PHYSICO-CHEMICAL STUDIES OF SOME ORGANIC MOLECULES.

A thesis submitted to the University of Glasgow,

in fulfilment of the requirements for the

degree of

Doctor of Philosophy

by

Andrew L. Porte B.Sc..

The University of Glasgow.   March, 1957.

ProQuest Number: 13849065

ProQuest 13849065

P R E F A C E

CONTENTS.

Page

SUMMARY

This thesis describes investigations of the molecular structures of a number of organic compounds using physico-chemical methods.   Two main methods are employed.   These are

1) Methods of X-ray structural analysis.

2) Theoretical methods (molecular orbital calulations).

Three compounds are considered under the first heading.   These are

a) Tiglic Acid   b) Angelic Acid

c) Tetrabenzo-2:3-6:7-2':3'-6':7'-heptafulvalene.

Tiglic and angelic acids are cis-trans isomers which were known from chemical evidence to have the formula $CH_3 . CH = C . COOH$ (CH_3). However, chemical methods of deciding which acid is the cis and which is the trans isomer were based on results which, by themselves, were of doubtful value although when considered collectively indicated the trans structure for tiglic acid and the cis structure for angelic acid.

X-ray analyses of both acids are described in the thesis. A novel method of distinguishing between cis and trans isomers by means of molecular Fourier transforms, which may be applicable in the case of planar molecules is described.

Unit cell and space group data are given for both compounds and a complete and accurate analysis of tiglic acid is described.   This analysis shows tiglic acid to be the trans acid I.

$$H_3C - CH$$
$$\|$$
$$H_3C - C - COOH.$$

I.

$$HC - CH_3$$
$$\|$$
$$H_3C - C - COOH$$
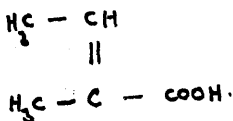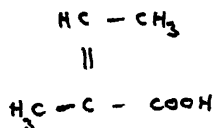
II.

thereby confirming the formula usually accepted — although
not proved — for this acid.

Owing to various experimental difficulties the analysis of
the angelic acid structure is not as accurate as that of tiglic
acid, but the work confirms the formula II for the acid.   The
complete determination of the structure of angelic acid has
been hindered somewhat by unusual difficulties involving
homometric projections.   Work is still proceeding on the compound
and it is believed that the described solution of the problem
is essentially correct.   Three dimensional analysis at low
temperatures may be necessary to achieve high accuracy.

The thesis also describes some work carried out on
Tetrabenzo-2:3-6:7-2':3'-6':7'-heptafulvalene (III).   If a
planar model of this molecule is built up using aromatic bond
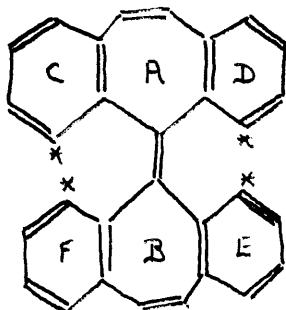
III

lengths of 1.4A., and assuming
rings A and B to be regular
heptagons and rings C, D, E and
F to be regular hexagons, then
it is found that the starred
carbon atoms approach to within
about 1.3A. of one another.
Such a close distance of approach
is impossible and so the atoms of the molecule cannot be coplanar.
The object of the work on this compound was to find the actual
shape of the molecule.   Unit cell and space group data for
III are described in the thesis and a discussion on the probable
conformation/

conformation is given.   This part of the work is not yet
finished but the structure postulated is chemically reasonable
and fits in with a number of experimental observations.

The second type of method used in the investigations was the
application of molecular orbital theory.   By this means,
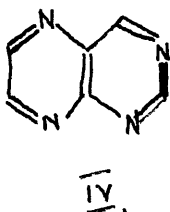
π-electron distributions, bond orders, and
bond lengths for pteridine(IV) have been
derived.   A ~~neew~~ new bond order — bond
length curve for carbon—nitrogen bonds
is proposed.

IV.

THE CRYSTAL AND MOLECULAR STRUCTURES

OF

TIGLIC AND ANGELIC ACIDS.
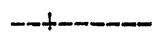
NOTE ON NOMENCLATURE

For the purpose of clarifying configurations, it will be convenient to use an adaption of the rules suggested by the I.U.P.A.C. Commission of Nomenclature of Organic Chemistry (1) and now generally adopted.   In these rules, in the case of 2-methyl-2-alkenoic acids the prefix cis- or trans- is used to describe the configuration of the unsubstituted alkenoic acid chain.   Thus the name 2-methyl-trans-2-butenic acid is obtained for I, while the acid II will be 2-methyl-cis-2-butenic acid.

$$CH_3 - CH$$
$$\| $$
$$CH_3 - C - COOH$$

I

2-methyl-trans-2-butenic acid.

$$CH_3 - CH$$
$$\|$$
$$HOOC - C - CH_3$$

II

2-methyl-cis-2-butenic acid.
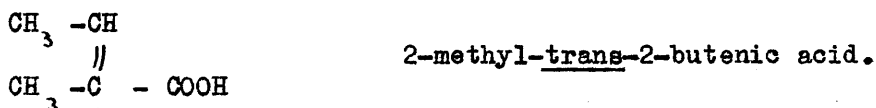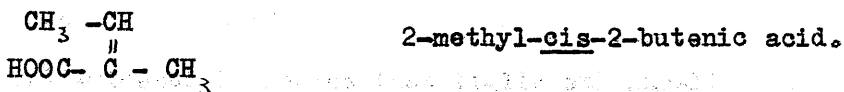
As a result of the identification of the biologically active
fraction of the fatty acids from the tuberculosis bacillus with
C_{27}-phthienoic acid(2) and the assignment of a structure including
the features of a 2-methyl-2-alkenoic acid(3,4) R.CH$_2$.CH=C.COOH

$\overset{\text{CH}_3}{}$

there has followed an extensive study of the methods
of preparation and properties of such unsaturated acids(5,6,7,8).
However, in spite of considerable synthetic activity in this field,
there has not yet developed a reliable method for assigning
geometrical configurations to either the synthetic 2-methyl-2-alkenoic
acids or to the naturally occurring C_{27}-phthienoic acid.

In two recent papers(7,9) assignments of geometic configuration
have been based primarily on correlation of infra-red and
ultra-violet spectral data with those of the simplest members
of the series – tiglic and angelic acids (R = H). However,
the geometric configuration to be assigned to tiglic and angelic
acids themselves were by no means conclusive, and it was with the
object of removing the uncertainties in the configurations
of these acids that this work was undertaken.

Chemical evidence indicates that tiglic and angelic acids are
2-methyl-2-butenic acids. Tiglic acid is believed to be the
trans- acid and angelic acid is believed to be the cis- acid thus

        CH$_3$ – CH   HC – CH$_3$

             ‖   ‖

        CH$_3$ – C – COOH   H$_3$C – C – COOH

           Tiglic Acid   Angelic Acid

4.
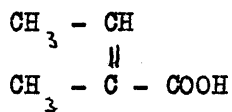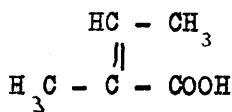
The main evidence on which these configurations are assigned is based mainly on the work of Sudborough and Davies (10) on esterification of the acids with methanol and hydrochloric acid, on the work of Von Auwers and Wissebach (11) on the configurations of the crotonic acids and on the work of Pfeiffer(12) on the dibromo derivatives of tiglic and angelic acids.

Sudborough and Davies showed that angelic acid is esterified at a slower rate than tiglic acid. Consequently, they said, angelic acid must have more groups blocking the carboxyl group than has tiglic acid. Hence angelic acid is the cis- acid.

Von Auwers and Wissebach established the configurations of the crotonic acids by relating them via the intermediate γ,γ,γ-trichlorocrotonic acids to maleic and fumaric acids. They thus showed that solid crotonic acid is the trans- form and isocrotonic acid is the cis- form. By comparing the physical properties of this pair of isomers with the same properties of the pair of isomers, tiglic and angelic acids, they then came to the conclusion that tiglic acid is the 2-methyl-trans-crotonic acid and angelic acid is the 2-methyl-cis-crotonic acid thus.

```
H C - CH          H - C - CH
 3   ||                   ||   3
H - C - COOH      H - C - COOH
```
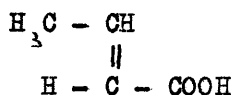
Crotonic Acid   Isocrotonic Acid

m.pt. 72°C; pK$_a$= 4.70   m.pt. 15.5°C; pK$_a$= 4.44

Heat of combustion   Heat of combustion

= 478 Kg.cals/mole.   = 486 Kg.cals/mole.

```
H₃C - C - H              H - C - CH₃
      ‖                        ‖
H₃C - C - COOH           H₃C - C - COOH
```
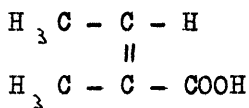
    __Tiglic Acid__   __Angelic Acid__

m.pt. 64.5°C;   pK$_a$ = 5.02   m.pt. 45°C;   pK$_a$ = 4.30.

Heat of combustion   Heat of combustion

  = 627 Kg.cals/mole.   = 635 Kg.cals/mole.

     Pfeiffer studied the elimination of hydrogen bromide from
a number of bromo-substituted elefins and came to the conclusion
that __trans__ elimination occurred more readily than __cis__ elimination.
On this basis, he assigned the __trans-methyl__ configuration to the
2-bromo-2-butene which boiled at 87 - 88°C., since it gave
crotonylene more readily than the 2-bromo-2-butene which boiled at
93 - 94°C..

```
H₃C - C- H        KOH        H₃C - C         H - C - CH₃
      ‖          ⟶                ‖‖              ‖‖
Br — C - CH₃      ⟵          C - CH₃         Br- C - CH₃
                 HBr
```
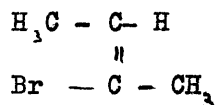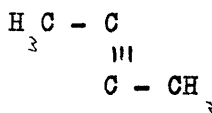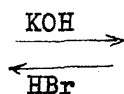
   __2-bromo-__   Crotonylene   __2-bromo-__
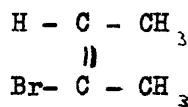
__trans-2-butene__   __cis-2-butene.__

    b.pt. = 87 - 88°C.   b.pt. = 93 - 94°C.

Having established the configurations of the 2-bromo-2-butenes in
this manner, he then went on to deduce the configurations of tiglic
and angelic acids as follows:

```
                        a    CH₃
                              ↕
H₃C - CH         Br₂  b  HC - Br      Na₂CO₃    H₃C - C - Br     +CO₂ +NaBr
      ‖          ⟶          |          ⟶              ‖
H₃C - C - COOH        c  Br- C - COOH          H - C - CH₃
                              |
                        d    CH₃
```

  __Tiglic Acid__   I   II

H – C – CH$_3$ Br$_2$ a) Br – C – H   Na$_2$CO$_3$ H$_3$C – C–Br
‖   ⟶   b)   CH$_3$ ⟶   ‖   + CO$_2$ + NaBr
H$_3$C – C – COOH   c) Br – C – COOH   H$_3$C – C – H
d)   CH$_3$
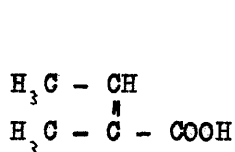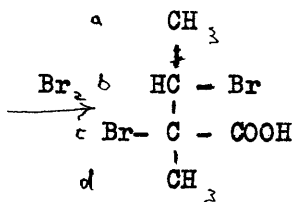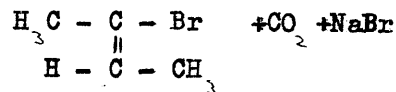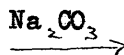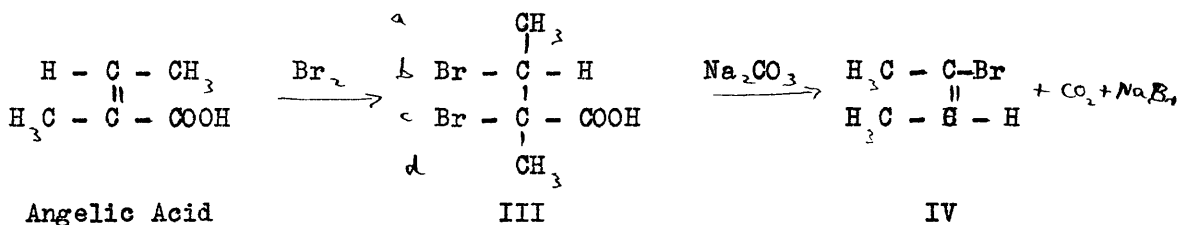
Angelic Acid   III   IV

Tiglic and angelic acids on treatment with bromine yield different
racemic dibromo derivatives. On treatment of these with sodium
carbonate CO_2 and $\overline{Br}$ ion are eliminated to yield the
2-bromo-2-butenes of known configuration. Tiglic acid gives
2-bromo-<u>trans</u>-2-butene(II) whilst angelic acid gives
2-bromo-<u>cis</u>-2-butene(IV). Pfeiffer assumed that the elimination
of the CO_2 and $\overline{Br}$ ion from the dibromo-acids could only take place
if the bromine atom and the carboxyl group concerned were in a
trans-conformation. If this were so, then tiglic acid and angelic
acid must have the configurations shown. Pfeiffer's ideas are
easily followed if the dibromo-acids are viewed along the $C_b - C_c$
bonds. Along this view the reactions become

H$_3$C – C – H   Br$_2$
‖   ⟶
H$_3$C – C – COOH
Tiglic Acid

$\overline{I}$.

Na$_2$CO$_3$

$\overline{II}$

+ CO$_2$ + NaBr.
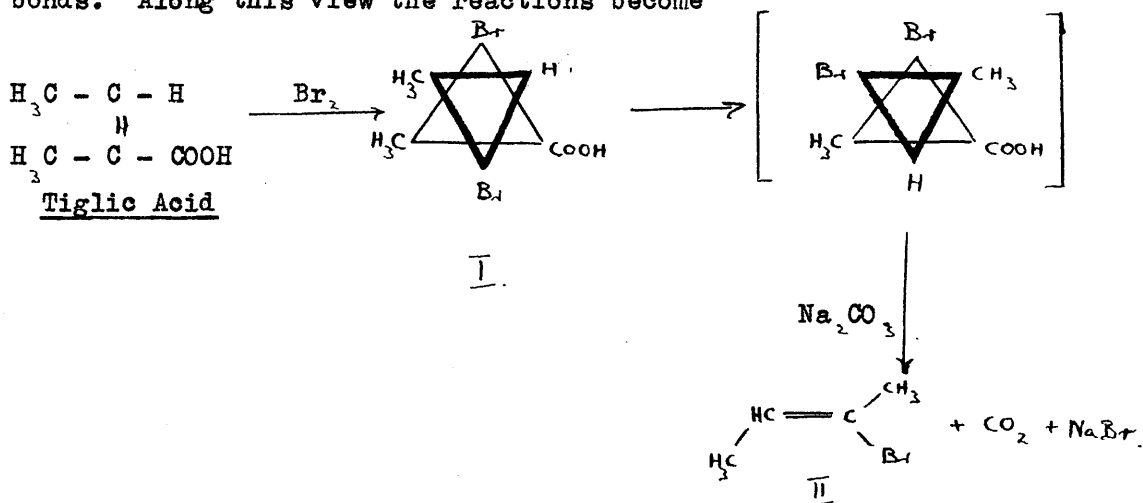
i.e. Bromine adds on <u>trans</u> across the double bond of tiglic acid
to give the dibromotiglic acid (I). For elimination of CO_2 and $\overline{Br}$
ion a trans conformation of –Br and –COOH is required so rotation
about the bond C –C takes place. The –CH$_3$ groups are now in a
<u>trans</u> conformation also and so on elimination of CO_2 and $\overline{Br}$

ion 2-bromo-trans-2-butene (II) is obtained.

The similar reaction in the case of angelic acid is

angelic acid.

III

$$Br - C - CH_3$$
$$\parallel \qquad\qquad + CO_2 + NaBr$$
$$H - C - CH_3$$

IV

These are the main lines of evidence on which the suggested
configurations for tiglic and angelic acids were based when this
work was started but more recently Dreiding and Pratt(13) have
obtained more conclusive chemical evidence from a study of the
carboxylation of trans-2-butenyl-lithium, a reaction which they
consider to afford a stereospecific synthesis of angelic acid.

2-bromo-trans-
2-butene.

Angelic Acid

None of these is by itself, conclusive evidence as to the
configurations of tiglic and angelic acids, although when taken
together they would indicate that tiglic acid is the trans- acid and
angelic acid is the cis- acid.   Nevertheless, a much more direct
determination of the configurations of these acids is clearly
desirable.

The methods of X-ray structure analysis afford an excellent means of solving such a problem since by this method the spacial positions of all the atoms in the molecules can be uniquely determined and hence the configurations obtained directly.

In particular it would be expected, at least to a first approximation, that these molecules would be planar and so the methods of the molecular Fourier transform could be employed. The problem would then reduce to finding out which reciprocal lattice fitted which transform.

<u>The Fourier Transforms of the two 2-methyl-2-butenic acids</u>(14, 15)

Experience shows that most organic acids occur as centrosymmetric dimers in the crystal, with hydrogen bonding between adjacent carboxyl groups.   If this is so, we can then say that the structure factor F(hkl) is given by

$$F(hkl) = 2\sum_{j=1}^{n} f_j \cos 2\pi(hx_j + ky_j + lz_j)$$

Here h, k and l are indices representing points on a reciprocal lattice and x_j, y_j and z_j are scaled coordinates(with centre of symmetry as origin) representing the position of the atom j in the unit cell. f_j is the atomic scattering factor of atom j for the reflection (hkl) and the summation is taken over half the number of atoms of the unit cell.   The structure factor F(hkl) is a discontinuous function and when solving a crystal structure the whole problem degenerates to finding both the amplitude and the phase angle of the structure factors for each of the points of the reciprocal lattice.

Suppose now we take any centrosymmetric constellation of atoms.   Choose an origin at the centre of symmetry through which pass any three non-coincident axes defined by the unit vectors $\underline{\xi}, \underline{\eta}, \underline{\zeta}$ and consider the function.

$$G(X^*Y^*Z^*) = 2\sum_{j=1}^{n} f_j \cos 2\pi(\xi_j X^* + \eta_j Y^* + \zeta_j Z^*)$$

where ξ_j, η_j, ζ_j are coordinates of the atom j referred to the set of axes $\underline{\xi}, \underline{\eta}, \underline{\zeta}$ and X^*, Y^* and Z^* are points on a set of axes reciprocal to the frame $\underline{\xi}, \underline{\eta}, \underline{\zeta}$.   Unlike h,k and l, the values of X^*, Y^* and Z^* are not restricted to integers.

$G(X^*Y^*Z^*)$ is the Fourier transform of the given constellation of atoms. It will be obvious that although $F(hkl)$ is a discontinuous function, $G(X^*Y^*Z^*)$ is continuous and in fact $F(hkl)$ can be obtained by placing $F(hkl)$ and $G(X^*Y^*Z^*)$ on the same scale, origin to origin, and sampling $G(X^*Y^*Z^*)$ at the points of the reciprocal lattice. It should be noted that the axes of the reciprocal unit cell of the crystal and the axes of the function $G(X^*Y^*Z^*)$ need not necessarily be the same so that the reciprocal unit cell of the crystal and the function $G(X^*Y^*Z^*)$ may be tilted during the sampling process.

It would be an extremely laborious task to work out the three-dimensional functions $G(X^*Y^*Z^*)$ over the whole of reciprocal space but the problem is very much simplified when the atoms of the constellation concerned are coplanar. When this is the case, take $\underline{\xi}$ and $\underline{\eta}$ in the plane of the constellation and $\underline{\zeta}$ at right angles to it. All values of ζ_j are now zero and the transform becomes

$$G(X^*Y^*Z^*) = 2 \sum_{j=1}^{\tilde{}} f_j \cos 2\pi (\xi_j X^* + \eta_j Y^*)$$

i.e. the function now consists of "cylinders" of contours with constant sections since its value is independent of Z^*.

