
closely to those indicated in bold outline in Fig. 1. In
the final refinements, the two phenyl rings were also
constrained to possess $D_{6h}$ symmetry in order to improve
the ratio of the number of observations to the number
of varied parameters. The orthogonal axes for the rigid
ethoxy and phenyl groups were as defined in Table 2.
The thermal motions of the halogen atoms were treated
anisotropically, whereas all other atoms, independent
or within groups, received individual isotropic thermal
parameters. In the following, numerical results of the
refinement for the chloro derivative are followed by
those for the bromo derivative in parentheses. Based
on the 487(522) reflections for which $F_o^2 \geq 3\sigma(F_o^2)$, $R(F)$
was  0·105  (0·090);  $R_w(F)$,  defined  as  $[\sum w||F_o| -
|F_c||^2/\sum w|F_o|^2]^{1/2}$, was 0·112 (0·111); and the standard
deviation of an observation of unit weight, defined as
$[\sum w|\Delta|F||^2/(n-p)]^{1/2}$, where $n$ is the number of observa-
tions and $p$ the number of parameters fitted to the data
set, was 3·84 (1·77). The parameter shifts in the final
cycle were all less than 20 (20)% of the corresponding
standard deviations. Final difference Fourier syntheses

were essentially flat except in the immediate vicinity of
the ethyl group, where the highest residual was 0·4 (0·6)
e.Å$^{-3}$, or about 20% of the average peak height of a
carbon atom in these structures. No attempt was made
to locate the positions of the 22 hydrogen atoms in
either derivative.

The scattering factors used were those of Thomas &
Umeda (1957) for the bromine atom, those of Dawson
(1960) for the chlorine atom, and those of Berghuis,
Haanappel, Potters, Loopstra, MacGillavry & Vee-
nendaal (1955) for carbon, nitrogen, and oxygen. The
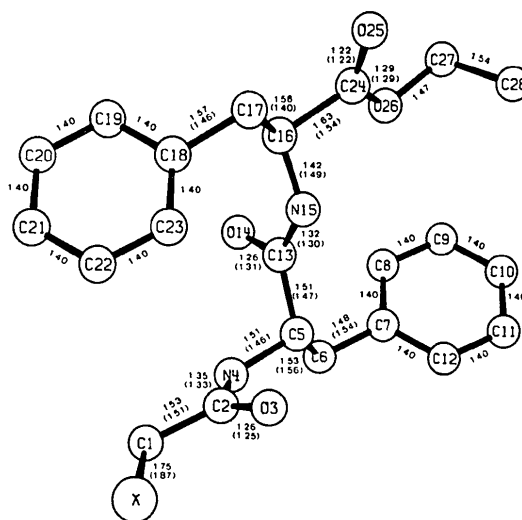anomalous scattering corrections of Templeton (1962)



Fig. 2. Molecular configurations with bond distances (Å) for
the two derivatives with values for the bromo derivative in
parentheses.