If all the atoms are assumed to be the same, say carbon atoms, then we can write

$$\frac{G(X^*Y^*Z^*)}{f} = T(X^*Y^*Z^*) = 2 \sum_{j=1}^{\tilde{}} \cos 2\pi (\xi_j X^* + \eta_j Y^*)$$

or, if the atoms are not the same but of nearly the same atomic number

$$\frac{G(X^*Y^*Z^*)}{f} = T(X^*Y^*Z^*) = 2 \sum_{j=1}^{\tilde{}} w_j \cos 2\pi (\xi_j X^* + \eta_j Y^*)$$

where f is an atomic scattering factor for carbon
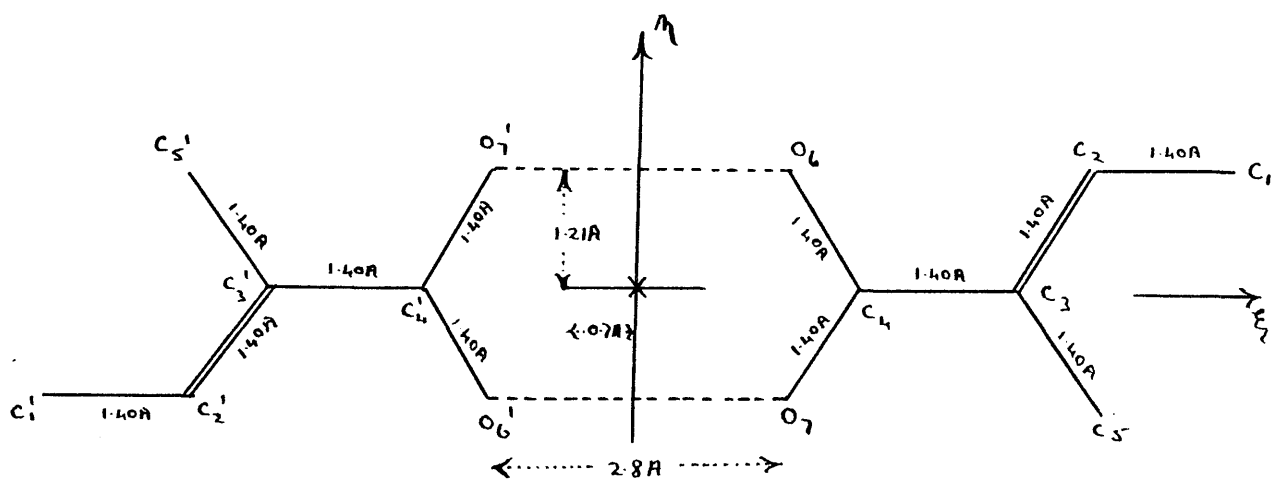
Idealised Dimer of 2-methyl-trans-2-butenic acid   (I).
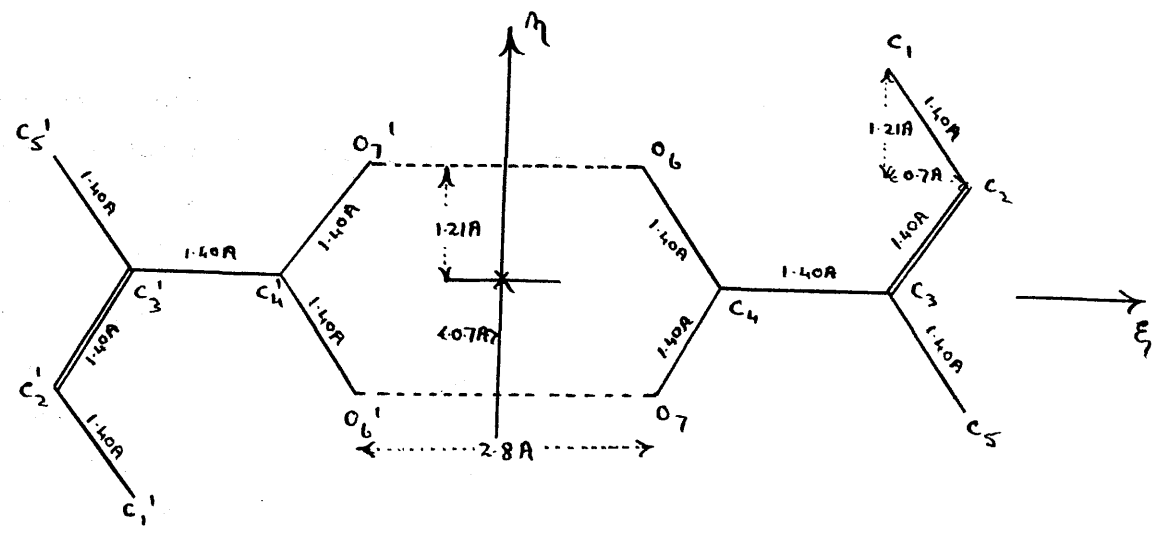
Idealised Dimer of 2-methyl-cis-2-butenic acid   (II).

Figure I.

w_j is a weighting factor whose value is 1 for j = carbon and whose value is $\frac{N_j}{6}$ for any other atom of atomic number N_j. i.e. instead of comparing the observed values of $F(hkl)$ with $G(X^*Y^*Z^*)$, in practice the observed value of $\frac{F(hkl)}{f}$ is compared with the calculated function $T(X^*Y^*Z^*)$.

Even with this simplification the calculation would be very laborious and so a further simplification is introduced. The way in which this is done is shown for the examples actually used here.

. We are interested in calculating the Fourier transforms of the two 2-methyl-2-butenic acids. If instead of calculating the transforms of the actual acids we calculate the transforms of the two idealised constellations of atoms shown, then a great simplification in labour will result. (See Figure I)

It will be seen that in these idealised constellations all coordinates can be expressed as multiples of $0.7A$ along the ξ - direction and $1.21A$ along the η - direction. Hence it follows that the function

$$T(X^*Y^*Z^*) = 2 \sum_{j=1}^{m} w_j \cos 2\pi (\xi_j X^* + \eta_j Y^*)$$

will be periodic with period $\frac{1}{0.7} A^{-1}$ along the X^* direction and period $\frac{1}{1.21}A^{-1}$ along the Y^* direction, and can easily be evaluated by means of Beevers-Lipson strips in a manner analogous to that used for estimating electron densities by the normal method of Fourier analysis.

The problem of deciding which of the two 2-methyl-2-butenic acids is tiglic acid and which angelic acid then degenerates to finding out whether the reciprocal lattice of the acid concerned fits the $T(X^*Y^*Z^*)$ function of I or II.

The first acid to be dealt with is tiglic acid.

THE CRYSTAL AND MOLECULAR STRUCTURE

OF

TIGLIC ACID

The Crystal and Molecular Structure of Tiglic Acid.

1. Crystal Data

Tiglic Acid, $C_5H_8O_2$;   Molecular weight = 100.1; m.pt. = 64.5°C.
Density calculated = 1.162 gms/cc..
Density observed (by flotation in potassium iodide solution)
         = 1.163 gms/cc..

$a = 7.70 \pm 0.02A$ $b = 5.23 \pm 0.02A$ $c = 7.42 \pm 0.02A$
$\alpha = 96.6° \pm 0.5°$ $\beta = 86.5° \pm 0.5°$ $\gamma = 106° \pm 0.5°$

No systematic absences.   Space group is P1 or $\overline{P1}$.
Two molecules per unit cell.   No molecular symmetry required.
Volume of unit cell(V)
 $= abc(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)^{\frac{1}{2}}$
 $= 285.6A^3$.
Absorption coefficient for X-rays($\lambda = 1.542A$) $\mu = 8.56cm^{-1}$.
Total number of electrons per unit cell = F(ooo) = 108.
The crystals were obtained by slow recrystallation from 60-80
petroleum-ether. This resulted in thick needles developed along
the b- axis.
The morphology has been described by Haushofer(16) who gives
        $a : b : c = 1.4927 : 1 : 1.5493$.
        $\alpha = 99°1'$; $\beta = 137°54'$; $\gamma = 85°39'$
No correlation between the two cells has been found.

2. Experimental Measurements

Tiglic acid is extremely volatile and so some method of protecting the crystal whilst the X-ray photographs are being taken is necessary. This was done by encapsulating the crystal on its glass fibre with a Park-Davis number 0 gelatine capsule and adding extra crystals of tiglic acid under the capsule to build up a vapour pressure. Rotation, oscillation and moving-film photographic measurements were employed using copper K_α radiation ($\lambda = 1.542A$). The cell dimensions were determined by rotation and equatorial layer line moving-film photographs. The angles α, β and γ of the unit cell were obtained by two methods.

a). Directly, by measuring the unit translations along diagonals of the unit cell and thence employing the simple triangulation formulae of the type

$$[1\overline{1}0]^2 = a^2 + b^2 - 2ab\cos\gamma$$

b). Indirectly, by means of reciprocal angles $\alpha*$, $\beta*$, $\gamma*$ obtained from equatorial layer line moving-film photographs and thence employing formulae of the type

$$\cos\alpha = \frac{\cos\beta* \cos\gamma* - \cos\alpha*}{\sin\beta* \sin\gamma*}$$

The mean values of α, β and γ obtained by these methods have been quoted.

The reflexions were estimated visually by means of the multiple film technique(17). The structure factors evaluated by the usual mosaic crystal formula from the estimated intensities are given in table VI(pp. 44 -49).

3. Structure Determination.

In the solid state, most organic acids exist as centro-symmetric dimers. Tiglic acid contains two molecules per unit cell and it was assumed that these would form dimers in the usual manner, i.e. the space group was assumed to be $P\bar{1}$, with the centre of inversion between the carboxyl groups of a symmetrical dimer. Hence the discussion already given on Fourier transforms applies to this problem.

Since the b-axis is the shortest of the unit cell translations it was expected that a projection of the structure down this axis would give the best resolution and so this projection was considered first.

109 planes out of a possible 151 were observed in the (hol) section of the reciprocal lattice and it was found that this reciprocal lattice net fitted the Fourier transform of I but a fit could not be obtained on the transform of II (fig.1. p. *ll*.).

It is obvious that, since the constellations of atoms, I and II, h ave much in common, their Fourier transforms will also have some common features. However, although parts of the (hol) reciprocal lattice net of tiglic acid fitted parts of the transform of II, serious discrepancies existed and there was no doubt that this reciprocal lattice net was a sample from the transform of I. Hence tiglic acid is 2-methyl-trans-2-butenic acid thus confirming the formula assigned to it on chemical grounds.

The orientation of the (hol) reciprocal lattice net on the transform of I is shown on fig.II.

Computation of the electron density projected on to a plane

Fig. II.
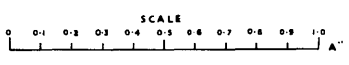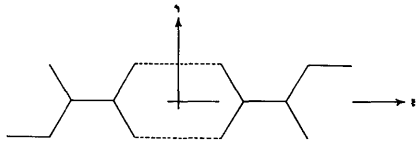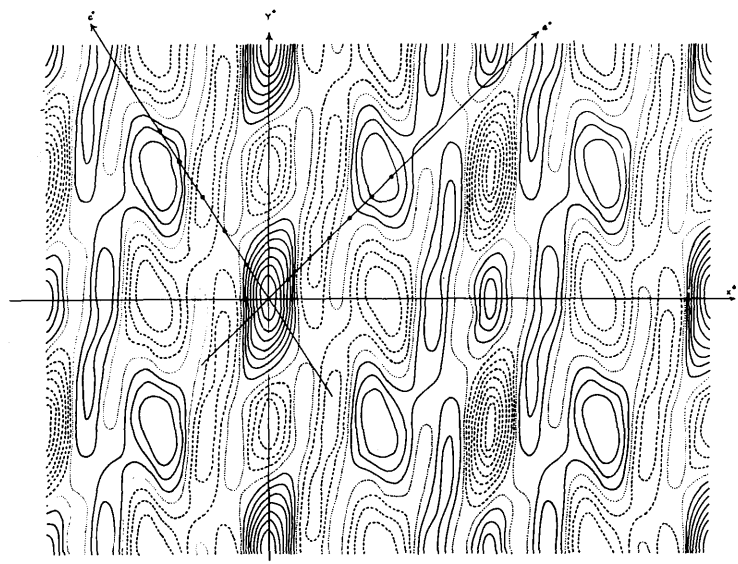
Fourier Transform of Idealised Dimer of
2-methyl - trans - 2-butenic acid
with the a* and c* reciprocal axes
of Tiglic acid superimposed.

perpendicular to the b-axis by summation of the Fourier series

$$\rho(xoz) = \frac{1}{A} \sum_{h,l=-\infty}^{+\infty} F(hol)\cos 2\pi(hx + lz)$$

using the phase angles derived from the transform and the observed structure factor amplitudes, resulted in an electron density map which showed all of the atoms to be clearly resolved. Structure factors calculated using atomic coordinates obtained by picking out atomic centres from the electron density map resulted in a discrepancy

$$R = \frac{100 \sum \left| |F_o| - |F_c| \right|}{\sum |F_o|}$$

of 21% at this stage.

4. Refinement of the projection down the b-axis and location of the hydrogen atoms.

Further refinement of this projection of the structure was carried out by successive difference Fourier synthesis,

$$D(xoz) = \rho_o(xoz) - \rho_c(xoz)$$
$$= \frac{1}{A} \sum_{h,l=-\infty}^{+\infty} \left\{ F_o(hol) - F_c(hol) \right\} \cos 2\pi(hx + lz)$$

Difference methods have been used by several authors from time to time(e.g.18,19,20) and Booth(21), in a theoretical investigation of the function advocated its use. However, the full potentialities of the method were established by Cochran(22) only in 1951.

The corrections to the atomic coordinates may be derived as follows. The electron density distribution of an atom, to a first approximation,is Gaussian and may be represented by the equation

$$\rho_o(r) = \rho_o(o) \exp\left[-pr^2\right]$$

where $\rho_o(r)$ is the electron density at a distance r from the centre of the atom where the electron density is $\rho_o(o)$.

i.e. $\quad \rho_o(r) = \rho_o(o) \left\{ 1 - pr^2 + \dfrac{p^2 r^4}{2!} - \cdots \right\}$

$\qquad\qquad \approx \rho_o(o) \left\{ 1 - pr^2 \right\} \qquad$ when r is small.

If the origin is taken at the point assumed to be the atomic centre in calculating the structure factors then the calculated electron density at $\underline{r}$ will differ from the true electron density at this point by an amount D(r) given by

$$D(r) = \rho_o(r) - \rho_c(r).$$

where $\rho_o(r)$ is the observed electron density at $\underline{r}$, calculated by using the observed values of the structure factors and $\rho_c(r)$ is the electron density at $\underline{r}$ obtained by making use of the calculated values of the structure factors. The position of $\rho_o(o)$ will differ from that of $\rho_c(o)$ by a small amount Δ, where Δ is the required correction to the atomic centre. Hence we may now write

$$D(r) = \rho_o(r) - \rho_c(r)$$
$$\quad\; = \rho_o(o) \left\{ 1 - p(r - \Delta)^2 \right\} - \rho_c(o) \left\{ 1 - p'r^2 \right\}.$$

If the correct scattering curve has been used so that $\rho_o(o) = \rho_c(o)$ and $p = p'$ then

$$D(r) = 2\rho_o(o)pr\Delta$$

$$\therefore \quad \Delta = \dfrac{\left(\frac{dD}{dr} \right)_{r=o}}{2p \cdot \rho_o(o)}$$

The factor p depends upon the thermal motion of the atoms and is derived experimentally from the equation

$$\rho_o(r) = \rho_o(o) \exp\left[-pr^2 \right]$$

i.e. $\log_e \rho_o(r) = \log_e \rho_o(o) - pr^2$

or $\quad \log_e \rho_o(r) = - pr^2 + \text{constant.}$

Hence if $\log_e \rho_o(r)$ is plotted against r^2, a straight line graph should be obtained. The gradient of this line is $-p$, and and at $r = o$, $\log_e \rho_o(o)$ and hence $\rho_o(o)$ may be obtained.

The direction of shift of the atomic coordinates is along the line of steepest ascent and the magnitude is obtained by measuring the slope of $D(r)$ at the atomic site and dividing by $2p \cdot \rho_o(o)$.

The process of refinement will come to an end when the function D has zero slope at the atomic centres. The coordinates will then be free from termination of series errors except in so far as the scattering curves used to calculate the F_c's are incorrect.

The atomic scattering curves employed during the preliminary stages of the refinement process were of the form

$$ f_j = f_{j\,o} \, \exp\left\{ -B\left(\frac{\sin\theta}{\lambda}\right)^2 \right\} $$

where $f_{j\,o}$ is the theoretical curve for atom j (valence states) at rest given by McWeeny(23) and the Debye–Waller factor B was assumed to be 4.8×10^{-16} $\overset{2}{\underset{cms}{\text{Å}}}$ for all atoms.

The first difference syntheses showed that the coordinates of several atoms required to be adjusted but when these atomic shifts were allowed for, it soon became obvious that the Debye–Waller factors for all the atoms were not the same, so the value of B for each atom was adjusted until the difference between the observed and calculated electron densities($\rho_o - \rho_c$) at each atom was as small as possible. The method of adjusting the B values was that employed by Cochran(22) who defined a quantity $\alpha_j = \dfrac{B}{\lambda^2}$ and showed that on the average a small change $\Delta\alpha_j$ in α_j of 0.1 for carbon(or oxygen) atoms results in a change in $(\rho_o - \rho_c)$ of about
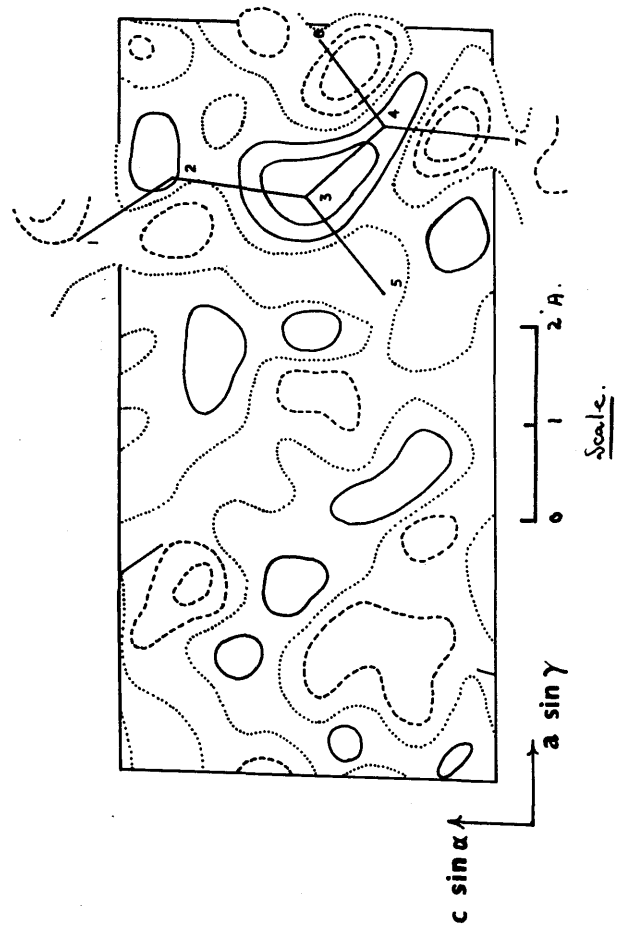
21.

Fig. III a.

Tiglic acid : b-axis projection.

Difference between the electron density in the crystal and that calculated for carbon, oxygen and hydrogen atoms.

Contours at intervals of $0.2 e/Å^2$. Zero contour dotted.
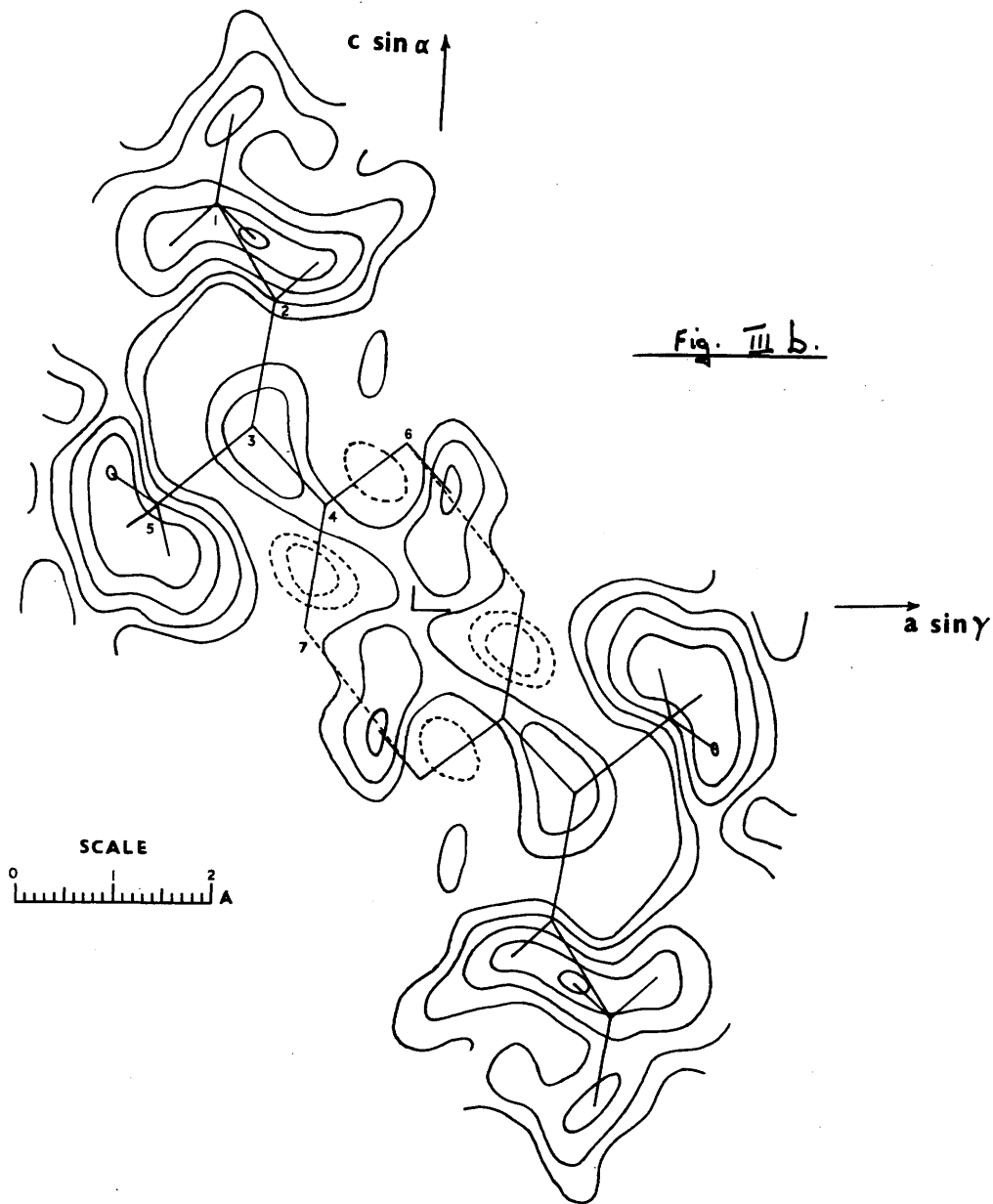
c sin α

a sin γ

Scale.