Table 3. *Derived parameters of group atoms for*
$C_{22}H_{25}XN_2O_2$

| Group | Atom | | x | y | z | B |
|---|---|---|---|---|---|---|
| PH(1) | C(7) | Cl | −0·109 | 0·326 | 0·412 | 6·6 (9) |
| | | Br | −0·108 | 0·342 | 0·424 | 6·2 (9) |
| PH(1) | C(8) | Cl | −0·167 | 0·384 | 0·501 | 8·6 (11) |
| | | Br | −0·166 | 0·401 | 0·503 | 7·0 (10) |
| PH(1) | C(9) | Cl | −0·273 | 0·398 | 0·356 | 9·7 (12) |
| | | Br | −0·271 | 0·414 | 0·360 | 9·9 (13) |
| PH(1) | C(10) | Cl | −0·319 | 0·355 | 0·123 | 8·3 (11) |
| | | Br | −0·318 | 0·367 | 0·137 | 7·2 (10) |
| PH(1) | C(11) | Cl | −0·261 | 0·297 | 0·034 | 8·0 (10) |
| | | Br | −0·260 | 0·307 | 0·058 | 10·8 (14) |
| PH(1) | C(12) | Cl | −0·156 | 0·282 | 0·179 | 7·3 (9) |
| | | Br | −0·155 | 0·295 | 0·201 | 7·7 (11) |
| PH(2) | C(18) | Cl | 0·291 | 0·567 | 0·445 | 5·4 (8) |
| | | Br | 0·284 | 0·583 | 0·429 | 5·2 (9) |
| PH(2) | C(19) | Cl | 0·336 | 0·610 | 0·682 | 8·5 (11) |
| | | Br | 0·333 | 0·622 | 0·670 | 6·1 (9) |
| PH(2) | C(20) | Cl | 0·433 | 0·587 | 0·857 | 8·3 (10) |
| | | Br | 0·430 | 0·596 | 0·832 | 9·9 (13) |
| PH(2) | C(21) | Cl | 0·485 | 0·522 | 0·794 | 10·1 (12) |
| | | Br | 0·478 | 0·531 | 0·751 | 9·4 (13) |
| PH(2) | C(22) | Cl | 0·440 | 0·479 | 0·557 | 12·5 (17) |
| | | Br | 0·428 | 0·492 | 0·509 | 12·9 (19) |
| PH(2) | C(23) | Cl | 0·343 | 0·502 | 0·382 | 11·6 (15) |
| | | Br | 0·332 | 0·518 | 0·348 | 7·8 (11) |
| ETO | O(26) | Cl | −0·098 | 0·568 | 0·170 | 13·9 (12) |
| | | Br | −0·098 | 0·582 | 0·164 | 11·4 (10) |
| ETO | C(27) | Cl | −0·200 | 0·604 | 0·036 | 13·8 (17) |
| | | Br | −0·201 | 0·618 | 0·046 | 12·2 (16) |
| ETO | C(28) | Cl | −0·243 | 0·570 | −0·254 | 19·8 (27) |
| | | Br | −0·250 | 0·583 | −0·240 | 27·9 (41) |
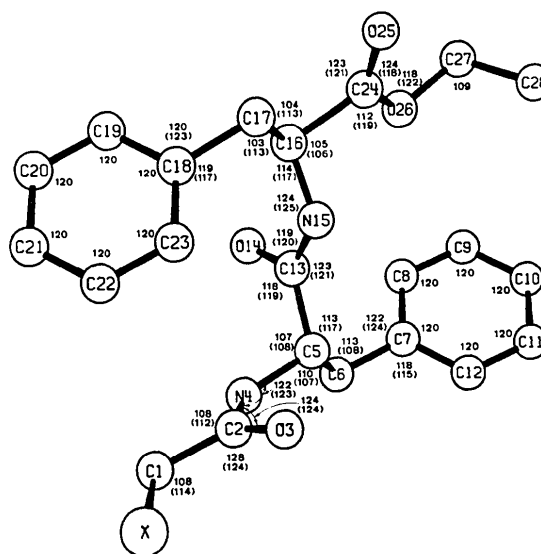


Fig. 3. Molecular configurations with bond angles (°) for the
two derivatives with values for the bromo derivative in
parentheses.

for bromine and of Cromer (1965) for chlorine were employed.

The final parameters for nongroup atoms and rigid groups, together with the corresponding standard deviations, are given in Table 2 for both compounds. The derived parameters for the group atoms are given in Table 3. Observed and calculated structure amplitudes ($\times 10$) for the reflections used in the refinements are listed for the chloro derivative and the bromo derivative in Tables 4 and 5, respectively. Bond lengths and angles, calculated with program *ORFFE* (Busing, Martin & Levy, 1964) are illustrated in Figs. 2 and 3. All figures in this paper were prepared with the use of the *ORTEP* program (Johnson, 1965). The calculations were done on the CDC 1604 and IBM 360 computers.