0   1   2 Å.

Fig. III b.

Tiglic Acid - b-axis projection.

Difference between the electron density in the crystal and that calculated for carbon and oxygen atoms.

Contours at intervals of $0.2 e/A^2$. Zero contour dotted.

0.35 e/$\overset{2}{\text{A}}$ at the centre of the atom.

Application of the methods outlined above to successive difference syntheses resulted in a final discrepancy R for the (hol) net of the reciprocal lattice of 10.8%.
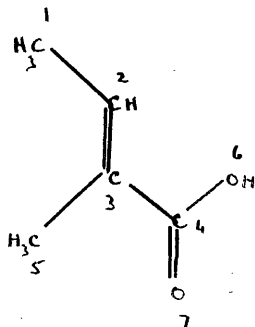
At this stage, there still remained small difference peaks on atoms C_3 and C_4 but since the main object of refining the structure this far was to obtain accurate coordinates for the carbon and oxygen atoms and to detect the hydrogen atoms, neither of which is affected by these small peaks, they were not removed. The values of B obtained in this way were

Atom	C_1	C_2	C_3	C_4	C_5	O_6	O_7
B	5.1	4.8	4.2	4.4	4.8	5.1	5.1

The final ($\rho_o - \rho_c$) maps are shown in figs. IIIa and IIIb. IIIb represents the difference between the electron density existing in the crystal and that calculated for carbon and oxygen atoms only. The significant peaks are due to the hydrogen atoms. There is very little resolution of the individual atoms on this map but although it is not possible to say exactly where a hydrogen atom is situated it is possible to say how many such atoms are in a given neighbourhood. The electron density of the hydrogen atom of the carboxyl group which is responsible for the hydrogen bonding between the molecules of the dimer is very much smeared out, as would be expected since the -O-H bond has a high electrovalent character, but the map indicates clearly a peak of 0.6e much nearer to O_6 than to O_7 i.e. the hydrogen bond is asymmetric and the hydrogen atom concerned is bonded to O_6 .
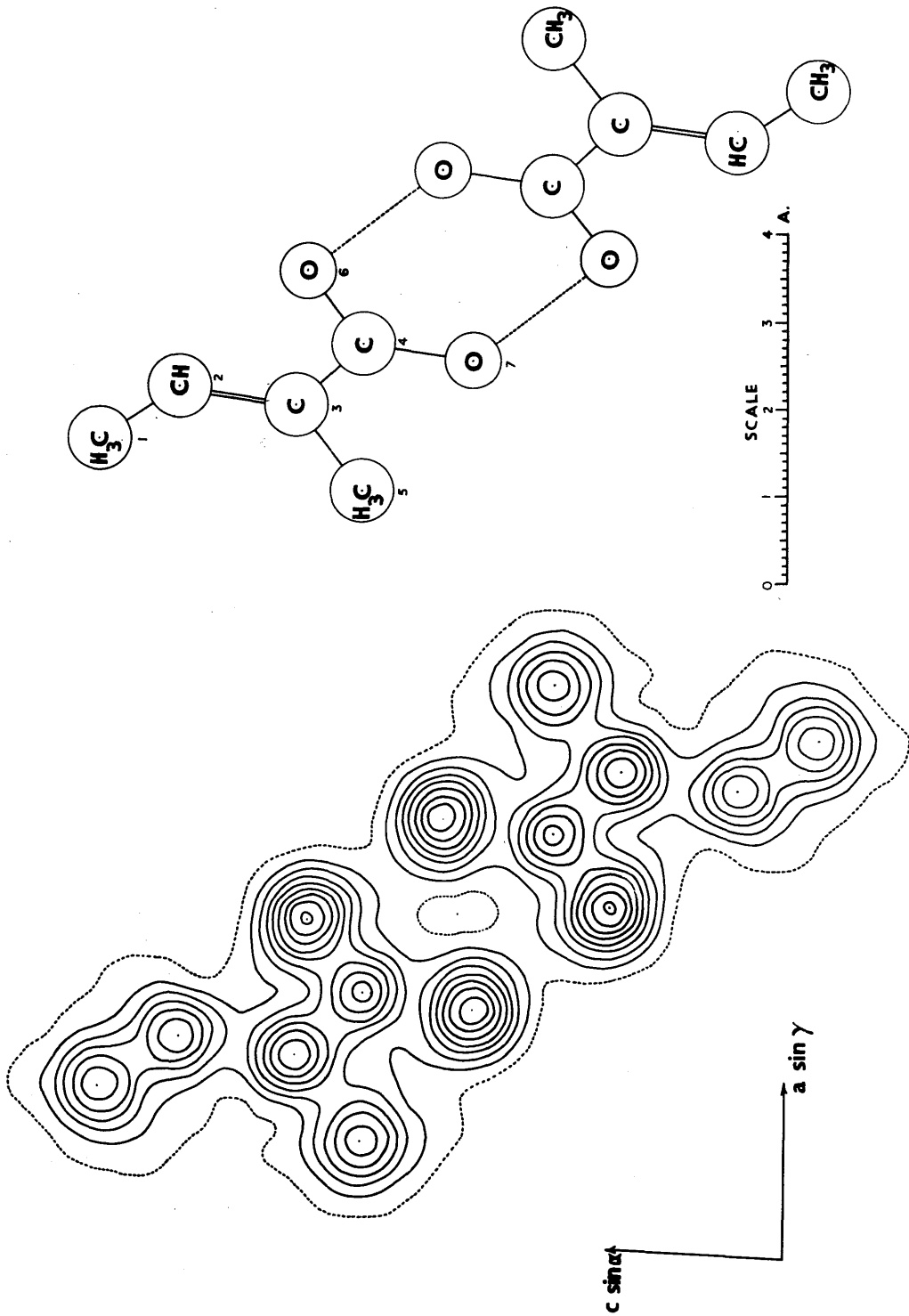
24.

Fig. IV.

Tiglic acid – b-axis projection. Final electron density map.
Contours at intervals of $1e/Å^2$. $1e/Å^2$ contour dotted.

The final electron density map of the projection of the structure on to the plane perpendicular to the b-axis is shown in fig. **IV.** This
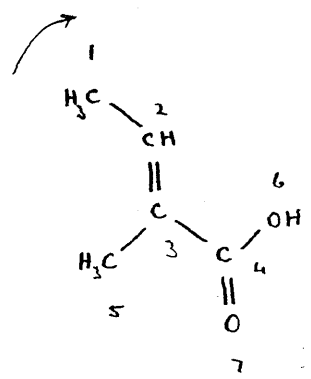
in addition to showing all the atoms in the molecule to be clearly resolved, shows evidence of repulsion between the methyl groups attached to C_2 and C_3. In the projection, bonds $C_2 = C_3$ and $C_4 = O_7$ are parallel to one another. So also are $C_5 - C_3$ and $C_4 - O_6$. However, bonds $C_1 - C_2$ and $C_3 - C_4$ are not parallel in projection. Hence they cannot be parallel in space, i.e. repulsion of the methyl groups has resulted in distortion of the sp^2 bond angles of the idealised formula. This point will be dealt with in detail again.

5. Projection of the structure on to the plane perpendicular to the a-axis of the triclinic cell.

The method of dealing with this projection was the same as that already employed for the b-axis projection. 73(okl) reflections out of a total of 97 possible were observed and estimated. Phase angles for these were obtained from the transform and combined with the observed structure amplitudes to give a representation of the electron density in the crystal projected on to the plane perpendicular to the a-axis. Atomic coordinates were picked out from the resultant electron density map, structure factors calculated, and the structure refined in the usual way by successive difference syntheses. After four cycles of refinement the discrepancy R, fell to 13.2% and stopped there. Again, McWeeny atomic scattering factor curves corrected for temperature vibration were used and the value of B employed for each atom was

Atom	C_1	C_2	C_3	C_4	C_5	O_6	O_7
B	6.2	5.1	4.0	3.8	5.8	5.1	5.1

Although the temperature factors for this projection were not investigated as thoroughly as in the case of the projection down the b-axis, they are of the correct order of magnitude and when combined with the values found in the first projection support the idea that in this structure, the molecule as a unit is vibrating randomly about its centre of mass. Superimposed on top of this main vibration there is a tendency for each atom to execute random vibrations of its own, the amplitude of vibration being greater the fewer the number of carbon(or oxygen) atoms to which the atom
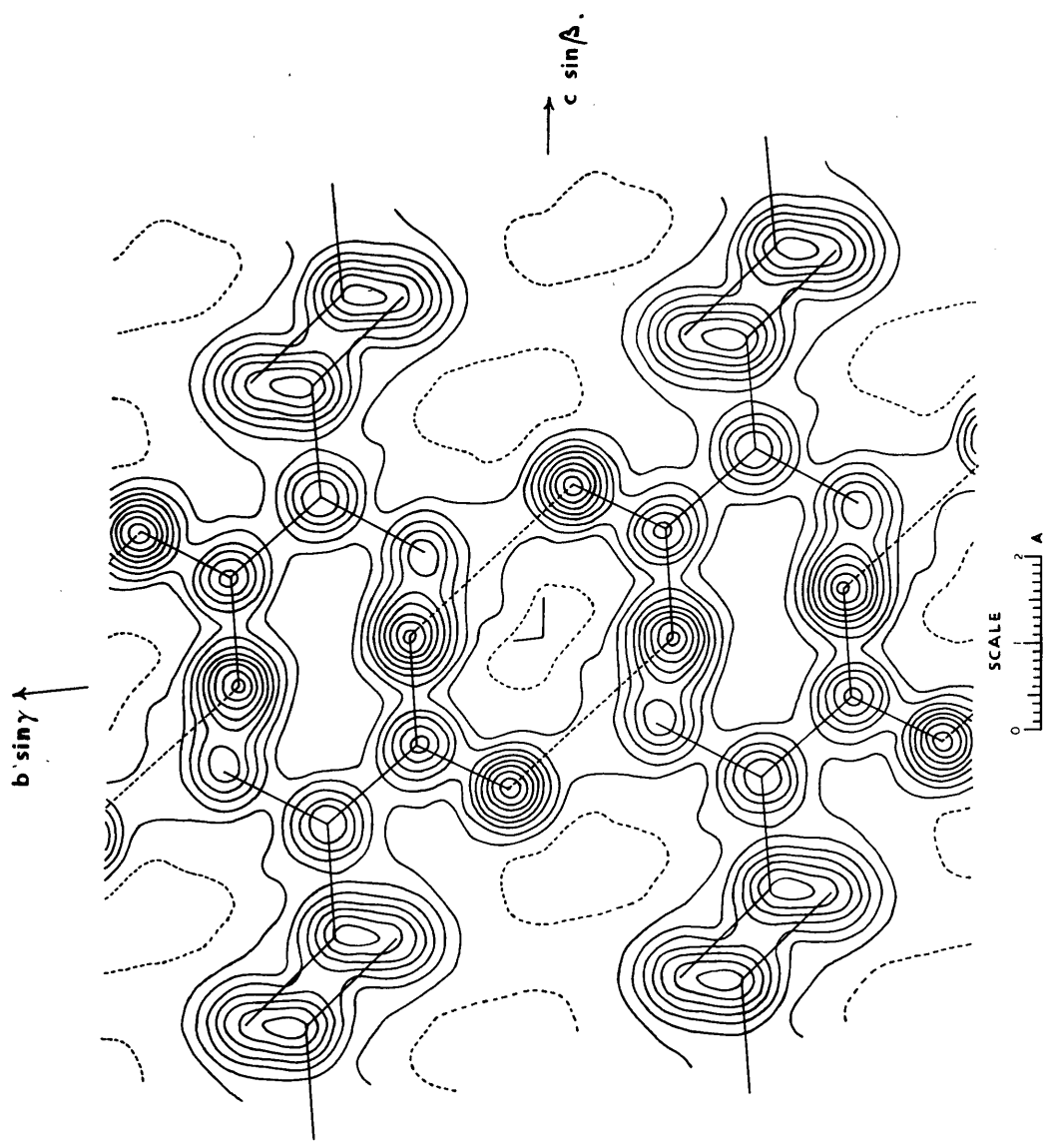
b'sin γ

c sin β.

SCALE

0   1   2 Å

Tiglic Acid — a -axis projection. Final electron density map.
Contours at intervals of $1e/Å^2$. $1e/Å^2$ contour dotted.

concerned is bonded.

The final electron density map of this projection is shown in fig. V.

6.   Projection of the structure on to a plane perpendicular to the c-axis of the triclinic cell.

As a check on the results already obtained, coordinates from the a and b-axes projections were combined to give coordinates for a projection of the structure on to a plane perpendicular to the c-axis. These were then used to calculate structure factors and the structure factors thus obtained were employed to compute the electron density projected on to the plane under consideration.   The atomic scattering factors used were of the form

$$f_j = f_{j_o} \exp \left\{ -5.7 \left(\frac{\sin \theta}{\lambda} \right)^2 \right\}$$

for all atoms.   This resulted in a discrepancy R = 15.9% over 76 planes observed and estimated out of a possible 109.

The electron density map computed in this way (fig.VI) shows no resolution of the individual atoms but demonstrates quite clearly the arrangement of the molecules in the unit cell.   In particular, it shows clearly the existence of a very pronounced cleavage plane, the $(2\overline{1}0)$ plane of the unit cell.

The observed and calculated structure factors are tabulated on pp. 44-49.

29.

a sin β

b sin α

SCALE
0 ——————— 2 Å
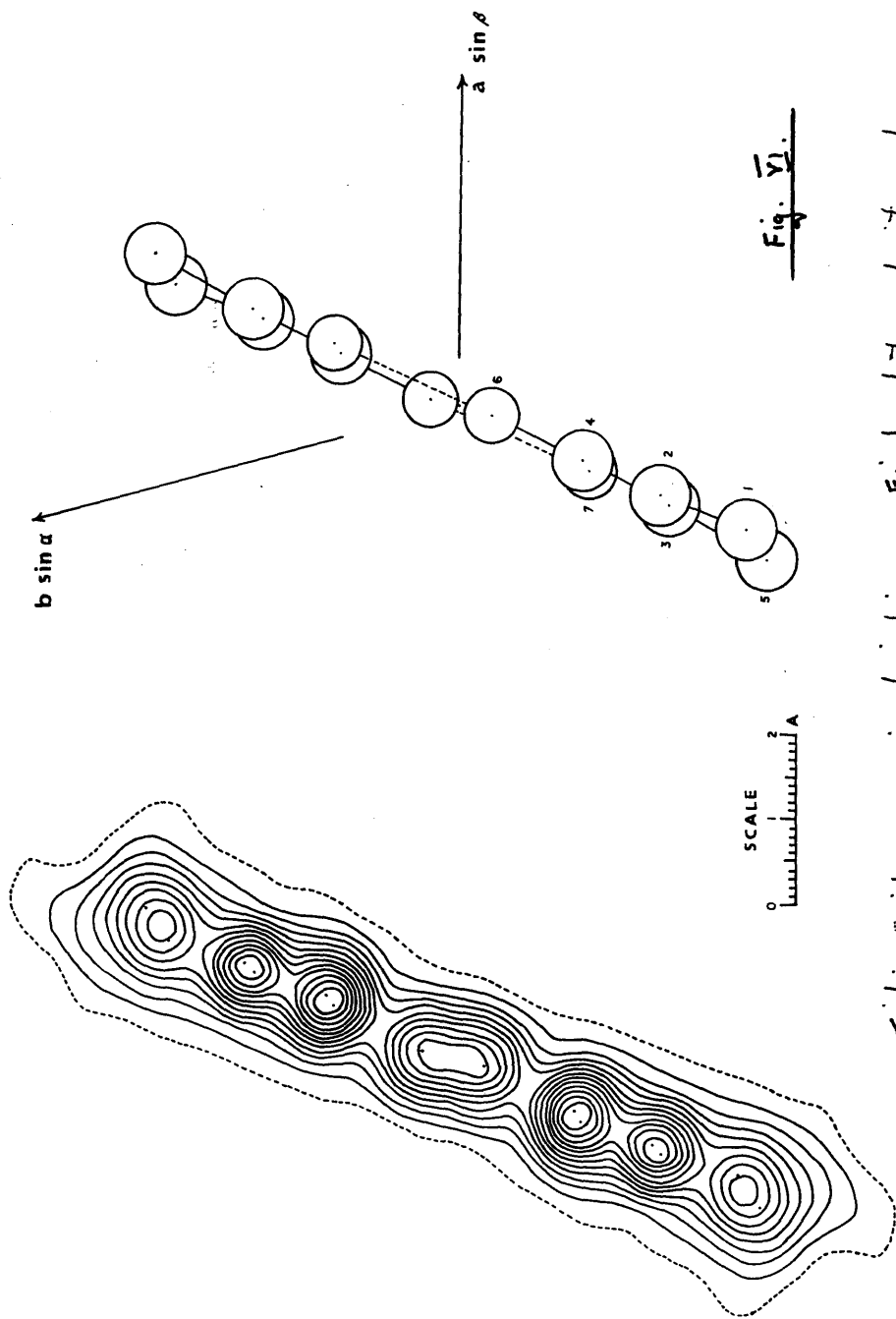
Fig. Ⅵ.

Tiglic acid – c – axis projection. Final electron density map.
Contours at intervals of 1 e/Å². 1 e/Å². contour dotted.

7.   Coordinates, Molecular Dimensions and Orientation of the molecule in the unit cell.

The final coordinates of the carbon and oxygen atoms are shown in table 1. x_P, y_P, z_P are fractional coordinates of an atom P and X_P, Y_P, Z_P are corresponding coordinates in Angstrom units referred to the oblique frame $\underline{a}$, $\underline{b}$, $\underline{c}$ of the triclinic unit cell.   There are two independent estimates of Z_P (from projections down the $\underline{a}$ and $\underline{b}$ axes respectively).   The average deviation from the mean of the two independent determinations is 0.008A and can be taken as a measure of the accuracy to be expected of the atomic coordinates.   The mean value of Z_P is quoted in table 1.

In order to make use of the usual formulae of 3-dimensional coordinate geometry it was necessary to transform (X_P, Y_P, Z_P) into rectangular coordinates (X'_P, Y'_P, Z'_P).   This was done in the following manner.
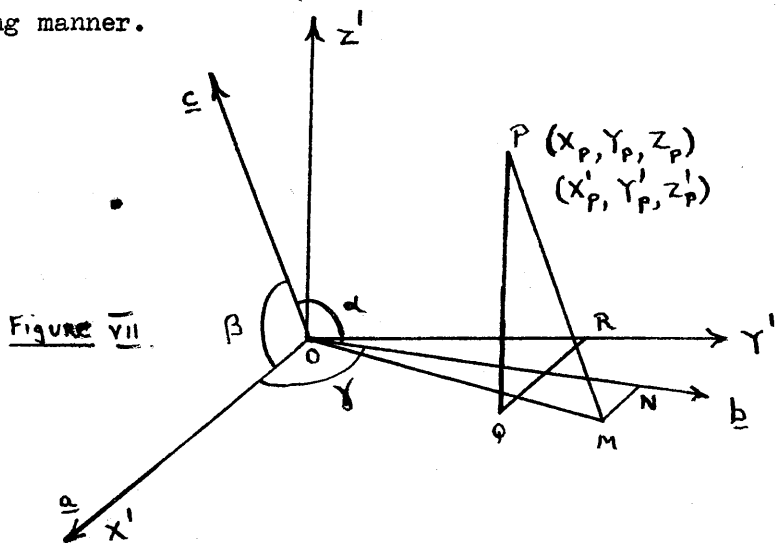
Figure VII

We are given a point P whose rectangular coordinates are (X'_P, Y'_P, Z'_P) and whose coordinates referred to oblique axes

$\underline{a}$, $\underline{b}$, $\underline{c}$ at angles α, β, γ to one another, as shown in the diagram, are (X_p, Y_p, Z_p). Oa and OX' are coincident and the aOb and X'OY' planes are identical.

$$QR = X'_p \quad ; \quad OR = Y'_p \quad ; \quad QP = Z'_p$$
$$MN = X_p \quad ; \quad ON = Y_p \quad ; \quad MP = Z_p$$

X_p, Y_p, Z_p are known, and X'_p, Y'_p, Z'_p are required in terms of these. This can be done quite easily by means of ordinary three dimensional geometry and the transformation formulae turn out to be

$$X'_p = X_p + Y_p \cos\gamma + Z_p \cos\beta$$

$$Y'_p = Y_p \sin\gamma + Z_p \left\{ \frac{\cos\alpha - \cos\beta\cos\gamma}{\sin\gamma} \right\}$$

$$Z'_p = Z_p \csc\gamma \left(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma \right)^{\frac{1}{2}}$$

Table 1. X_p, Y_p, Z_p, X'_p, Y'_p, Z'_p, are in Angstrom units.