## Discussion

### Interatomic distances and bond angles: accuracy of the structure determination

For both derivatives the estimated standard deviations (e.s.d.'s) are in the ranges 0·03–0·05 Å for distances and 2–4° for angles. This relatively poor precision for counter data may be attributed to the small number of intensity data, and to the disorder for which no satisfactory model could be refined. However, the structures for the two derivatives agree closely with each other and, in many details, with previously determined peptide structures. Corresponding distances and angles for the two derivatives involving independently refined (nongroup) atoms agree to within 3 e.s.d.'s

Table 4. *Observed and calculated structure amplitudes ($\times 10$) for* $C_{22}H_{25}ClN_2O_4$

[Dense tabulated structure-factor data omitted.]

Table 5. *Observed and calculated structure amplitudes ($\times 10$) for* $C_{22}H_{25}BrN_2O_4$

[Dense tabulated structure-factor data omitted.]

(except for the carbon–halogen bond lengths). In a comparison of both derivatives of the main-chain distances and angles with the corresponding weighted-average values calculated by Marsh & Donohue (1967), the root-mean-square (r.m.s.) difference for bond distances is 0·03 Å, and the maximum difference is 0·07 Å; the r.m.s. difference for bond angles is 4°, and the maximum difference is 7·5°. At positions where rigid groups are attached to independent atoms, the interatomic distances and bond angles are not significantly different (considering the low precision) from normal values. The parameters of interest at the ester linkage are close to published values for three other esters, as shown ·n Table 6. Thus, it appears certain that the molecular conformations are essentially correct as determined, except for some uncertainty in the region of the disordered ethyl group.

The C(1)–Cl distance of 1·75 Å and C(1)–Br distance of 1·87 Å are roughly in agreement with the average values of 1·77 and 1·94 Å, respectively, given in *Tables of Interatomic Distances and Configuration in Molecules and Ions* (1965).

*Molecular structure and packing*

A stereoscopic view of the molecule is shown in Fig. 4. Fig. 5 shows the hydrogen bonding of molecules separated by c translations, and Fig. 6 shows the pack-

ing in a stereoscopic view down the c axis. The structure is of the parallel-chain pleated-sheet (PCPS)* type (Pauling & Corey, 1951), as has also been found for the tripeptide Gly-L-Phe-Gly (GPG)* (Marsh & Glusker, 1961). The molecule may be considered to contain two complete peptide units (as does a tripeptide), where we use the term 'peptide unit' to denote a group $C^\alpha$.CO.NH.$C^\alpha$, and where the atom C(1) takes the place of an α-carbon atom. As expected (Marsh & Donohue, 1967), both peptide units have *trans* configuiation and are planar within experimental error [maximum out-of-plane deviation 0·05 (0·08) Å, calculated by means of the Smith (1962) program]. Parallel chains related by c-axis translations are connected by NH···O hydrogen bonds to form a sheet with a single pleat. The NH···O distances of 2·90 (2·86) Å and 2·78 (2·74) Å are in rough accord with the value of 2·87 Å for the PCPS, and with the value of 2·90 Å suggested by Marsh & Donohue (1967) as normal for the bonds. The distance between adjacent chains within the sheet is the c-axis period, 4·93 (4·95) Å,

---

* Abbreviations used in the text are PCPS: parallel-chain pleated sheet (Pauling & Corey, 1951, 1953); GPG: glycyl-phenylalanyl-glycine (Marsh & Glusker, 1961); and TPPNBE HBr: threonyl-phenylalanine-p-nitrobenzyl ester hydrobromide (Mallikarjunan, Rao & Venkatesan, 1969).

Table 6. *Comparison of molecular parameters at ester linkages in several structures*

$$
\begin{array}{c}
O \\
|b \\
C \qquad C \\
a/ \quad \backslash c \; d/ \\
C \qquad O
\end{array}
$$

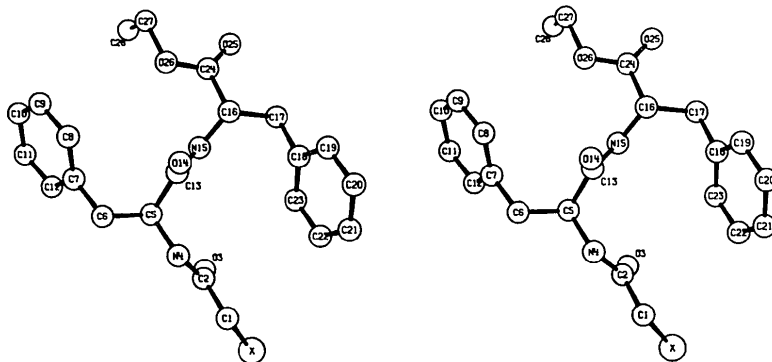|  | c(Å) | ∠ac(°) | ∠bc(°) | ∠cd(°) | Reference |
|---|---|---|---|---|---|
| Chloro derivative | 1·29 | 112 | 124 | 118 | This work |
| Bromo derivative | 1·29 | 119 | 118 | 122 | This work |
| Methyl formate | 1·334 | — | 125·9 | 114·8 | Curl (1959). |
| TPPNBE HBr | 1·34 | 113 | 119 | 118 | Mallikarjunan, Rao & Venkatesan (1969) |
| L-Cysteine ethyl ester HCl-urea (1:1) | 1·31 | 109 | 127 | 117 | Haas (1966) |



Fig. 4. Stereoscopic view of the molecule.

which is close to the values 4·85 Å for the PCPS and 4·90 Å for GPG.

Whereas the peptide chain in GPG is extended in comparison with the PCPS structure, here the chain is considerably contracted (Table 7) with respect to the PCPS model. The pleat angle, or dihedral angle between the peptide-unit planes, is smaller here by 11·5 (11·8)° than in the PCPS, and accordingly, the C(1)–C(16) distance, corresponding to the repeat distance along chains in the PCPS, is shortened by 0·41 (0·44) Å. For all three structures in Table 7, the lengths of the peptide units are close to 3·78 Å, the length of a peptide unit of average dimensions given by Marsh & Donohue (1967). In each case of Table 7, the repeat distance or its equivalent is approximately equal to 2 × (peptide unit length) × sin ($\frac{1}{2}$ pleat angle), a relation which holds exactly if the plane of the three successive $\alpha$-carbon atoms is perpendicular to the planes of the two peptide units.

The two side chains of the molecule lie on opposite sides of the pleated sheet, and are very roughly related by a pseudocenter of symmetry midway between the $\alpha$-carbon atoms of the second peptide unit, so that the molecule is compact and in the form of the letter H (see Figs. 4 & 6). The phenyl rings are nearly parallel with dihedral angle 11·1 (10·8)°.

The carboxy group, comprised of C(16), C(24), O(25), and O(26), is nearly planar with a maximum out-of-plane deviation of 0·03 (0·07) Å.

Both phenyl rings of the molecule stack along the c direction. The closest C · · · C contacts are 3·49 (3·64) Å and 3·42 (3·39) Å, for the rings in residues 1 and 2, respectively. (Residues are numbered sequentially from the haloacetyl-terminal to the carboxyl-terminal end of the chain.)

The atoms of a molecule lie not far from a plane roughly perpendicular to c. The closest contacts between the two vertical columns of molecules in Fig. 6 involve halogen atoms, phenyl rings, and methyl groups. The smallest interatomic distances between neighboring molecules fall within normal van der Waals limits.