Atom	x_p	y_p	z_p	X_p	Y_p	Z_p	X'_p	Y'_p	Z'_p
C_1	0.7000	0.3254	0.5557	5.3900	1.7018	4.1233	5.1725	1.2153	4.0957
C_2	0.7863	0.5333	0.4298	6.0545	2.7892	3.1891	5.4803	2.3560	3.1677
C_3	0.7667	0.5095	0.2523	5.9036	2.6647	1.8721	5.2834	2.3706	1.8596
C_4	0.8707	0.7155	0.1450	6.7044	3.7421	1.0759	5.7387	3.4876	1.0687
C_5	0.6440	0.2710	0.1506	4.9588	1.4173	1.1175	4.6364	1.2485	1.1100
O_6	0.9773	0.9290	0.2310	7.5252	4.8587	1.7140	6.2907	4.4959	1.7025
O_7	0.8530	0.6980	0.0217	6.5681	3.6505	0.1610	5.5522	3.5256	0.1599

The bond lengths and valency angles in the tiglic acid molecule calculated from the values of X'_p, Y'_p, Z'_p are shown in fig.VIII. It will be obvious that the hydrogen of the carboxyl group is to be associated with O_6 rather than with O_7 thus confirming the position assigned to this atom on the basis of difference syntheses.
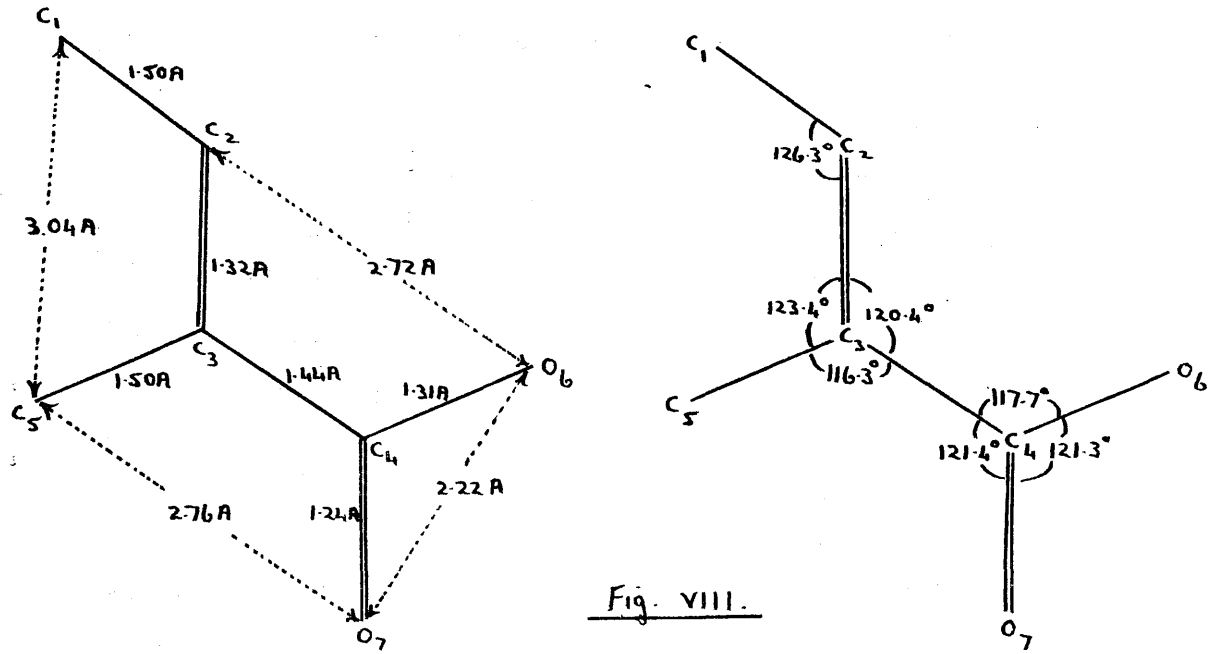
Fig. VIII.

It will be noticed that

$$C_2\widehat{C_3}C_4 \;+\; C_2\widehat{C_3}C_5 \;+\; C_5\widehat{C_3}C_4 \;=\; 360.1^\circ$$
$$C_3\widehat{C_4}O_6 \;+\; O_6\widehat{C_4}O_7 \;+\; O_7\widehat{C_4}C_3 \;=\; 360.4^\circ$$

Hence it follows that bonds C_2-C_3, C_3-C_4, C_3-C_5 lie approximately in one plane and bonds C_3-C_4, C_4-O_6, C_4-O_7 in another plane. It does not follow that these planes are the same since the one may be rotated with respect to the other about their common line C_3-C_4. However, calculation showed the two planes to be identical and the equation of the common plane of atoms C_2, C_3, C_4, C_5, O_6 and O_7 to be

$$0.8907X' \;-\; 0.4348Y' \;-\; 0.1324Z' \;-\; 3.4373 \;=\; 0 \dots\dots(i)$$

This equation is given in the form $lx + my + nz + p = 0$ where
l, m, n, are the direction cosines of the normal to the plane
and p is the perpendicular distance from the origin to the plane.
The mean distance of atoms C_2, C_3, C_4, C_5, O_6 and O_7 from the
plane is 0.008A whereas C_1 is 0.098A away from it. Hence, it
follows that all of the atoms of the tiglic acid molecule lie in
one plane exdept for C_1 .

Table II   Distances (d) of the atoms from the plane defined by (i)

Atom	C_1	C_2	C_3	C_4	C_5	O_6	O_7
d(A)	+0.098	–	$\overline{0.008}$	+0.016	+0.003	$\overline{0.014}$	$\overline{0.004}$

The orientation of the molecule in the cell may be defined by
  a) stating the equation of the molecular plane.
and
  b) fixing the position of a line in the molecule in this plane.

The equation of the plane has already been given above, and,
as will be shown later is very nearly, although not quite, the
$(\overline{2}10)$ plane.   The orientation of the molecule in this plane will be
defined by stating the angle between $C_3 - C_4$ and the c-axis of the
triclinic cell. This angle was found to be 52.8°.

8. Estimation of Accuracy.

The standard deviations of the atomic coordinates, bond lengths, bond angles and electron density were estimated by the methods described by Cruickshank[24,25], Ahmed and Cruickshank[26], Cruickshank and Robertson[27]. The essentials of these methods are as follows

a) <u>Standard Deviation of Atomic Coordinates</u> is given by the equation

$$\sigma(X_P) = \frac{2\pi\left\{\sum_q h^2 (F_o - F_c)^2\right\}^{\frac{1}{2}}}{a.A.C_a}$$

where $\sigma(X_P)$ is the standard deviation of the atomic coordinate of an atom P measured in the direction of the <u>a</u>-axis of the triclinic cell.

a is the length of the projected <u>a</u>-axis of the cell.

A is the area of the plane of projection.

C_a is the curvature of atom P measured in the direction of the projected <u>a</u>-axis.

h, F_o and F_c have their usual meanings.

Similar expressions obtain for $\sigma(Y_P)$ and $\sigma(Z_P)$.

Use of these expressions give $\sigma(X_P) = \sigma(Y_P) = \sigma(Z_P)$.

Atom	C_1	C_2	C_3	C_4	C_5	O_6	O_7
$\sigma(X_P)$	0.012A	0.009A	0.009A	0.009A	0.012A	0.007A	0.007A
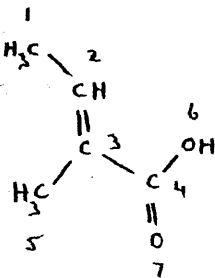
Table III

i.e. the accuracy of the positions obtained from the atoms is in the order oxygen more accurate than sp^2 carbon which in turn is more accurate than sp^3 carbon.

Standard Deviation of Bond Lengths.

If $\sigma(X_1')$, $\sigma(Y_1')$, $\sigma(Z_1')$ represent the standard deviations of the atomic coordinates of atom 1 and $\sigma(X_2')$, $\sigma(Y_2')$, $\sigma(Z_2')$ represent the standard deviations of the atomic coordinates of atom 2, and l, m, n are the direction cosines of the bond joining atoms 1 and 2 then the standard deviation of bond lengths $\sigma(d)$ is given by

$$\sigma^2(d) = \left\{\sigma^2(X_1') + \sigma^2(X_2')\right\} l^2 + \left\{\sigma^2(Y_1') + \sigma^2(Y_2')\right\} m^2$$
$$+ \left\{\sigma^2(Z_1') + \sigma^2(Z_2')\right\} n^2$$

if, as is very nearly the case in this analysis, $\sigma(X_1') = \sigma(Y_1') = \sigma(Z_1')$
$$\sigma(X_2') = \sigma(Y_2') = \sigma(Z_2')$$

then $\quad \sigma^2(d) = \left\{\sigma^2(X_1') + \sigma^2(X_2')\right\}\left\{l^2 + m^2 + n^2\right\}$

i.e. $\quad \sigma^2(d) = \sigma^2(X_1') + \sigma^2(X_2')$

This formula gives the following values for the standard deviations of bond lengths.

Bond	C_1-C_2	$C_2=C_3$	C_3-C_4	C_3-C_5	C_4-O_6	$C_4=O_7$
$\sigma(d)$	0.015A	0.013A	0.013A	0.015A	0.012A	0.012A

Table IV

Standard Deviation of Bond Angles
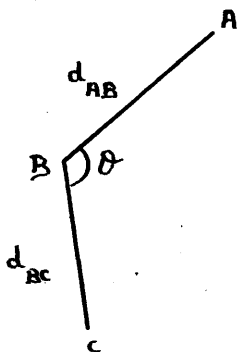
The standard deviation of bond angles is given by

$$\sigma^2(\theta) = \frac{\sigma^2(A)}{d_{AB}^2} + \sigma^2(B)\left\{\frac{1}{d_{AB}^2} - \frac{2\cos\theta}{d_{AB}d_{BC}} + \frac{1}{d_{BC}^2}\right\} + \frac{\sigma^2(c)}{d_{BC}^2}$$

where d_{AB} and d_{BC} are the lengths of the bonds joining atoms A,B and C.

θ is the angle between these bonds.

$\sigma(A)$, $\sigma(B)$, $\sigma(C)$ are the standard deviations of the positions of A, B and C.

$\sigma(\theta)$ radians is the standard deviation of θ

Use of this equation gives values of $\sigma(\theta)$ for the bond angles of tiglic acid of the order of $0.8°$.

Standard Deviation of Electron Density.

The standard deviation of the projected electron density $\sigma(\rho_o)$ is given by

$$\sigma(\rho_o) = \frac{1}{A} \left\{ \sum (F_o - F_c)^2 \right\}^{\frac{1}{2}}$$

giving a value of $0.14 e/A^2$ for the b-axis projection and a value of $0.3 e/A^2$ for the a-axis projection.

9. Discussion of the Structure and Molecular Dimensions.
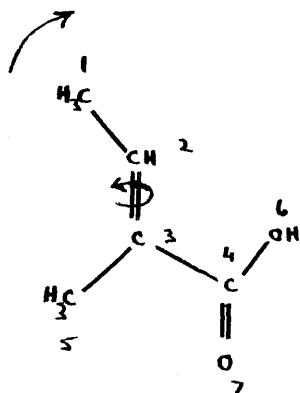
The displacement of atoms C_2, C_3, C_4, C_5, O_6 and O_7 from the plane of the molecule varies from zero to 0.016A.   These deviations have no significance.   However, C_1 is situated 0.098A away from the plane. Hence it follows that C_1 is not coplanar with the rest of the molecule.   If all the atoms lay in the same plane and all the bond angles were $120°$, as would be expected if the structure were that of the idealised chemical formula, then C_1 and C_5 would approach to within 2.8A.   Apparently this distance of approach is far too close and the methyl groups have repelled one another until they are separated by a distance of 3.04A. C_5 obviously cannot move very far otherwise it would come into the field of O_7 and so most of the repulsion has centred round C_1 .   The bond angle $C_1\widehat{C_2}C_3$ has widened out until it becomes $126.3°$.   Bond C_2-C_1 has rotated slightly about C_3-C_2 as axis in an anticlockwise direction until its direction makes an angle of $4°$ with the plane of the rest of the molecule.

The angle $C_2\widehat{C_3}C_5$ has widened out to $123.4°$ i.e. it is wider than the sp^2 bond angle although the distortion is not as large as in the case of $C_1\widehat{C_2}C_3$.

$C_2=C_3$ and C_4-O_7 are conjugated. Consequently it would be expected that C_3-C_4 should differ from a pure single bond.   A shortening of 0.06A from the pure single sp^2 bond length of 1.50A has been observed and since this difference is five times the standard deviation to be expected in this bond the result shows that a considerable degree of conjugation does in fact exist.

The system

$$\overset{a}{\underset{b}{>}}C = \overset{OH}{\underset{O}{\overset{|}{C} - C}}$$
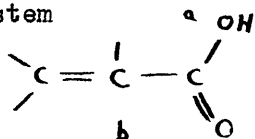
is found in other acids whose structures
have been accurately worked out by
X-ray methods and it is interesting
to compare the results observed for tiglic acid with some other acids
of similar type.   Thus

	$C_\alpha{-}C_b$	$C_\alpha{-}OH$	$C_\alpha{=}O$	$\widehat{C_b{-}C_\alpha{-}OH}$	$\widehat{C_b{-}C_\alpha{=}O}$	$\widehat{HO{-}C_\alpha{=}O}$	hydrogen bonded distance
Salicylic Acid(28)	1.46A	1.33A	1.24A	117.0°	122.7°	120.2°	2.63A
Benzoic Acid (29)	1.48A	1.29A	1.24A	118.0°	122.0°	122.0°	2.64A
Tiglic Acid	1.44A	1.31A	1.24A	117.7°	121.4°	121.3°	2.64A

Table V.

The coordinates listed in table I (p. 31) are those of a molecule lying
at the $\big[\,[110]\,\big]$ corner of the triclinic cell.   The corner of the unit
cell is the centre of symmetry of the dimer.   However, the centre of
symmetry does not lie on the molecular plane defined by equation(1).
It lies at a perpendicular distance of 0.05A away from the plane.
It follows therefore that the two molecules comprising a dimer of
tiglic acid do not lie in the same plane.   Instead, the planes of
the two halves of the dimer are displaced 0.1A from one another to
give what might be called a "stepped" arrangement of hydrogen bonding
in the dimers.
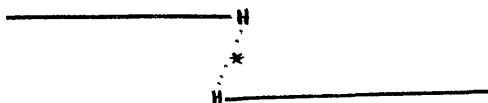
The molecular plane which has been defined by the equation

$$0.8907X' - 0.4348Y' - 0.1324Z' - 3.4373 = 0 \qquad \ldots\ldots\ldots\ldots(i)$$

is very nearly, although not quite, the $(2\bar{1}0)$ plane.   It lies
0.05A from the $\left[\left[110\right]\right]$ corner of the triclinic cell and is 3.44A
from the origin of this cell.
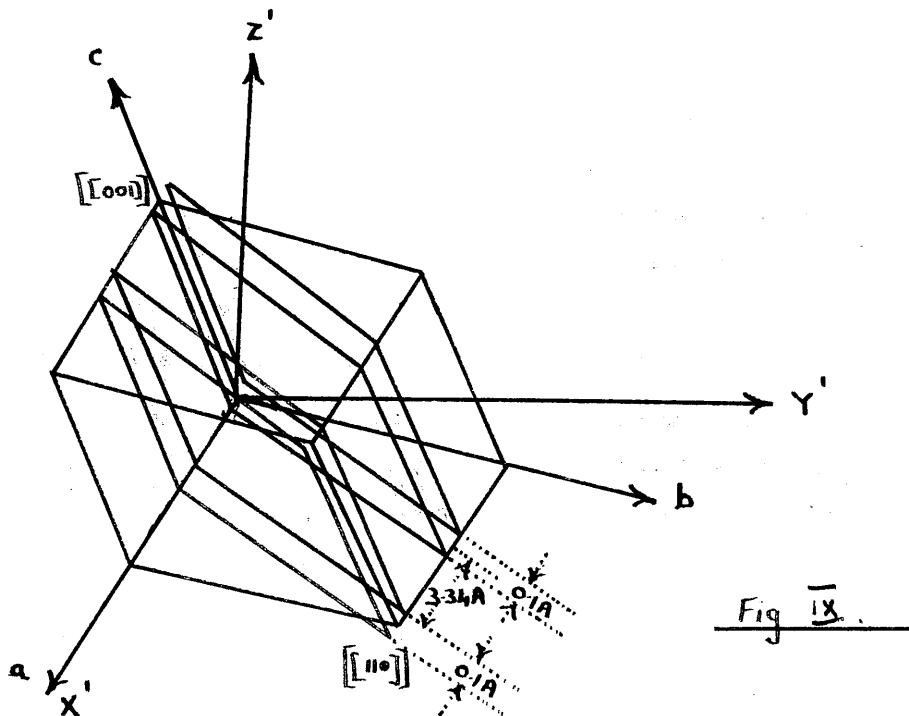
Fig $\overline{IX}$.

It now follows that the plane of the associated molecule comprising
the dimer at $\left[\left[110\right]\right]$ must lie at a perpendicular distance of 3.34A
from the origin and these planes are parallel to one another.   This
stepped arrangement exists right through the crystal.   Consequently
the structure can be considered as being made up of alternating pairs
of sheets of molecules.   The members of each pair are joined by
hydrogen bonds and are separated by a perpendicular distance of 0.1A
from each other.   The pairs of sheets are held together by

Van der Waals forces and are 3.29A apart (fig. X).   However, it does not follow that non-bonded atoms are 3.29A apart.
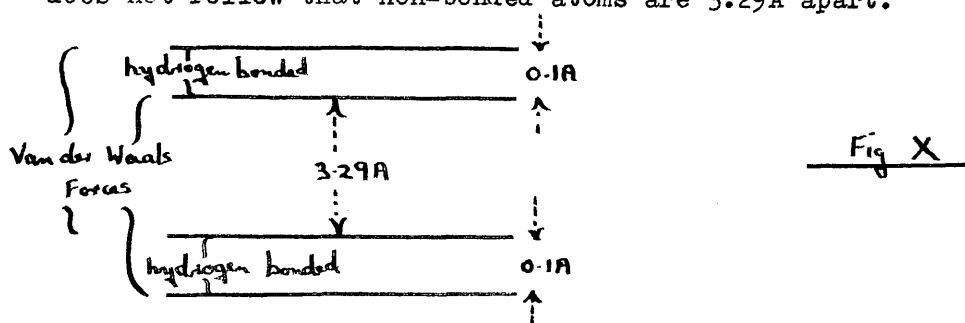
Fig X

Fig.XI shows the arrangement of the dimers whose centres of symmetry lie on the (010) plane.   In this arrangement, the shortest intermolecular distances are those between carbon and oxygen atoms, the shortest being a methyl to oxygen distance of 3.58A.

Fig.XII shows the arrangement of the dimers whose centres of symmetry lie on the (001) plane, and demonstrates the closest distance of approach between the doubled sheets of molecules.   The closest distance of approach is again a methyl to oxygen distance – this time of 3.63A.

The closest carbon to carbon non-bonded distance of approach is demonstrated in fig.XII and is between methyl groups related to each other by centres of symmetry.   This distance is 3.70A and is somewhat less than should be expected from the radius of 2A assigned to a methyl group by Pauling(30) and the corresponding value of 3.87A observed in Durene and Hexamethylbenzene (31,32).   However, shorter non-bonded methyl to methyl distances than these have been observed before e.g. in 2, 3, 5, 6 – Tetramethyl Pyrazine (33) a distance of 3.72A has been observed and a value of 3.64A has been quoted for dl – Alanine (34).   As Corey has stated (34) when longer non-bonded methyl to methyl distances are involved it is probable that the rotations of the methyl groups are independent of one another.   As the
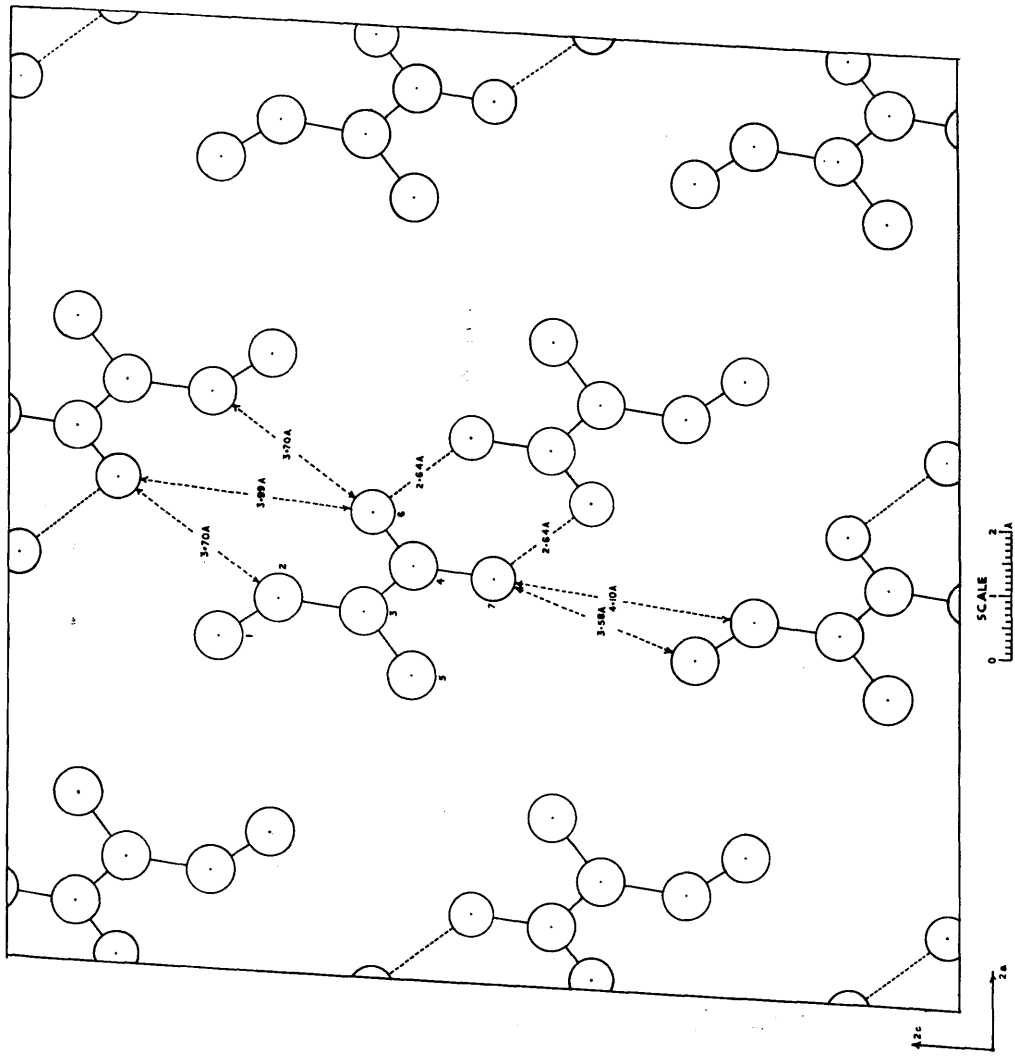
Fig XI.

Tiglic Acid.

Arrangement of molecules on (010) plane.

Fig. XII.

Arrangement of molecules on (001) plane.

Tiglic acid.

non-bonded methyl to methyl distance decreases the rotations will tend to synchronise and the carbon – hydrogen bonds will act like the teeth of two cog-wheels revolving round each other.

Whether or not this explanation for the close distance of approach of methyl groups belonging to adjacent molecules is correct, it must be true for methyl groups belonging to the same molecule if these are rotating. The distance between methyl groups in the same molecule in this case, is only 3.04A and so it follows immediately that the hydrogen atoms of these methyl groups must have a staggered configuration. If these methyl groups rotate independently of one another then sooner or later the hydrogen atoms will approach to within about 1.8A. Since the Van der Waals radius of hydrogen is 1.2A such a close distance of approach is inconceivable and consequently if the methyl groups are rotating it must follow that the rotations of the methyl groups in the same molecule are synchronised.

It is interesting to note that a similar explanation of the mode of rotation of the methyl groups in Hexamethylbenzene was put forward by Mack as long ago as 1932.(36).

44.

Table VI

Observed (|Fo|) and Calculated (Fc) Structure Factors of Tiglic Acid.

F(okl)

Plane.	$2\sin\theta$.	|Fo|	Fc	Plane.	$2\sin\theta$.	|Fo|	Fc
001	0.207	6.3	+ 7.3	020	0.615	5.6	− 7.2
002	0.415	1.1	+ 0.5	021	0.675	11.1	− 11.7
003	0.623	10.3	− 11.8	022	0.785	7.6	− 6.6
004	0.835	9.0	+ 9.4	023	0.930	5.4	− 5.2
005	1.050	6.5	+ 6.4	024	1.100	8.5	+ 7.9
006	1.255	1.6	− 1.2	025	1.285	1.2	+ 1.4
007	1.465	4.2	+ 4.0	026	1.470	3.4	− 2.8
008	1.672	1.2	+ 1.5	027	1.667	2.6	− 2.0
009	1.878	1.2	+ 0.8	02$\bar{1}$	0.622	16.1	− 16.8
010	0.305	16.8	− 18.0	02$\bar{2}$	0.697	16.4	− 19.8
011	0.395	14.0	+ 15.8	02$\bar{3}$	0.820	7.8	+ 8.6
012	0.550	17.1	− 20.6	02$\bar{4}$	0.973	1.9	+ 1.8
013	0.730	4.1	− 3.2	02$\bar{6}$	1.325	10.7	− 10.5
014	0.928	7.5	− 5.7	02$\bar{7}$	1.515	2.4	− 1.9
016	1.332	2.3	− 1.7	030	0.915	5.9	+ 5.8
017	1.535	5.2	− 4.1	031	0.963	10.9	+ 11.4
018	1.738	0.6	− 0.8	032	1.050	4.6	+ 5.1
01$\bar{1}$	0.348	2.3	+ 3.1	033	1.170	2.7	− 1.9
01$\bar{2}$	0.490	1.7	− 2.9	034	1.315	2.1	− 1.6
01$\bar{3}$	0.665	1.8	− 1.2	035	1.480	4.2	+ 3.5
01$\bar{4}$	0.855	9.2	+ 8.3	036	1.650	3.8	+ 3.4
01$\bar{5}$	1.055	1.4	− 1.2	03$\bar{1}$	0.910	1.4	− 0.9
01$\bar{6}$	1.255	4.2	+ 4.0	03$\bar{2}$	0.957	5.5	+ 5.2
01$\bar{7}$	1.460	2.3	− 1.8	03$\bar{3}$	1.040	1.8	+ 2.1
01$\bar{8}$	1.645	2.1	+ 1.4	03$\bar{4}$	1.155	0.8	+ 0.6

Plane.	2sinθ.	$\|Fo\|$	Fc	Plane.	2sinθ.	$\|Fo\|$	Fc
03$\bar{5}$	1.300	2.9	− 2.7	04$\bar{8}$	1.930	2.2	+ 1.6
03$\bar{6}$	1.455	1.0	+ 1.6	050	1.530	5.9	− 5.9
03$\bar{7}$	1.610	1.2	+ 1.1.	054	1.835	1.1	− 1.1
040	1.220	5.9	+ 5.5	05$\bar{1}$	1.517	3.8	− 3.0
041	1.265	4.1	+ 3.8	05$\bar{3}$	1.573	2.0	+ 1.6
043	1.435	3.8	− 2.6	05$\bar{4}$	1.650	0.8	− 0.8
044	1.555	1.6	− 1.4	05$\bar{5}$	1.737	1.8	− 1.9
04$\bar{2}$	1.240	1.4	− 1.7	061	1.870	0.8	− 1.0
04$\bar{3}$	1.300	1.6	+ 1.0	062	1.925	1.1	+ 0.7
04$\bar{4}$	1.390	1.6	+ 1.6	06$\bar{1}$	1.820	0.7	− 0.3
04$\bar{6}$	1.625	0.6	+ 0.6	06$\bar{3}$	1.860	1.2	+ 1.2
04$\bar{7}$	1.770	2.3	+ 2.4				

$$R = 100 \left\{ \frac{\sum \left| |Fo| - |Fc| \right|}{\sum |Fo|} \right\}.$$

$$= 13.2\%$$

F(hol)

Plane.	$2\sin\theta$.	|Fo|	Fc	Plane.	$2\sin\theta$.	|Fo|	Fc
001	0.215	7.4	+ 7.9	201	0.460	16.2	+ 16.1
002	0.420	1.2	+ 1.6	202	0.580	20.7	− 22.7
003	0.630	12.6	−13.2	203	0.745	3.3	− 3.0
004	0.835	9.1	+ 8.7	204	0.920	7.8	− 5.5
005	1.050	5.9	+ 6.5	206	1.300	3.0	− 1.3
006	1.250	1.7	− 1.7	207	1.495	5.4	− 5.0
007	1.460	3.9	+ 3.9	$20\bar{1}$	0.480	5.2	+ 6.2
008	1.670	1.3	+ 1.0	$20\bar{2}$	0.605	4.4	− 5.3
009	1.880	1.1	+ 0.9	$20\bar{3}$	0.770	1.2	− 0.1
100	0.210	21.0	+21.9	$20\bar{4}$	0.950	6.6	+ 7.1
101	0.295	35.5	+ 34.6	$20\bar{5}$	1.145	0.9	− 1.4
102	0.460	4.5	+ 4.6	$20\bar{6}$	1.340	3.4	+ 3.3
103	0.660	3.4	− 3.9	$20\bar{7}$	1.540	1.8	− 1.8
104	0.850	2.3	+ 2.2	$20\bar{8}$	1.740	1.2	+ 1.0
105	1.060	1.3	+ 0.9	300	0.630	0.6	+ 0.7
106	1.260	3.3	− 3.3	301	0.655	4.1	− 4.5
107	1.460	3.4	− 3.3	302	0.740	5.9	− 5.7
108	1.670	1.6	+ 1.9	303	0.870	7.4	− 7.2
109	1.880	1.1	+ 1.1	304	1.020	2.9	− 3.5
$10\bar{1}$	0.310	12.9	−11.8	305	1.190	5.2	− 4.7
$10\bar{2}$	0.480	16.5	−16.4	306	1.370	3.4	− 2.7
$10\bar{3}$	0.675	7.9	+ 8.2	307	1.560	1.7	− 0.5
$10\bar{4}$	0.870	13.5	+12.9	$30\bar{1}$	0.670	7.6	− 9.4
$10\bar{5}$	1.080	6.1	+ 6.5	$30\bar{2}$	0.770	14.4	−15.5
$10\bar{6}$	1.280	1.8	+ 1.3	$30\bar{3}$	0.910	8.9	− 8.9
$10\bar{7}$	1.480	2.5	+ 2.0	$30\bar{4}$	1.070	2.2	+ 2.1
200	0.420	17.3	−16.9	$30\bar{5}$	1.250	3.0	+ 2.5

Plane.	$2\sin\vartheta$.	$|F_o|$	Fc		Plane.	$2\sin\vartheta$.	$|F_o|$	Fc
$30\overline{7}$	1.625	1.7	− 2.0		600	1.250	2.5	+ 3.4
401	0.855	13.3	− 12.5		601	1.260	6.5	+ 7.1
402	0.910	2.9	− 2.9		602	1.300	< 1.0	− 0.6
403	1.020	3.9	− 5.1		604	1.470	2.1	− 1.9
404	1.150	10.7	+ 8.5		605	1.590	5.2	+ 4.5
405	1.300	< 1.0	− 0.4		606	1.725	1.5	+ 0.9
406	1.470	3.7	− 3.1		$60\overline{1}$	1.280	1.8	− 1.1
407	1.640	2.9	− 1.6		$60\overline{2}$	1.340	3.7	+ 2.8
$40\overline{1}$	0.880	9.1	− 11.5		$60\overline{3}$	1.425	< 1.0	− 0.6
$40\overline{2}$	0.955	14.4	− 14.2		$60\overline{4}$	1.540	1.0	+ 1.8
$40\overline{3}$	1.075	3.0	+ 3.3		$60\overline{5}$	1.670	1.9	− 2.2
$40\overline{4}$	1.210	1.0	+ 1.6		706	1.875	1.3	+ 1.6
$40\overline{5}$	1.375	2.1	+ 2.1		$70\overline{1}$	1.480	2.5	+ 1.8
$40\overline{6}$	1.545	6.3	− 6.2		$70\overline{2}$	1.540	1.0	+ 1.4
$40\overline{7}$	1.730	< 0.9	− 1.4		800	1.665	1.6	+ 1.7
500	1.045	3.9	+ 2.7		801	1.660	< 0.9	+ 0.9
501	1.060	1.6	+ 1.7		802	1.690	0.9	− 1.0
502	1.105	8.5	− 7.7		803	1.755	2.6	− 1.6
503	1.200	3.5	− 3.0		805	1.920	1.0	− 0.8
504	1.300	4.5	+ 3.6		$80\overline{1}$	1.690	2.7	+ 2.6
505	1.040	6.7	+ 6.7		$80\overline{2}$	1.740	1.7	− 1.6
507	1.750	1.4	− 1.3		900	1.880	< 0.7	+ 0.9
$50\overline{1}$	1.080	2.2	− 2.2		901	1.880	< 0.7	+ 1.2
$50\overline{2}$	1.150	1.0	− 1.0		903	1.950	0.4	− 0.7
$50\overline{3}$	1.245	3.0	+ 3.3		$90\overline{1}$	1.905	1.0	− 1.5
$50\overline{5}$	1.520	3.1	− 3.5		$90\overline{2}$	1.950	0.4	− 1.0
$50\overline{6}$	1.670	2.8	− 2.0					

$$R = 100\left\{\frac{\sum\big||F_o| - |F_c|\big|}{\sum|F_o|}\right\}$$

$$= 10.8\%$$

F(hko)

Plane.	$2\sin\theta$.	$\lvert Fo\rvert$	Fc		Plane.	$2\sin\theta$.	$\lvert Fo\rvert$	Fc
010	0.305	16.7	− 17.8		300	0.630	0.6	+ 0.8
020	0.612	5.0	− 7.3		310	0.765	5.8	+ 4.4
030	0.916	5.6	+ 6.0		320	0.980	2.0	+ 2.3
040	1.224	5.2	+ 4.6		330	1.230	1.8	+ 1.6
050	1.525	5.4	− 4.8		340	1.505	4.4	− 3.5
100	0.210	21.7	+ 21.8		350	1.785	< 0.9	+ 0.8
110	0.410	5.9	+ 4.8		$\bar{3}10$	0.625	22.5	+ 21.3
120	0.697	6.4	+ 6.5		$\bar{3}20$	0.760	3.0	+ 0.8
130	0.990	1.5	+ 1.1		410	0.960	5.2	+ 6.5
140	1.295	< 1.0	+ 0.5		420	1.155	5.3	+ 4.4
150	1.592	3.2	− 2.4		430	1.390	4.3	− 4.6
$\bar{1}10$	0.325	7.7	+ 7.1		440	1.645	< 1.0	− 0.1
$\bar{1}20$	0.595	1.5	+ 3.1		$\bar{4}10$	1.815	6.6	− 6.0
$\bar{1}30$	0.885	1.8	+ 3.2		$\bar{4}20$	0.900	20.5	+ 25.7
$\bar{1}40$	1.190	< 1.0	+ 0.1		$\bar{4}30$	1.060	3.2	− 2.7
$\bar{1}60$	1.790	1.3	− 1.2		$\bar{4}60$	1.810	1.4	+ 1.5
200	0.420	16.0	− 14.3		500	1.045	4.0	+ 2.1
210	0.580	4.9	− 4.5		510	1.160	2.4	+ 2.2
220	0.825	7.6	+ 8.7		520	1.335	2.6	+ 1.8
230	1.100	7.1	+ 5.3		530	1.550	4.3	− 3.2
240	1.390	5.8	− 5.7		540	1.790	< 0.8	+ 0.9
$\bar{2}10$	0.455	67.7	+ 70.6		$\bar{5}10$	1.010	2.7	− 0.2
$\bar{2}20$	0.650	7.5	− 7.4		$\bar{5}20$	1.070	10.1	+ 10.7
$\bar{2}30$	0.905	1.1	− 4.7		600	1.255	2.6	+ 3.4
$\bar{2}40$	1.185	1.4	+ 2.2		610	1.365	2.8	+ 2.8
$\bar{2}50$	1.470	3.0	+ 2.9		620	1.525	2.1	− 2.4
$\bar{2}60$	1.775	2.4	− 3.0		630	1.730	< 0.9	− 0.4

Plane.	2sinθ.	$|Fo|$	Fc	Plane.	2sinθ.	$|Fo|$	Fc
$\overline{6}20$	1.250	2.3	− 1.9	$\overline{7}20$	1.430	1.9	− 0.9
$\overline{6}30$	1.345	5.8	+ 7.8	$\overline{7}30$	1.510	3.0	+ 4.4
$\overline{6}40$	1.505	2.4	− 1.1	800	1.665	1.7	+ 1.4
710	1.560	2.1	+ 1.4	$\overline{8}40$	1.800	1.1	+ 2.2
720	1.710	2.6	− 2.1	900	1.880	< 0.7	+ 0.8
730	1.905	1.1	+ 0.8	$\overline{9}40$	1.960	0.6	+ 1.6
$\overline{7}10$	1.410	1.5	+ 1.2				

$$R = 100\left\{\frac{\Sigma \left||Fo| - |Fc|\right|}{\Sigma |Fo|}\right\}$$

$$= 15.9\%$$

THE CRYSTAL AND MOLECULAR STRUCTURE

OF

ANGELIC ACID

The Crystal and Molecular Structure of Angelic Acid.

1.   Crystal Data.

Angelic Acid, $C_5H_8O_2$;   Molecular Weight = 100.1;   m.pt. = 45.5°C.
Density calculated = 1.139 gms/cc.
Density observed(by flotation in phosphoric acid solution)
                = 1.141gms/cc.

a = 7.66 $\pm$0.02A   b = 11.60 $\pm$0.02A   c = 6.67 $\pm$ 0.02A

α = 90° β = 100° $\pm$ 0.5° γ = 90°

Systematic absences:- (hol) when h = 2n + 1;   (oko) when k = 2n + 1
Crystal system:- Monoclinic.
Space group:- P 2₁/a.
Four molecules per unit cell.   No molecular symmetry required.
Centre of symmetry between two molecules.
Volume of unit cell (V) = abc $\sin\beta$ = 583.7A^3.
Absorption coefficient for X-rays(λ= 1.542A)μ = 8.39cm^{-1}.
Total number of electrons per unit cell = F(000) = 216.

The crystals were obtained by slow recrystallisation from 60 - 80
petroleum ether.   This resulted in thin plates with the $\{001\}$ face
prominent.

    The above data agree reasonably well with that found by
Schimper(37) who gives
        a : b : c : = 0.6494 : 1 : 1.1393
                β = 100°33'
The X-ray data give
        a : b : c = 0.6603 : 1 : 0.5750
i.e. Schimper's cell is twice the size of the true unit cell.

2   Experimental Measurements.

Angelic acid is very volatile and exists as soft, distorted plates.   These plates cleave readily along the (20$\bar{1}$) plane and so thin needles with the needle axis parallel to the b-axis of the unit cell can easily be cut.

However, difficulty was experienced in obtaining specimens suitable for taking rotation photographs with the crystals rotating about either the a or c axes since attempts to cut the crystals in any other direction usually resulted in grossly distorted specimens.

Since the acid is volatile, some method of protecting the crystals whilst the X-ray photographs are being taken is necessary. As in the case of tiglic acid, the crystal on its glass fibre, was encapsulated in a Park-Davis gelatine capsule and extra crystals of angelic acid were added under the capsule to build up the vapour pressure.   It was found the capsule acted as a permeable membrane and the crystal volatilised steadily whilst the photographs were being taken.   As a result, the crystals had to be sealed in thin walled glass capillary tubes.   This method proved to be perfectly satisfactory and in this way the crystals could be kept indefinitely.

Rotation, oscillation and moving-film photographs were taken using Cu – Kα radiation(λ = 1.542A.).   The unit cell dimensions were obtained by means of rotation and equatorial layer line moving-film photographs.   The reflections were estimated visually by/

by the multiple film technique(17).   The structure factors
evaluated from the estimated intensities by means of the usual
mosaic crystal formula are given in table (VIII).

3. <u>Correlation of the Reciprocal Lattice of Angelic Acid with the</u>
   <u>Fourier Transform of the idealised 2-methyl-cis-2-butenic acid</u>
   <u>the dimensions of which are given on page II</u>

The Fourier Transforms of the idealised dimers of the
two 2-methyl-2-butenic acids have already been calculated and it has
been shown that tiglic acid is 2-methyl-<u>trans</u>-2-butenic acid.
Consequently, it follows on the basis of chemical evidence, that
angelic acid should be 2-methyl-<u>cis</u>-2-butenic acid.

However, the unit cell of angelic acid contains four molecules,
i.e. two centrosymmetric dimers of angelic acid, and so the
reciprocal lattice of the acid can not be obtained by simple
sampling of the transform of the idealised dimer.   Instead, the
weighted reciprocal lattice is obtained by sampling the sum of
two transforms correctly orientated with respect to one another.
However, there is no need to consider two separate transforms.   The
solution of the structure can be obtained from the transform of
one dimer by the following methods.

In the space group P 2_1/a , projection of the structure on to
the (010) plane results in a projected arrangement of molecules
similar to that shown in fig. ($\overline{XIII}$).
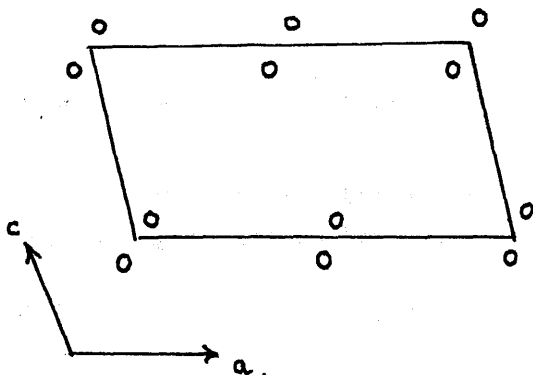
Figure $\overline{XIII}$.

i.e.   The problem now reduces to consideration of a simple
oblique cell half the size of the true cell and so the procedure
employed in fitting the triclinic reciprocal lattice of tiglic
acid to the Fourier Transform of the dimer of 2-methyl-trans-2-
butenic acid can again be used.

Unfortunately, in the particular case of angelic acid, it was
obvious from the strength of the reflection from the $(20\bar{1})$ plane
that on looking down the b-axis the molecule would be seen end on
and it would be useless to try to solve the structure by making
use of this projection.