### Conformation of the molecule

The peptide-chain conformation may be described according to the recommendations of the *IUPAC-IUB Commission on Biological Nomenclature* (1970). In these definitions, the torsion angle about bond $B$–$C$ in a system of four atoms $\overset{A}{\diagdown}B$–$C\overset{D}{\diagup}$ is considered positive (negative) if in a view along $B$–$C$ the front bond $A$–$B$ must be rotated clockwise (counterclockwise) in

Table 7. *Comparison of peptide-chain parameters with parameters of related structures*

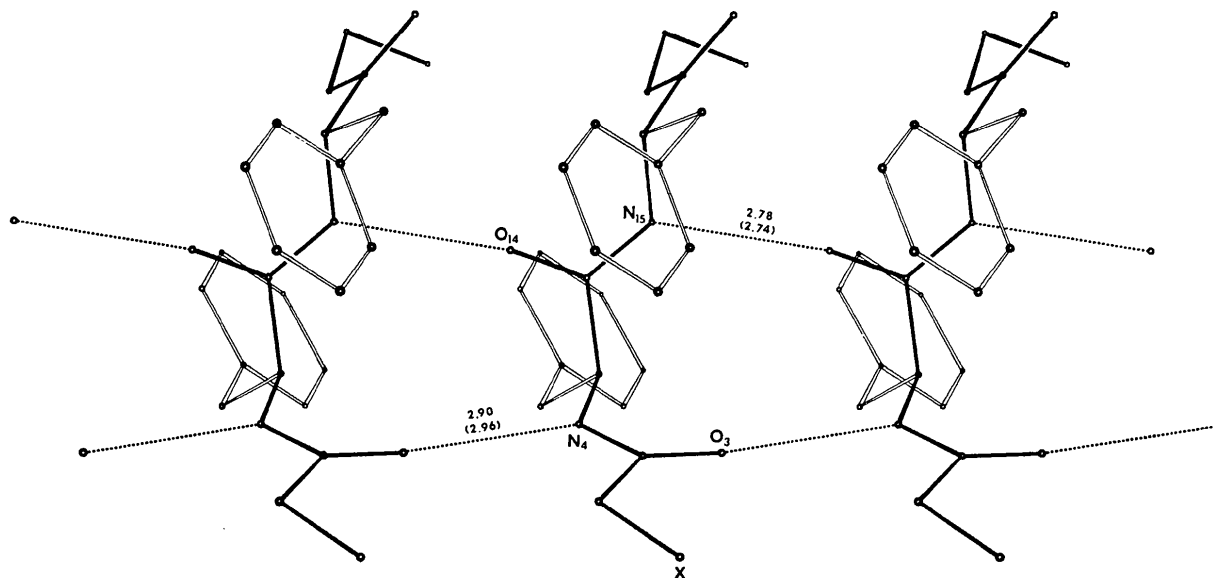| Structure | Length of peptide unit(s) | Dihedral angle | Repeat distance or equivalent | Reference |
|---|---|---|---|---|
| PCPS | 3·80 Å | 117·8° | 6·50 Å | Pauling & Corey (1951, 1953) |
| GPG | 3·80, 3·77 | 126·2 | 6·74 | Marsh & Glusker (1961) |
| Chloro derivative | 3·79, 3·81 | 106·3 | 6·09 | This work |
| Bromo derivative | 3·76, 3·82 | 106·0 | 6·06 | This work |



Fig. 5. Hydrogen-bonding of parallel chains. The c axis is horizontal and to the left. The direction of view is almost down the $a^*$ axis. For clarity, the main-chain bonds are indicated by solid lines.

order to eclipse the rear bond *C–D*. The pertinent torsion angles are illustrated in Fig. 7.

The backbone chain conformation around the α-carbon C(5) is given by the torsion angles $\varphi_1 = -102$ ($-100$)°, $\psi_1 = 106$ (110)°. These values can be compared with those for the PCPS ($-126$, 132°) and for the phenylalanyl residue in GPG ($-119$, 113°). The pair of values lies within an allowed region of the $\varphi$–$\psi$ diagram of Ramachandran and co-workers as redrawn with the conventions of the *IUPAC-IUB Commission*