Consequently, solution of the problem was approached by
considering projection on to a plane perpendicular to the a-axis
instead.

In this projection of the space group P 2_1/a the projected
arrangement of the molecules- is as shown in fig. ($\overline{XIV}$).
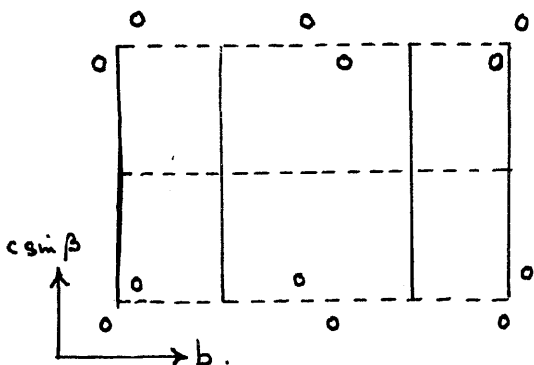
Figure $\overline{XIV}$.

The symmetry of the projection gives rise to the equivalent
points

$$[y, z], [\bar{y}, \bar{z}], [(\tfrac{1}{2}+y), \bar{z}], [(\tfrac{1}{2}-y), z]$$

Hence, using the same notation as before, we get

$$\frac{F(0kl)}{fc} = \sum_{j_1} w_{j_1} \cos 2\pi (ky_{j_1} + lz_{j_1}) + \sum_{j_2} w_{j_2} \cos 2\pi (ky_{j_2} + lz_{j_2})$$

$$\underset{\text{dimer 1}}{\quad} \qquad \underset{\text{dimer 2}}{\quad}$$

where the first term on the right hand side is to be summed over the

atoms of the first dimer and the second term is summed over the dimer 2 related to the first by means of the reflection line $y = \frac{1}{4}$.

Expanding the expression and dropping the subscripts 1 and 2 gives

$$\frac{F(okl)}{fc} = \sum_{dimer} w_j \cos 2\pi(ky_j + lz_j) + \sum_{dimer} w_j \cos 2\pi\left\{ k(\tfrac{1}{2} + y_j) - lz_j\right\}$$

$$= \sum_{dimer} w_j \cos 2\pi(ky_j + lz_j) + \sum_{dimer} w_j \cos\left\{ 2\pi(ky_j - lz_j) + k\pi\right\}.$$

$$= \sum_{dimer} w_j \cos 2\pi(ky_j + lz_j) + (-1) \sum_{dimer} w_j \cos 2\pi(ky_j - lz_j)$$

Suppose now that the correct orientation of the unweighted reciprocal lattice frame of angelic acid on the Fourier Transform of the dimer of 2-methyl-cis-2-butenic acid is known and consider the right hand side of the last expression. The first term is equivalent to the $T(X*Y*Z*)$ function at the point (okl) and the second term is that corresponding to the same function at the point $(ok\bar{l})$ i.e. we may now write

$$\frac{F(okl)}{fc} = \underset{dimer}{T(okl)} + (-1)^k \underset{dimer}{T(ok\bar{l})}$$

Hence the approximate value of $\frac{F(okl)}{fc}$ for angelic acid can be obtained by correctly orientating the reciprocal lattice frame of the acid on the Fourier Transform of the idealised 2-methyl-cis-2-butenic acid, reading off the values of the transform at the points (okl) and $(ok\bar{l})$, and adding these if k is even or subtracting if k is odd.

i.e. when $k = 2n$, $\frac{F(okl)}{fc} = \underset{dimer}{T(okl)} + \underset{dimer}{T(ok\bar{l})}$

when k = 2n + 1, $\dfrac{F(ok1)}{fc}$ = T(ok1) − T(ok1)

It should be noted that

$$\dfrac{F(oo1)}{fc} = 2T(oo1) \text{ for all values of } 1,$$

and $\quad \dfrac{F(oko)}{fc} = 2T(oko) \text{ when } k = \text{even}$

$$= 0 \text{ when } k = \text{odd.}$$

Although $\dfrac{F(oko)}{fc} = 0$ when k = odd, it does not follow that the value of the transform T(X∗Y∗Z∗) at the point (oko) is zero when k = odd. The zero values for $\dfrac{F(oko)}{fc}$ arise because a particular value of the T(X∗Y∗Z∗) function is being subtracted from itself.

Fig. XV.
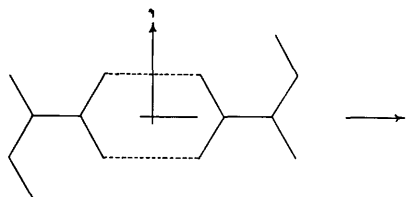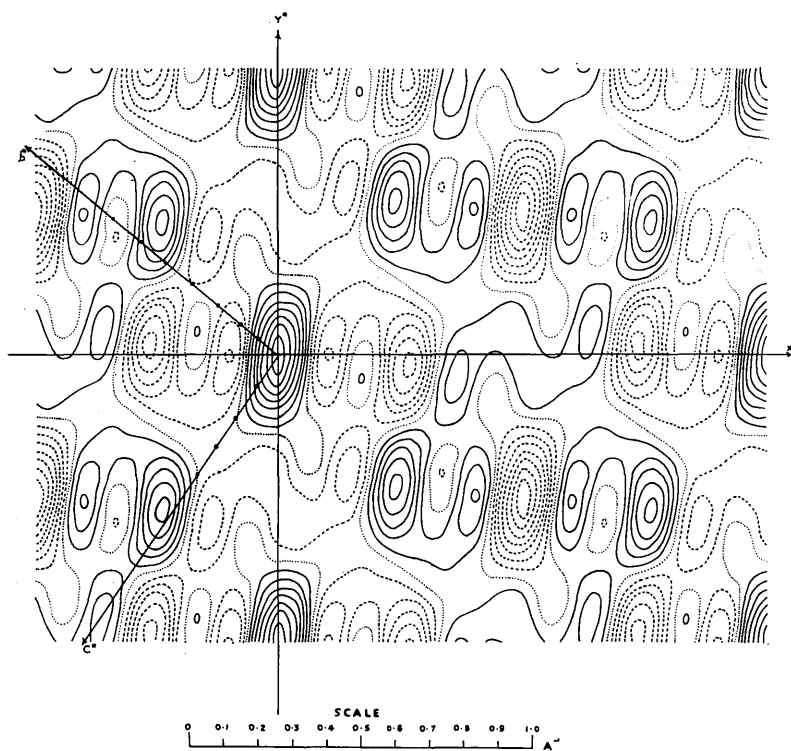
SCALE

0   0·1   0·2   0·3   0·4   0·5   0·6   0·7   0·8   0·9   1·0
A*

Fourier transform of idealised dimer of
2-methyl-cis-2-butenic acid
with the b* and c* reciprocal axes of
angelic acid superimposed.

4.   The Structure of Angelic Acid.

Projection on to the Plane perpendicular to the a-axis of the unit cell.

The discussion just given, holds only if the orientation of the reciprocal lattice on the Fourier transform is known.   This is precisely the problem being considered and so, at first sight it would seem that what has been said would be of no immediate use in solving the structure of angelic acid.   Nevertheless, a number of clues to the way in which the (okl) reciprocal lattice net is arranged on the transform, are given by the discussion.

First of all, the values of $\frac{F(oko)}{fc}$ and $\frac{F(ool)}{fc}$ are twice the the values of $T(oko)$ and $T(ool)$ read off from the transform and so addition and subtraction of terms need not be considered for these when seeking the correct orientation.

Secondly, strong values of $\frac{F(okl)}{fc}$ must correspond with large sums $T(okl) + (-1)^k T(ok\bar{l})$ and weak values must correspond with small sums.

Forty-six independent (okl) reflections out of a possible 102 were observed and estimated.   The intensities of these reflections were placed on an approximately absolute scale by means of the Wilson method(38), and with the above rules concerning the (oko) and (ool) reflections, coupled with the stronger general reflections (011), (032), (041), (042), (055), a reasonable fit of the reciprocal lattice net on to the transform was soon obtained.

Fig.($\overline{XV}$) shows the Fourier transform of an idealised dimer of

2/

Fig. XVI.
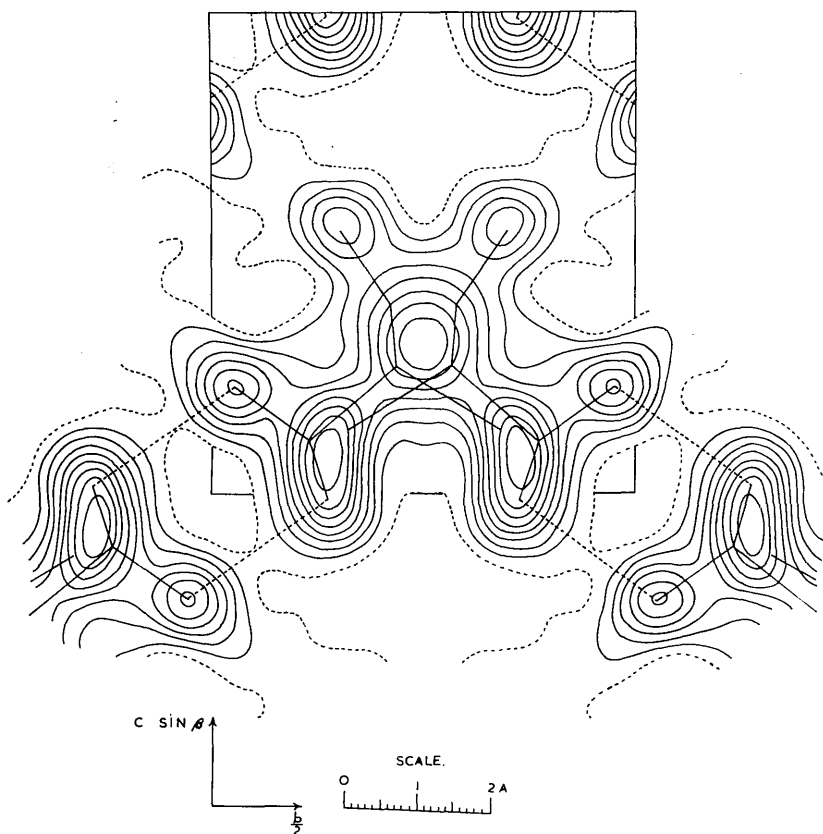
Angelic acid. a -axis projection.

The electron density was computed using phase angles obtained
from the Fourier transform of the idealised dimer of
2-methyl - cis - 2 - butenic acid.

Contours at intervals of 1e /A². 1e /A² contour dotted.

2-methyl-cis-2-butenic acid with the b* and c* reciprocal axes of angelic acid superimposed on it.

Phase angles obtained from the transform, combined with the observed structure factor amplitudes were then used to calculate the electron density projected on to the plane perpendicular to the a-axis.

Atomic centres in positions consistent with the expected geometry of the angelic acid molecule were then picked out from this map and structure factors were evaluated from these.

The discrepancy (R) at this stage was 36.2%.

At a similar stage in the analysis of the tiglic acid structure, the discrepancy was 21% and so it seemed likely at this point that the true angelic acid structure differed more from the idealised dimer of 2-methyl-cis-2-butenic acid than did tiglic acid from the corresponding trans acid.

This projection was then refined by successive difference syntheses until no further refinement could be effected.   At this stage, the final value of the discrepancy was ~~16.4%~~ 15.8% .

An electron density map of the projected structure is shown in fig.(XVI).

The projection of the crystal structure on to the plane perpendicular to the a-axis should have given two reasonably accurate atomic coordinates (y_p, z_p) for each atom P.   The dimensions of the angelic/

angelic acid molecule should be expected to be roughly those of the idealised 2-methyl-cis-2-butenic acid and are therefore known approximately. Hence, it should be quite a simple matter to calculate the third coordinate (x_p) for each atom and so refine the other two views of the molecule obtained by projecting the structure down the b and c - axes respectively.

This process was carried out and attempts were made to refine the c-axis projection. In this, the calculated coordinates x_p, were combined with the coordinates y_p obtained from the a-axis projection. The resultant discrepancy between the observed and calculated structure factors F(hko) was 39.8%. Successive refinements by means of difference syntheses brought down the value (R) in the sequence 39.8% – 35.2% – 31.5% – 27.9% and when hydrogen atoms were included in the calculations R fell to 23% at which point the projection stopped refining any further. A map of the electron density computed at this point, exhibited so much distortion that it was obvious that something was wrong with the structure.

A re-examination of the electron density map obtained from the a-axis projection gave no indication what could be wrong. There seemed no other way of choosing a constellation of atoms from this map which was consistent with the X-ray data and at the same time was chemically reasonable. Consequently, the work up to this point was abandoned for the moment and it was decided to try to find a new structure.

Fig. XVII.
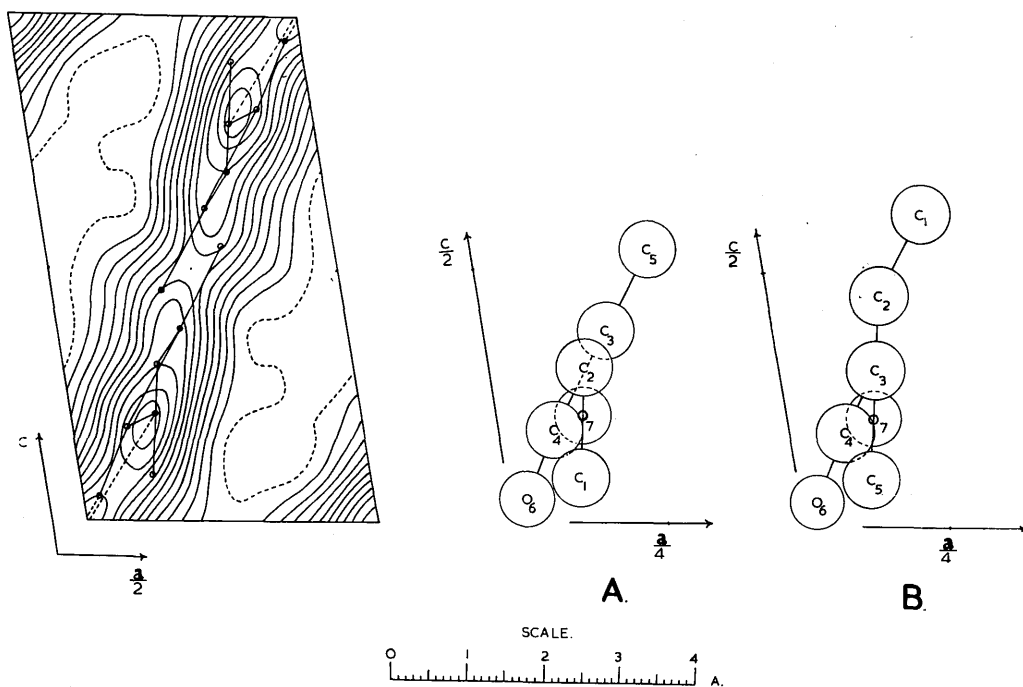
Angelic acid — b — axis projection showing final electron density map (corresponding to structure A), and the alternative structures A and B.

Contours at intervals of $1e/A^2$.. $1e/A^2$ contour dotted. ⸗

Projection on to the (010) plane of the unit cell.

At this stage, the best hope of finding a new structure seemed to be to tackle the b-axis projection. The reflected beam obtained from the (20$\bar{1}$) plane is so very strong that all the atoms in the unit cell must lie on, or very nearly on this plane.

To solve the structure therefore, it is only necessary to place a model of the idealised 2-methyl-cis-2-butenic acid on this plane, with the centre of symmetry of the dimer coincident with the crystallographic centre of symmetry, and rotate the dimer in this plane about an axis through the centre of symmetry until the observed and calculated structure factors agree.

This was done and it soon became obvious that there were only two arrangements in space which gave reasonable agreement between observed and calculated structure factors. These orientations are shown in figure (XVII). It will be noted that arrangements A and B are almost identical in this projection, and as a result, on the basis of the (hol) reflections alone, the X-ray data do not give a unique solution.

However, when the (hko) reflections are considered it becomes quite obvious that arrangement B is unacceptable and orientation A is consistent with both the (hol) and (hko) structure factors.

Structure A was refined by means of successive difference syntheses. The refinement process was accomplished very rapidly at this point and the discrepancy R fell in the sequence R = 31.6% —— 19.6% ——— 13.2%, at which stage no further refinement/

refinement was attempted.   All the atoms in the unit cell
(carbon, hydrogen and oxygen) were included in the structure
factor calculations and a Debye-Waller factor $B = 8.5 \times 10^{-16}$ cms.^{2}
was employed.

Only 22 (hol) reflections out of a possible total of 63
in the Cu - Kα sphere of reflection were observed and estimated.
This small percentage of observed reflections is due to two
effects.   First, the large amount of overlapping of atoms in the
projection and secondly, the apparently high temperature factor
will both cause the structure factors of planes with small spacings
to be very small and hence the X-ray beams reflected from these
planes will be so weak that they will be unobserved on the
photographs.

Because of this, it should be expected that large errors might
be associated with coordinates derived from this projection, and
this appears to be borne out by the observation that in this
projection, moving an atom quite a considerable distance in one
direction can be compensated for in the difference syntheses by
movements of neighbouring atoms so that the overall discrepancy is
varied by only a small amount.

Even so, because of the good agreement between observed and
calculated structure factors the structure appears to be
essentially correct.

The final electron density map obtained by projecting the
structure on to the (010) plane is shown in figure(XVII) and the
atomic coordinates obtained from the projection are given in
table (VII).

Fig. XVIII
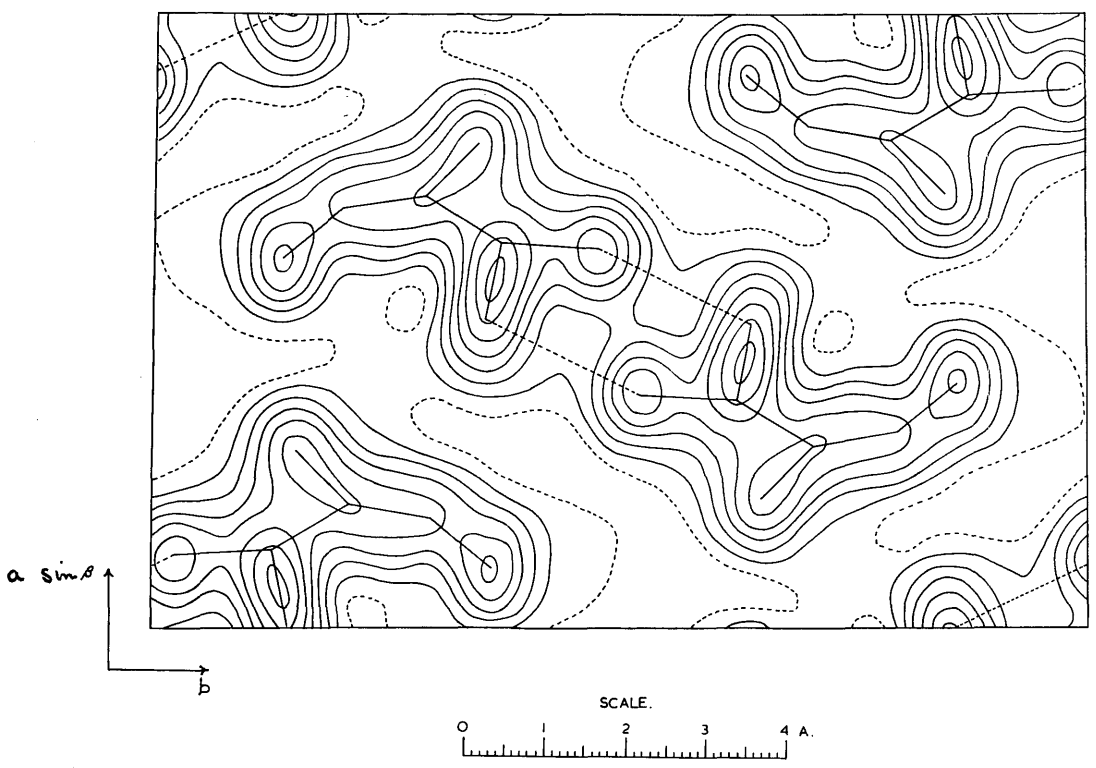
Angelic acid. c - axis projection. Final electron density map.
Contours at intervals of $1e/A^2$. $1e/A^2$ contour dotted.

Projection of the Structure on to the plane perpendicular to
the c-axis of the unit cell.

The atomic coordinates $(x_p$ and $z_p)$ obtained from the b-axis
projection should not be expected to be very accurate but they
should be accurate enough to provide an approximate y_p coordinate
to be used in a model set up for refinement of the c-axis
projection.

Such a model was obtained and refined by successive difference
syntheses. The refinement process proceeded in the sequence
R = 39.9% — 31.4% — 27.8% — 24.2% — 22.3% — 19.9%. No
further refinement was attempted because it was found that the
calculated structure factors were insensitive to movements of up
to 0.1A. in the coordinates of C_2 and it was felt that the small
percentage of observed reflections (39 independent reflections out
of a possible 112) was insufficient to justify any claim of
higher accuracy.

The temperature factor used was B = 8.5 $\times 10^{-16}$ cms.^{2} for all
atoms except C_4 for which a value of 3.5 x 10^{-16} cms.^{2} was used.

The final electron density map of the c-axis projection is
shown in figure (XVIII). The calculated and observed structure
factors are quoted in table (VIII).

When the high temperature factor is considered together with
the facts that angelic acid crystals melt at 45°C., are extremely
volatile, and, in general, are very distorted, it would appear
that at room temperatures the crystal lattice is tending to break
down/

down and disorder is setting in.   The result is that only
reflections from planes of low order spacings are observed.   If
the disorder is increased by addition of a small amount of
impurity to the acid then the melting point falls sharply.   This
explains why impure acids of this type tend to exist as oils
and will only solidify if the impurity is removed.

Coordinates, Molecular Dimensions and Orientation of the molecule in the unit cell.

The final coordinates obtained by projection of the structure down the b and c-axes respectively, are quoted in table (vii).

These two projections were refined independently of one another and since errors in atomic coordinates of up to 0.1A. can be expected in both cases, the sets of coordinates obtained from the projections are given independently.

(x_p, y_p, z_p) are the fractional coordinates of atom P. (X_p, Y_p, Z_p) are the same coordinates expressed in Angstom units, $(X_p$ is obtained from the average value of x_p from the two independent measurements) referred to the frame of the monoclinic unit cell.

(X'_p, Y'_p, Z'_p) are coordinates of P referred to a rectangular frame in which the unit vectors OY' and OZ' are coincident with the unit cell translations b and c and the vector OX' lies in the (010) plane at right angles to the plane Y'OZ'.

The transformations from the monoclinic to the rectangular frame are

$$X'_p = X_p \sin \beta$$
$$Y'_p = Y_p$$
$$Z'_p = Z_p + X_p \cos \beta .$$

Table (VII)

Fractional Coordinates

<table>
<tr><td colspan="3">From b-axis projection</td><td colspan="3">From c-axis projection</td></tr>
<tr><th>Atom</th><th>x_p</th><th>z_p</th><th>Atom</th><th>x_p</th><th>y_p</th></tr>
<tr><td>C_1</td><td>0.1257</td><td>0.0913</td><td>C_1</td><td>0.1000</td><td>0.3633</td></tr>
<tr><td>C_2</td><td>0.1667</td><td>0.3100</td><td>C_2</td><td>0.1833</td><td>0.3000</td></tr>
<tr><td>C_3</td><td>0.2127</td><td>0.3790</td><td>C_3</td><td>0.2080</td><td>0.2103</td></tr>
<tr><td>C_4</td><td>0.0967</td><td>0.1833</td><td>C_4</td><td>0.1330</td><td>0.1273</td></tr>
<tr><td>C_5</td><td>0.3083</td><td>0.5427</td><td>C_5</td><td>0.2943</td><td>0.1527</td></tr>
<tr><td>O_6</td><td>0.0300</td><td>0.0467</td><td>O_6</td><td>0.0033</td><td>0.1447</td></tr>
<tr><td>O_7</td><td>0.1500</td><td>0.2117</td><td>O_7</td><td>0.1230</td><td>0.0240</td></tr>
</table>

Coordinates in A.

<table>
<tr><th>Atom</th><th colspan="3">Monoclinic Coords.</th><th colspan="3">Rectangular Coords.</th></tr>
<tr><th></th><th>X_p</th><th>Y_p</th><th>Z_p</th><th>X'_p</th><th>Y'_p</th><th>Z'_p</th></tr>
<tr><td>C_1</td><td>0.865</td><td>4.214</td><td>0.609</td><td>0.852</td><td>4.214</td><td>0.459</td></tr>
<tr><td>C_2</td><td>1.341</td><td>3.480</td><td>2.068</td><td>1.320</td><td>3.480</td><td>1.835</td></tr>
<tr><td>C_3</td><td>1.612</td><td>2.440</td><td>2.528</td><td>1.587</td><td>2.440</td><td>2.248</td></tr>
<tr><td>C_4</td><td>0.880</td><td>1.477</td><td>1.223</td><td>0.867</td><td>1.477</td><td>1.070</td></tr>
<tr><td>C_5</td><td>2.308</td><td>1.771</td><td>3.620</td><td>2.273</td><td>1.771</td><td>3.219</td></tr>
<tr><td>O_6</td><td>0.128</td><td>1.679</td><td>0.312</td><td>0.126</td><td>1.679</td><td>0.289</td></tr>
<tr><td>O_7</td><td>1.046</td><td>0.278</td><td>1.412</td><td>1.030</td><td>0.278</td><td>1.231</td></tr>
</table>

Molecular Dimensions.

The standard deviation from the mean of the two independent measurements of X_p is 0.08A. and the largest deviation is 0.14A.. This means, therefore, that a standard deviation of about 0.12A. should be associated with measured bond lengths and differences between observed and expected bond lengths of two and a half times this will have no significance.

The standard deviation in bond angles is 10°.

The observed bond lengths and bond angles are
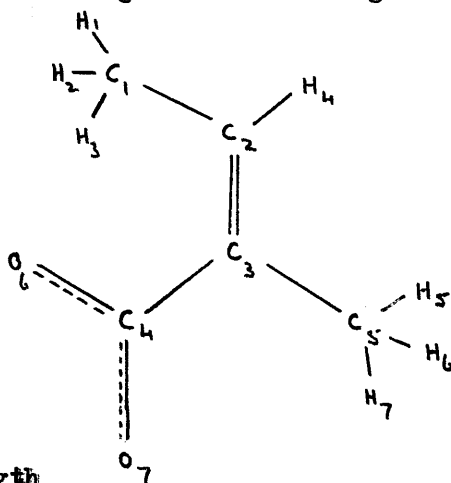
Bond	Bond Length		Bond Angles	
$C_1 - C_2$	1.63A		$C_1C_2C_3$	140°
$C_2 = C_3$	1.16A		$C_2C_3C_4$	99°
$C_3 - C_4$	1.69A		$C_2C_3C_5$	142°
$C_3 - C_5$	1.36A		$C_4C_3C_5$	116°
$C_4 - O_6$	1.10A		$C_3C_4O_6$	132°
$C_4 - O_7$	1.22A		$C_3C_4O_7$	114°
			$O_6C_4O_7$	111°

The equation of the plane through atoms C_2, C_3, C_4, C_5, O_6, and O_7 is

$$0.8154\ X' + 0.0669\ Y' - 0.5742\ Z' - 0.1675 = 0 \qquad \ldots\ldots\ldots(i)$$

The distance (d) of the atoms from this plane are

Atom	C_1	C_2	C_3	C_4	C_5	O_6	O_7
d(A)	0.55	0.09	–	0.02	−0.04	−0.12	0.02

The hydrogen bonded O.............O distance in the dimers is 2.73A.. Hence it follows that the molecule has essentially the shape expected from the chemical formula with one important difference. C_1 does not lie in the plane(1) of the rest of the molecule. Bond $C_2 - C_1$ has rotated about $C_2 = C_3$ as axis, in a clockwise direction until C_1 is 0.55A. from the plane (1) and although large errors are associated with the atomic coordinates this rotation is highly significant.

The consequences of this rotation of bond $C_1 - C_2$ are rather interesting. It means that if the sp^2 hybridisation at carbon atoms C_2 and C_3 is preserved, and the axes of the π-orbitals are still perpendicular to the planes of the sp^2 orbitals, then in angelic acid, the axes of the π-orbitals of C_2 and C_3 are not parallel to one another. Hence the overlap of these π-orbitals is not as extensive in the planar system $\begin{matrix} \diagdown \\ \diagup \end{matrix} C = C \begin{matrix} \diagup \\ \diagdown \end{matrix}$ of the normal double bond.

This probably accounts, at least to some extent, for the differences in the relative stabilities of tiglic and angelic acids and of their derivatives.

If/

If some mechanical stress is exerted on C_1 so that $C_1 - C_2$ is caused to rotate even further (e.g. by overcrowding of C_1 due to a large group attempting to attach itself to the carboxyl group) or, if the distortion of the $\diagdown C = C \diagup$ system is increased by attempted substitution, then the resultant overlap of the π-orbitals may be so small that effectively the π-electrons will be localised on C_2 and C_3. Free rotation will then take place about $C_2 - C_3$ ~~and~~ and a mixture of tiglic acid and angelic acid(or derivatives of these) will be obtained with the tiglic acid (or tiglic acid derivative) predominating.

The ready isomerization of angelic acid(and its derivatives) to tiglic acid (and its derivatives) is observed in practice. (39).

Molecular Orientation.

The molecular orientation can be defined by stating
a) the mean plane through atoms C_2, C_3, C_4, C_5, O_6, and O_7 and
b) fixing the position of a line in the molecule in this plane.

The equation of the plane has already been given and it is very nearly, although not quite, the $(20\bar{1})$ plane. The orientation of the molecule in this plane will be defined by stating the angle between $C_3 - C_4$ and the positive direction of the b-axis of the monoclinic unit cell. This angle is found to be 55°.

Conclusion.

The atomic coordinates obtained from the b and c axial projections respectively, are consistent with one another and/with the chemical formulation 2-methyl-cis-2-butenic acid assigned to angelic acid. The packing of the molecules in the unit cell is reasonable. The hydrogen bonded — O.....H – O — distance is 2.73A., and the closest distances of approach of non-bonded atoms of neighbouring dimers are all of the expected dimensions.

However, these results are not consistent with the a-axis projection orginally obtained by considering the Fourier transform of the idealised dimer.

There are two ways of explaining this inconsistency.

(1)   The projections down the b and c axes are completely wrong and the a axis projection is correct.

(2)   The projections down the b and c axes are correct and the a axis projection is wrong.

There are a number of reasons for assuming that explanation (2) is correct. It explains why the structure based on the a axis projection could not be refined when the other projections were considered. Combination of the b and c axial projections result in a structure which, if allowance is made for the inaccuracies involved, is both chemically and crystallographically reasonable. In these projections there are no serious discrepancies between the observed and calculated values of the structure factors. Furthermore, if the b and c axial projections were completely wrong then there would be no agreement between the observed and calculated structure/

Fig XIX.
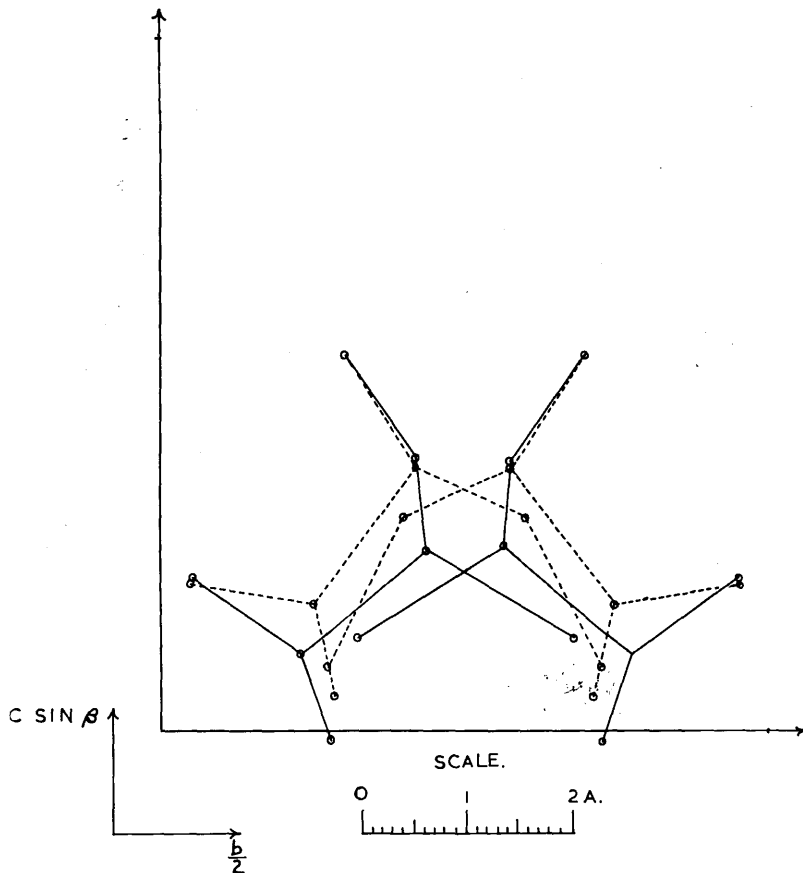
Angelic acid. a-axis projection.

—— = original interpretation of projected electron
density obtained from the Fourier Transform.

- - - = projected structure expected on the basis
of the b and c axial projections.

structure factors for the general (hkl) planes. A comparison of
the observed and calculated (hlλ) structure factors shows a
general agreement between the two sets of values and when it is
memembered that the standard deviation of the atomic coordinates
obtained from the b and c axial ~~reflections~~ projections is of the order 0.1A.,
then the agreement is very reasonable.

Hence, it follows that the b and c axial projections must be
essentially correct but the interpretation of the electron
density projection obtained on looking down the a-axis is wrong.

Figure (XIX) compares the original interpretation of this
projection with that expected on the basis of the b and c axial
projections. It will be observed that the alternative explanations
of the projected electron density have much in common. It is
thought that the wrong interpretation has arisen because the true
structure of the dimer is somewhat different from the idealised
dimer on which the Fourier transform was based. As a result,
a number of incorrect phase angles have been obtained from the
transform and the refinement process involving combination of
these incorrect phase angles with the structure factor moduli has
biased the refinement towards a pseudohomometric constellation which
has much in common with the correct structure.

The a-axis projection will, therefore, require to be
reinvestigated.

Table (viii) lists the observed and calculated structure factors
of the (hol), (hko), (hlλ), reflections. The atomic coordinates
obtained on projection down the a-axis and the calculated and
observed structure factors based on these coordinates are also given.

<u>Observed ($|F_o|$) and Calculated (F_c) Structure Factors of</u>
<u>Angelic Acid.</u> (Table VIII).

<u>h + k = 2n</u>

$$F(hkl) = F(h\bar{k}l)$$
$$F(\bar{h}kl) = F(hk\bar{l}) \neq F(hkl)$$

<u>h + k = 2n + 1</u>

$$F(hkl) = -F(h\bar{k}l)$$
$$F(\bar{h}kl) = -F(hk\bar{l}) \neq F(hkl)$$

$$F(hkl) = 0 \text{ when } h = l = 0 \text{ or } k = 0.$$

F(hol)

Plane	$2\sin\theta$	$|F_o|$	F_c	Plane	$2\sin\theta$	$|F_o|$	F_c
200	0.420	10.9	$\overline{13.1}$	$40\bar{1}$	0.830	13.8	$\overline{13.8}$
201	0.515	17.0	$\overline{17.7}$	$40\bar{2}$	0.890	26.2	27.1
202	0.685	< 1.3	0.1	$40\bar{3}$	0.990	8.8	10.5
203	0.880	4.0	$\overline{1.9}$	$40\bar{4}$	1.140	3.6	2.8
204	1.090	4.6	4.4	600	1.260	< 1.3	4.6
$20\bar{1}$	0.450	117.8	117.0	601	1.315	< 1.3	0.9
$20\bar{2}$	0.580	25.4	26.8	602	1.413	4.3	2.6
$20\bar{3}$	0.760	6.5	8.0	$60\bar{1}$	1.240	3.8	1.6
$20\bar{4}$	0.965	< 1.3	0.3	$60\bar{2}$	1.260	6.9	$\overline{8.2}$
$20\bar{5}$	1.180	< 2.8	2.0	$60\bar{3}$	1.330	< 1.3	4.7
400	0.840	12.5	$\overline{11.3}$	$60\bar{4}$	1.430	< 3.0	2.7
401	0.910	4.1	5.5	001	0.235	21.7	18.3
402	1.030	< 1.3	0.1	002	0.470	< 1.3	6.2
403	1.190	6.6	4.5	003	0.705	< 1.3	$\overline{1.1}$
404	1.360	< 1.3	$\overline{2.0}$	004	0.940	7.5	2.6
				005	1.180	< 2.8	2.9

$$R = \frac{100\left\{\Sigma\left||F_o| - |F_c|\right|\right\}}{\Sigma|F_o|}$$

$$= 13.2\%$$

F(hko)

Plane	$2\sin\theta$	$\lvert F_o\rvert$	F_c	Plane	$2\sin\theta$	$\lvert F_o\rvert$	F_c
020	0.267	24.2	$\overline{24.3}$	2100	1.385	2.7	$\overline{2.7}$
040	0.530	26.7	$\overline{31.4}$	2110	1.520	< 2.6	$\overline{4.0}$
060	0.800	19.5	26.0	310	0.625	5.1	1.7
080	1.060	13.3	9.9	320	0.670	11.1	6.0
0100	1.325	5.4	$\overline{2.6}$	330	0.730	18.2	$\overline{22.4}$
110	0.245	31.5	26.8	340	0.810	5.0	$\overline{8.0}$
120	0.340	24.0	$\overline{25.4}$	350	0.905	–	$\overline{0.3}$
130	0.450	5.6	8.3	360	1.000	–	$\overline{2.4}$
140	0.570	–	$\overline{0.6}$	370	1.110	4.9	2.0
150	0.700	2.0	2.8	380	1.225	–	$\overline{1.0}$
160	0.820	–	1.4	390	1.340	3.1	$\overline{4.0}$
170	0.955	7.8	9.1	400	0.820	9.1	$\overline{9.7}$
180	1.085	3.0	$\overline{4.3}$	410	0.830	4.0	4.3
200	0.410	9.5	$\overline{11.1}$	420	0.865	9.6	$\overline{12.2}$
210	0.430	35.3	$\overline{37.7}$	430	0.915	10.0	$\overline{8.0}$
220	0.490	18.6	19.9	440	0.980	2.2	$\overline{2.4}$
230	0.575	14.9	$\overline{12.3}$	450	1.060	7.1	5.0
240	0.675	12.8	$\overline{14.1}$	460	1.145	–	–
250	0.785	2.4	0.8	470	1.240	–	0.5
260	0.895	< 2.6	1.1	480	1.340	3.9	$\overline{4.1}$
270	1.015	–	$\overline{1.6}$	490	1.450	4.8	$\overline{2.4}$
280	1.135	4.4	4.5	510	1.030	5.3	2.7
290	1.265	6.1	$\overline{5.5}$	520	1.060	2.9	3.7

$$R = \frac{100\{\Sigma\lvert\,\lvert F_o\rvert - \lvert F_c\rvert\}}{\Sigma\lvert F_o\rvert}$$

$$= 19.9\%$$

F(h1ℓ)

Plane	$2\sin\theta$	$|F_o|$	F_c	Plane	$2\sin\theta$	$|F_o|$	F_c
011	0.270	55.6	$\overline{59.3}$	210	0.430	34.1	$\overline{43.4}$
012	0.478	15.8	$\overline{26.0}$	211	0.520	11.7	$\overline{8.0}$
013	0.715	9.6	$\overline{9.2}$	21$\bar{1}$	0.450	31.5	$\overline{27.2}$
014	0.955	< 4.8	$\overline{10.4}$	21$\bar{2}$	0.575	16.9	8.3
110	0.240	23.9	22.6	310	0.625	< 4.9	1.6
111	0.360	5.2	$\overline{0.1}$	311	0.700	7.9	$\overline{7.8}$
112	0.565	17.1	$\overline{17.4}$	312	0.840	8.9	12.6
113.	0.780	20.4	7.2	31$\bar{1}$	0.625	15.3	12.8
114	1.000	6.5	10.0	31$\bar{2}$	0.715	13.5	23.7
11$\bar{1}$	0.315	27.7	39.9	410	0.850	4.8	0.4
11$\bar{2}$	0.495	5.6	0.7	41$\bar{1}$	0.825	10.7	$\overline{11.3}$
11$\bar{3}$	0.715	11.4	15.8	41$\bar{2}$	0.880	20.3	$\overline{16.4}$
11$\bar{4}$	0.935	8.7	$\overline{1.7}$				

Atomic Coordinates and Structure Factors obtained from incorrect
interpretation of the projection down the a—axis.
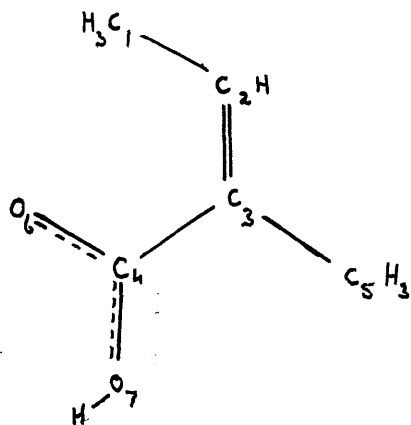
Atom	y_p	z_p
C_1	0.1513	0.5449
C_2	0.2113	0.3903
C_3	0.2177	0.2667
C_4	0.1155	0.1107
C_5	0.3400	0.1333
O_6	0.0253	0.2243
O_7	0.1394	0.0137

F(okl)

Plane	2sinθ	$|F_o|$	F_c	Plane	2sinθ	$|F_o|$	F_c
011	0.270	56.6	$\overline{59.1}$	060	0.800	20.4	21.3
012	0.490	13.5	$\overline{10.7}$	061	0.835	3.8	5.1
0$\overline{1}$3	0.725	9.8	4.9	062	0.930	5.9	6.0
014	0.950	2.4	$\overline{2.0}$	063	1.075	5.4	$\overline{4.3}$
015	1.185	4.9	5.0	065	1.430	3.3	2.4
020	0.270	24.1	$\overline{34.2}$	072	1.045	4.5	$\overline{2.1}$
021	0.360	16.3	17.5	073	1.170	5.2	5.2
022	0.545	20.4	$\overline{14.3}$	074	1.325	2.9	2.8
023	0.765	10.5	$\overline{12.4}$	080	1.065	12.6	12.4
024	0.985	6.7	6.4	081	1.090	8.2	8.2
031	0.465	7.5	6.2	082	1.165	< 3.3	3.5
032	0.620	27.2	$\overline{27.0}$	083	1.280	< 2.6	1.0
033	0.820	4.1	3.5	091	1.220	4.8	$\overline{3.7}$
040	0.535	29.4	$\overline{25.6}$	092	1.285	5.8	$\overline{6.0}$
041	0.590	31.2	$\overline{28.0}$	0100	1.335	5.1	$\overline{3.6}$
042	0.715	35.5	$\overline{38.9}$	0101	1.355	3.4	$\overline{2.6}$
043	0.895	3.7	$\overline{0.7}$	0102	1.415	5.1	$\overline{5.1}$
044	1.085	4.5	4.5	0103	1.520	< 2.1	0.9
051	0.705	9.2	$\overline{14.0}$	001	0.240	21.7	24.3
052	0.815	12.5	14.9	002	0.475	< 2.1	$\overline{2.8}$
053	0.980	3.8	5.4	003	0.710	< 2.1	$\overline{1.5}$
054	1.160	5.2	3.8	004	0.940	7.0	7.7
055	1.360	7.2	$\overline{9.4}$	005	1.180	< 2.1	5.7

$$R = \frac{100\left\{\Sigma \left| |F_o| - |F_c| \right| \right\}}{\Sigma |F_o|} = 15.8\%$$

THE CRYSTAL AND MOLECULAR STRUCTURE

OF

TETRABENZO-2:3-6:7-2':3'-6':7'-HEPTAFULVALENE.