(1970, Fig. 6). The torsion angle $\varphi_2$, $-122$ ($-118$)°, is close to the value for the PCPS structure. However, the torsion angles $\psi^1_T$ and $\psi^2_T$ at the carboxyl terminus have the uncommon values of $-133$ ($-126$)° and 53 (40)°; $\psi^2_T$ is usually approximately 0°, or in a minority of cases somewhat negative (down to $-56$°); and $\psi^1_T \simeq \psi^2_T + 180°$ (Lakshminarayanan, Sasisekharan & Ramachandran, 1967, Table 2). The same reference lists only two cases with large positive $\psi^2_T$: aspartic acid-HCl, $\psi^2_T = 42.3°$ (Dawson, private communica-
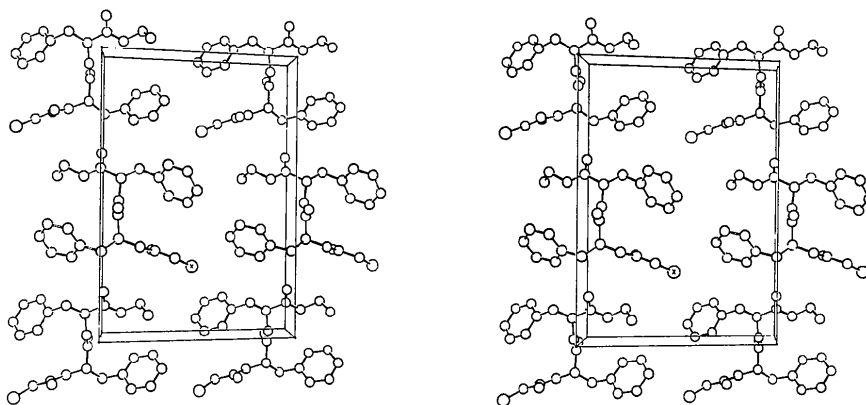


Fig. 6. Stereoscopic view down the *c* axis through the center of the unit cell. For clarity, only one row of the molecules approximately perpendicular to the viewing direction is shown. The origin is in the lower-left rear corner, and the *b* axis is vertical. Only the halogen atom of the basic molecule is labeled.
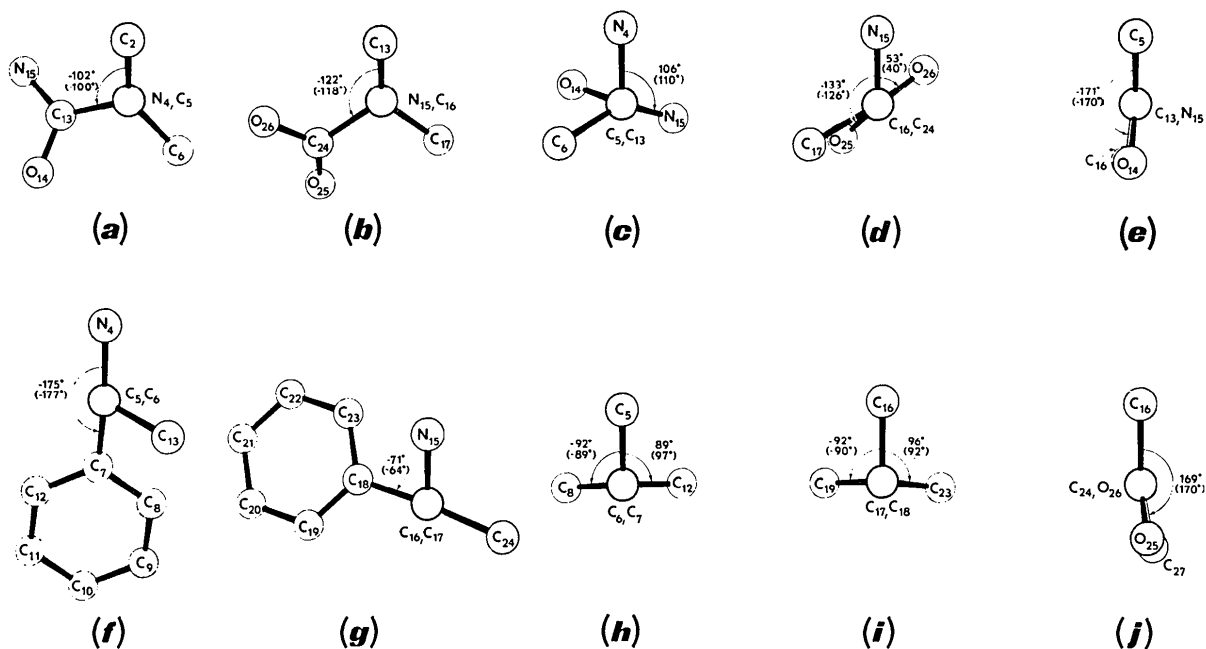


Fig. 7. Conformation angles for the two derivatives with values for the bromo derivative in parentheses: (a) $\varphi_1$; (b) $\varphi_2$; (c) $\psi_1$; (d) $\psi^1_T$ to O(25) and $\psi^2_T$ to O(26); (e) $\omega_1$; (f) $\chi^1_1$; (g) $\chi^1_2$; (h) $\chi^{2,1}_1$ to C(12) and $\chi^{2,2}_1$ to C(8); (i) $\chi^{2,1}_2$ to C(19) and $\chi^{2,2}_2$ to C(23); (j) $\theta^C_T$. According to the recommendations of the *IUB-IUPAC Commission* (1970, section 2·3·2), *i.e.* that the smallest positive or negative value of the principal torsion angle be assigned the highest priority and the lowest number, the $\chi^{2,1}_1$ and $\chi^{2,2}_1$ values for the bromo derivative should be reversed so that $\chi^{2,1}_1 = -89°$ and $\chi^{2,2}_1 = 97°$. However, to be compatible with the equivalent values for the chloro derivative, the present values for the bromo derivative are given instead.

tion to Lakshminarayanan *et al.*, 1967); and glycyl–asparagine, $\psi_T^2 = 63\cdot3°$ (Pasternak, Katz & Corey, 1954). In the latter structure the two terminal carboxyl oxygen atoms are involved in four hydrogen bonds in all, but here there are no hydrogen bonds involving O(25) or O(26). Packing may, however, explain the unusual conformation here: a rotation of the carboxyl group about the C(16)–C(24) bond to make $\psi_T^2 \sim 0°$, with atoms 1 through 23 kept stationary in the unit cell, would bring O(25) too close to one of the hydrogen atoms attached to C(6) in the neighboring molecule at $-x, \frac{1}{2}+y, 1-z$ (see Fig. 6). The estimated O$\cdots$H interatomic contact, assuming a normal tetrahedral configuration of hydrogen atoms about C(6), would then be less than 2 Å.

Conformational angles for the two side chains are shown in Fig. 7(*f*) through (*i*). The three general positions for a $C^\gamma$ atom have been labeled by Ramachandran & Lakshminarayanan (1966) as I, II, and III for $\chi$ values of 60, 180, and $-60°$, respectively. In this structure the $C^\gamma$ atom of the first residue, C(7), is within 5 (3)° of position II; and that of the second residue, C(18), is within 11 (4)° of position III. Position III has not previously been observed for a phenylalanyl residue in a small peptide. Position II has been found previously for the phenylalanyl residues in GPG and TPPNBE HBr, and position I in L-phenylalanine-HCl (Gurskaya, 1964).

Views down the $C^\beta$–$C^\gamma$ bonds of the two side chains [Fig. 7(*h*) and (*i*)] show that in both cases the planes of the phenyl rings are roughly perpendicular to the $C^\alpha$–$C^\beta$–$C^\gamma$ planes, so that all these $\chi$ angles approximate 90°; the perpendicular arrangement results in minimum interaction between $\alpha$- and $\delta$-carbon atoms, and not surprisingly has been found previously for several phenylalanyl residues (Lakshminarayanan *et al.*, 1967, Table 3).

## References

BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOPSTRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A. L. (1955). *Acta Cryst.* **8**, 478.

BLOUNT, J. F. (1965). *FOURIER, A Three-Dimensional Crystallographic Fourier Summation Program for the* CDC 1604 *Computer*. Ph.D. Thesis (Appendix), Univ. of Wisconsin, Madison, Wisconsin.

BROWN, G. M. (1969). *Acta Cryst.* B**25**, 1338.

BUSING, W. R., ELLISON, R. D., LEVY, H. A., KING, S. P. & ROSEBERRY, R. T. (1968). *The Oak Ridge Computer-Controlled X-ray Diffractometer*. Report ORNL-4143, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS, A Fortran Crystallographic Least-Squares Program*. Report ORNL-TM-305, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1964). *ORFFE, A Fortran Crystallographic Function and Error Program*. Report ORNL-TM-306, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

CROMER, D. T. (1965). *Acta Cryst.* **18**, 17.

CURL, R. F. JR (1959) *J. Chem. Phys.* **30**, 1529..

DAWSON, B. (1960). *Acta Cryst.* **13**, 403.

DOHERTY, D. G. (1971). To be published.

GURSKAYA, G. V. (1964). *Kristallografiya*, **9**, 839.

HAAS, D. J. (1966). *Acta Cryst.* **19**, 860.

*International Tables for X-ray Crystallography* (1965). Vol. I, pp. 78, 79. Birmingham: Kynoch Press.

IUPAC-IUB Commission on Biochemical Nomenclature (1970). *Biochemistry*, **9**, 3471. Article also appears in *Biochem. J.* (1971). **121**, 577; *Biochim. Biophys. Acta* (1971). **229**, 1; *Eur. J. Biochem.* (1970). **17**, 193; *J. Biol. Chem.* (1970). **245**, 6489; *J. Mol. Biol.* (1970). **52**, 1.

JOHNSON, C. K. (1965). *ORTEP, A Fortran Thermal-Ellipsoid Plot Program for Crystal-Structure Illustrations*. Report ORNL-3794, revised, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

KENNARD, O. & WATSON, D. G. (1970). *Molecular Structures and Dimensions,* Vol. 1 ('*Bibliography* 1935–69, *General Organic Crystal Structures*'), pp. 343–355. Utrecht: Oosthoek.

LAKSHMINARAYANAN, A. V., SASISEKHARAN, V. & RAMACHANDRAN, G. N. (1967). *Conformation of Biopolymers*. Edited by G. N. RAMACHANDRAN, Vol. 1, p. 61. London: Academic Press.

LEVY, H. A. (1966). Unpublished work.

MALLIKARJUNAN, M., RAO, S. T. & VENKATESAN, K. (1969). *Acta Cryst.* B**25**, 220.

MARSH, R. E. & DONOHUE, J. (1967). *Advanc. Protein Chem.* **22**, 235.

MARSH, R. E. & GLUSKER, J. P. (1961). *Acta Cryst.* **14**, 1110.

PAULING, L. & COREY, R. B. (1951). *Proc. Nat. Acad. Sci.* **37**, 251.

PAULING, L. & COREY, R. B. (1953). *Proc. Nat. Acad. Sci.* **39**, 253.

PASTERNAK, R. A., KATZ, L. & COREY, R. B. (1954). *Acta Cryst.* 7, 225.

PETERSON, S. W. & LEVY, H. A. (1957). *Acta Cryst.* **10**, 70.

RAMACHANDRAN, G. N. & LAKSHMINARAYANAN, A. V. (1966). *Biopolymers*, **4**, 495.

SMITH, D. L. (1962). *PLANE1, A Least-Squares Plane Program for the CDC* 1604 *Computer*. Ph.D. Thesis (Appendix IV), Univ. of Wisconsin, Madison, Wisconsin.

*Tables of Interatomic Distances and Configuration in Molecules and Ions* (1965). pp. S14s & S17s. London: The Chemical Society.

TEMPLETON, D. H. (1962). *International Tables for X-ray Crystallography*. Vol. III, p. 214. Birmingham: Kynoch Press.

THOMAS, L. H. & UMEDA, K. (1957). *J. Chem. Phys.* **26**, 293.