I

The Crystal and Molecular Structure of

Tetrabenzo-2:3-6:7-2':3'-6':7'-heptafulvalene.

1. Introduction.

As a result of much accurate work carried out on the structures of organic compounds by means of the methods of X-ray and electron diffraction analysis, coupled with parallel progress in theoretical calculations, it has become possible to predict with fair accuracy such things as bond lengths, bond angles, charge distributions and ultra-violet absorption bands of many molecules.

In particular, for conjugated condensed ring hydrocarbons, calculations by means of the molecular orbital theory have been found to be in reasonable agreement with observation.

Tetrabenzo-2:3-6:7-2':3'-6':7'-heptafulvalene (I) is such a condensed ring hydrocarbon and molecular orbital calculations have been reported for it (40).
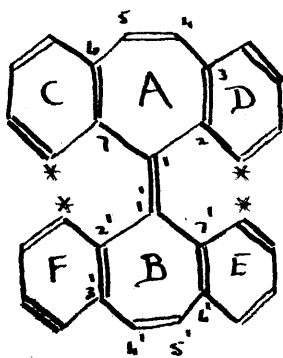
I

These calculations are based on the assumption that all the atoms lie in one plane. If, however, the atoms do not lie in the same plane then the calculations may be in serious error. The magnitude of this error will depend on how far the molecule deviates from planarity. If the deviations are marked then predictions based on the

assumption of a planar molecule may be quite wrong.

This is in fact what is found to happen in the case of tetrabenzo-2:3-6:7-2':3'-6':7'-heptafulvalene for certain predictions made on the basis of these calculations are not in accord with observation.

The reason for this is quite simple.   If a planar model is built up using bond lengths of 1.4A. and assuming the rings A and B to be regular heptagons and the rings C,D, E and F to be regular hexagons, then it is found that the centres of the starred pairs of carbon atoms approach to within 1.3A. of one another.

Such a close distance of approach of carbon atoms is impossible and so the atoms of the molecule can not be coplanar.

The object of this work is to determine the shape of the tetrabenzo-2:3-6:7-2':3'-6':7'-heptafulvalene molecule.

From now on, the compound will be referred to as T.B.H.F.

2.   Crystal Data.

T.B.H.F., $C_{30}H_{20}$:   Molecular Weight = 380.5.
Density calculated = 1.224 gms/cc.
Density observed(by flotation in potassium iodide solution)
                = 1.245 gms/cc.
a = 13.52 $\pm$ 0.02A.   b = 8.38 $\pm$ 0.02A.   c = 10.41 $\pm$ 0.02A.
α = 90° β = 118°50'$\pm$ 0.5° γ = 90°
Volume of unit cell = 1033A.^{3}
Number of molecules per unit cell = 2.
Systemmatic absences:- (hol) when h = 2n + 1.
                       (oko) when k = 2n + 1.

Space group = P 2_1/a

Absorption coefficient for X-rays (λ = 1.542A) μ = 6.4cm^{-1}.

Total number of electrons per unit cell F(000) = 400.

  Crystals were obtained in the form of small, colourless, diamond-shaped plates by refluxing the material with toluene until solution was complete and allowing the toluene to evaporate off overnight at room temperature. The diamond-shaped face is the (001) face and the <u>a</u> and <u>b</u> axes lie along the diagonals of the face.

  Harnik, Herbstein, Schmidt and Hirshfield([41]) have reported the existence of another form of T.B.H.F. with

  a = 11.3A   b = 16.6A   c = 11.6A

  α = 90° β = 107°30' γ = 90°

Space group = P 2_1/n.

Density calculated = 1.22 gms/cc. Density observed = 1.25 gms/cc.

4 molecules per unit cell.

  These authors call the compound Bisdibenzosuberenylidene.

  The existence of more than one form of these complex condensed ring hydrocarbons seems to be not unusual, e.g.

α and β forms of 3:4-5:6-dibenzophenanthrene ([41,42])

α and β forms of dixanthylidene ([41]).

α and β forms of difluorenylidene ([43,44]).

3. Experimental Measurements.

  Rotation, oscillation and moving-film photographic measurements have been employed. The radiation used was copper K$_\alpha$(λ = 1.542A). The cell dimensions were determined by means of rotation and equatorial layer line moving-film photographs. The reflections were estimated visually by means of the multiple film technique. Structure factors for the (h0l) reflections calculated by the usual mosaic crystal formula are given in pp(96, 97).

4. Attempts to determine the structure of T.B.H.F.

Since the space group is P 2_1/a with two molecules per unit cell, it follows that the asymmetric unit is half the molecule and the molecule possesses a centre of symmetry. The centre of symmetry must lie at the mid-point of the bond joining the seven-membered rings, and the asymmetric unit is
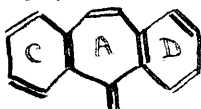
It should be expected that the most stable state of the molecule would be attained when the conjugation is a maximum and so, rather than merely have the starred atoms knock one another out of the plane of the rest of the molecule, it seemed more likely that the rings C, D, E and F would move out of the plane of rings A and B. In addition, it is unlikely that only rings C and E say, will move. It is more likely that at the overcrowded positions, one ring will be twisted up whilst the other is twisted down. If ring C is up out of the plane of rings A and B then ring F will be twisted downwards. Hence, since the molecule is centrosymmetrical, the plane of ring D will be up and the plane of ring E will be down from the plane of rings A and B, i.e. the molecule may have a shape similar to that shown below.
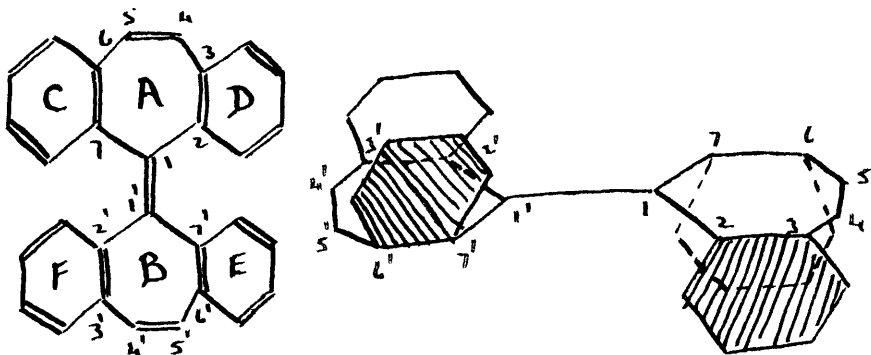
The next step forward in the investigation was an attempt to fit a model of this sort into the unit cell in such an orientation as to fit the observed data.

It was assumed, on analogy with results obtained for dibiphenylene-ethylene (II) (43), that the atoms at the overcrowded positions would have to be at least 2.5A apart -

II
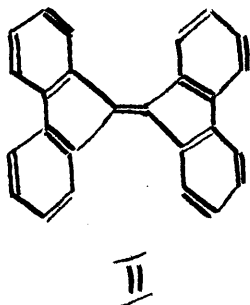
possibly further - and so, for the trial structure it was postulated that in the asymmetric unit consisting of rings C, A and D, ring A was planar and rings C and D were folded up out of the plane of ring A and at angles of about 45° to it.

In spite of a great deal of work carried out in trying to obtain the correct crystal structure based on such a model, no satisfactory structure could be obtained.

All structures based on this model gave very large discrepancies $\left(R = \dfrac{\Sigma||Fo| - |Fc||}{\Sigma|Fo|} > 65\% \right)$.

Then it was noticed that the ultra-violet absorption spectra of tetraphenyl-ethylene and T.B.H.F. were reported to be very similar (40).
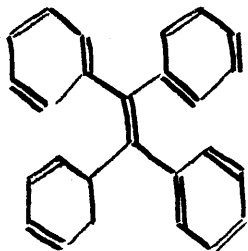
Tetraphenyl - Ethylene.
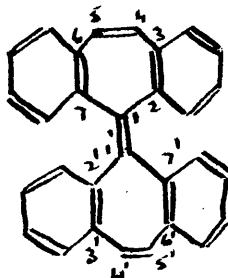
T·B·H·F.

The ultra-violet absorption spectra are defined by the
electronic energy levels of the compounds.   These energy levels
are defined, to the first approximation, by the extent of the
conjugation of the π-electrons of the molecule.   If the
ultra-violet spectra are similar then the π-electron conjugations
must be similar.   Tetraphenylethylene is just T.B.H.F. minus atoms
4, 5, 4' and 5'.   Consequently, in T.B.H.F., atoms 4, 5, 4' and 5'
can not be adding a large contribution towards the resonating
structure of the remaining π-electrons.

In tetraphenyl ethylene, the steric strain due to overcrowding
can be overcome by simply rotating the six membered rings about
axes along the bonds joining them to the central "ethylenic" system,
and so it seems reasonable to assume that something of this sort
will take place in T.B.H.F.

At this stage, the original trial structure was abandoned.

Inspection of the calculated structure factors of the trial
structures using the rejected molecular model revealed one structure
where the calculated values of F(ool) agreed reasonably well with
the scaled observed structure factors for these planes.   This
suggested that the z_p coordinates of this structure were correct but
did not indicate whether any particular coordinate had a positive
or negative value nor did it indicate which atom should be associated
with a particular value of z_p.
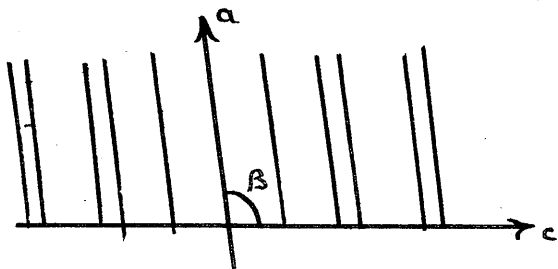
Consequently, on a scale diagram
of the (010) plane, a grid consisting
of lines drawn parallel to the
a-axis and passing through the z_p
coordinates of this trial structure
was drawn up.

The asymmetric unit consisting of rings A, C and D was then divided up into four parts

(i) Ring C   (ii) Ring D   (iii) Atom 1 of ring A
          (iv) Atoms 4 and 5 of Ring A.

The structure factors of the planes $(40\bar{1})$, (400), (401) and (402) are 32.0, 43.0, 5.8 and 22.6 respectively and because of the sudden fall in the value of $|F(401)|$ as compared with the others, it was felt that if a structure could be obtained which yielded calculated F values for these planes in this sequence, as well as accounting satisfactorily for the (00ℓ) planes, then this might prove to be the correct structure.

Experience with the numerous trial structures already considered showed an orientation of the molecule which yielded calculated F values for the above mentioned (40ℓ) planes in the correct sequence. On placing the four parts of the asymmetric unit on to the grid already discussed and sliding them into the approximate orientation to give calculated values of $F(40\ell)$ in agreement with the observed values, a feasible trial structure was obtained.   This gave a discrepancy of 49%.

This may or may not prove to be the correct structure.

The extent of the buckling is not certain but the broad features of this postulated molecular structure are as shown in figure (XX). The puckering of the seven membered rings in the diagram is very much over-exaggerated to make the distortions clear.

The main features of this structure are

Fig. xx.

Postulated Structure of T.B.H.F.

a)   There is a central ethylenic system, the atoms of which are coplanar or very nearly so.   This system is made up of atoms 1, 2, 7, 1', 2', 7'.

b)   The six membered rings are turned about axes along the bonds which attach them to the ethylenic system (a) until their planes are at about 40° to that of the plane 1, 2, 7, 1', 2', 7'.

c)   This results in atoms 2, 3, 6, 7 of ring A being approximately coplanar and the planes of rings C and D lie at about 40° to this plane.   Plane 2, 3, 6, 7 make an angle of about 40° with the plane 1, 2, 7, 1', 2', 7'.

d)   Atoms 3, 4, 5 and 6 lie approximately in one plane in such a way that 4 and 5 bridge across the space between 3 and 6 which is left as a consequence of operations (a), (b) and (c).   Rings C and D therefore, lie on opposite sides of the plane 2, 3, 6, 7 from atoms 1, 4 and 5.   Hence the seven-membered rings are puckered in a manner analogous to that of the boat form of the cyclohexane molecule.

This process means that bonds 1-1', 4-5, 4'-5' must be true double bonds.   Bromine does add on to this molecule although the number of atoms of bromine which do add to the structure is not known.

It is also interesting to note that in this structure there is no steric strain due to distortion of bond angles.   Neither is there distortion of the sp^2 bond angles of the six membered rings nor is there any associated with the ethylene systems.

There is no overcrowding of atoms.   The closest distance of approach of non-bonded carbon atoms is 2.8A and is between atoms 2 and 7'.   The next closest distance of approach is between atoms 2 and 15' which are about 3A apart.

The structure can be considered to be tetraphenyl ethylene plus atoms 4, 5, 4', 5'. Hence the similarity reported for the ultra-violet spectra of the two compounds.

It is interesting to note that this postulated structure is very similar to that found by Harnik and Schmidt (45) for dianthronylidene. In this compound, "A molecular model has been obtained approximating to symmetry 2/m "in which are centre to centre distance of 2.9A between the "overcrowded carbon atoms 4, 8' and 8, 4' is achieved by a 40° "rotation of the benzene rings out of the plane of the central "ethylenic system. This rotation is accompanied by a compression "of the $\overset{c}{\diagdown}\underset{c}{\diagup}c=$ angle to 113° and by
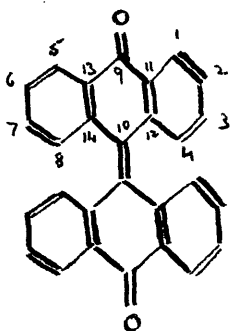
"deflection of the exocyclic bonds "to the carbonyl carbon by 8° out "of the planes of their respective "benzene rings".

Both dianthronylidene and T.B.H.F. have four six-membered rings and a central ethylenic system in common. They differ only in the nature of the group spanning the space between the six-membered rings. The postulated model of T.B.H.F. was not based on Harnik and Schmidt's work. It is a model which has emerged in the process of putting together parts of a molecule in such a way as to account for the experimentally observed structure factors.

Although at the moment there is no direct proof that the model proposed for the T.B.H.F. molecule is correct all the above pieces

dianthronylidene

of evidence, which by themselves do not mean very much, all fit in with it.   Consequently, the proposed structure may possibly be correct and it may be well worth while attempting to refine this structure further.

The fractional coordinates of the carbon atoms in this projection are given in table (IX).   A comparison of the calculated structure factors with those observed for the (hol) reciprocal lattice net of T.B.H.F. are given on pages(96-97).

Fractional Atomic Coordinates for the Atoms of the Asymmetric Unit
of T.B.H.F.

Table IX.
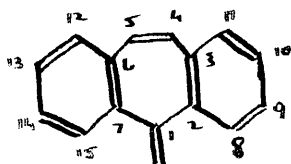
Atom	x_p	y_p
C_1	0.0500	0.0333
C_2	0.0967	0.1667
C_3	0.2100	0.2866
C_4	0.2934	0.3133
C_5	0.2984	0.2233
C_6	0.2400	0.1000
C_7	0.1333	0.0333
C_8	0.0091	0.2083
C_9	0.0334	0.3533
C_{10}	0.1501	0.4667
C_{11}	0.2400	0.4333
C_{12}	0.2617	0.0667
C_{13}	0.1733	$\overline{0.0333}$
C_{14}	0.0449	$\overline{0.1033}$
C_{15}	0.0234	$\overline{0.0667}$

Observed ($|Fo|$) and Calculated (Fc) Structure Factors of T.B.H.F.

F(hol)

Plane.	$2\sin\theta$..	$|Fo|$	Fc	Plane.	$2\sin\theta$.	$|Fo|$	Fc
200	0.255	7.2	-17	$40\bar{1}$	0.460	32.0	34
201	0.365	11.2	11	$40\bar{2}$	0.460	18.8	-15
202	0.510	5.8	8	$40\bar{3}$	0.510	13.3	-7
203	0.660	22.0	4	$40\bar{4}$	0.610	21.0	-15
204	0.820	18.1	-4	$40\bar{5}$	0.735	25.0	-25
205	0.980	9.1	12	$40\bar{6}$	0.865	11.6	9
206	1.135	8.9	14	$40\bar{7}$	1.010	13.0	2
207	1.305	13.8	-18	$40\bar{8}$	1.160	11.3	-4
$20\bar{1}$	0.230	15.4	31	$40\bar{9}$	1.290	10.7	17
$20\bar{2}$	0.300	8.3	-8	$40\overline{10}$	1.480	8.6	14
$20\bar{3}$	0.435	2.4	-4	$40\overline{11}$	1.630	5.0	–
$20\bar{4}$	0.580	12.6	-21	600	0.775	6.0	3
$20\bar{5}$	0.730	18.6	22	601	0.870	7.4	7
$20\bar{6}$	0.895	15.0	10	602	0.980	8.5	15
$20\bar{7}$	1.055	10.5	–	603	1.105	10.0	14
$20\bar{8}$	1.215	13.3	14	604	1.235	17.0	14
$20\bar{9}$	1.380	5.7	2	605	1.385	5.2	9
400	0.510	43.0	54	$60\bar{1}$	0.710	4.0	-10
401	0.615	5.8	-4	$60\bar{2}$	0.680	8.9	-20
402	0.735	22.6	-14	$60\bar{3}$	0.685	4.0	2
403	0.875	–	14	$60\bar{4}$	0.735	25.4	21
404	1.015	–	-4	$60\bar{5}$	0.815	18.0	-2
405	1.165	13.0	-4	$60\bar{6}$	0.915	12.1	11
406	1.320	6.2	-29	$60\bar{7}$	1.030	4.4	-12
407	1.480	10.8	-17	$60\bar{8}$	1.160	4.5	10
				$60\bar{9}$	1.300	5.0	6

97.

Plane	2sin θ.	|Fo|	Fc
800	1.030	5.2	7
801	1.125	4.8	2
802	1.230	5.8	17
803	1.345	6.0	-1
804	1.475	10.8	-3
805	1.610	5.0	1
80$\bar{1}$	0.965	4.4	-2
80$\bar{2}$	0.915	13.6	-10
80$\bar{4}$	0.920	8.6	-9
80$\bar{5}$	0.955	9.8	5
80$\bar{6}$	1.025	11.8	3
80$\bar{9}$	1.340	8.6	9
80$\bar{10}$	1.460	5.4	6
80$\bar{11}$	1.595	8.1	7
80$\bar{13}$	1.880	3.4	-3
1001	1.380	5.4	-6
1003	1.595	5.0	-4
100$\bar{3}$	1.135	4.9	-3
100$\bar{4}$	1.130	16.0	10
100$\bar{5}$	1.145	9.1	10
100$\bar{7}$	1.245	9.6	7
100$\bar{10}$	1.525	10.2	13
100$\bar{12}$	1.760	7.9	12

Plane.	2sin θ.	|Fo|	Fc
1200	1.550	10.7	-4
120$\bar{1}$	1.580	10.9	-4
120$\bar{2}$	1.420	5.4	7
120$\bar{4}$	1.355	7.2	-5
120$\bar{5}$	1.355	7.2	-9
120$\bar{6}$	1.370	8.6	-1
120$\bar{10}$	1.620	5.0	-4
140$\bar{5}$	1.580	10.5	-8
140$\bar{7}$	1.605	10.0	-1
140$\bar{8}$	1.635	5.4	-4
140$\bar{9}$	1.690	5.7	-12
160$\bar{6}$	1.810	8.1	-4
160$\bar{8}$	1.840	–	3
001	0.165	46.4	75*
002	0.325	37.0	36
003	0.495	9.2	21
004	0.660	17.3	17
005	0.825	–	-3
006	0.995	–	-2
007	1.555	5.5	-7

* This may be due to extinction.

A MOLECULAR ORBITAL STUDY OF PTERIDINE

AND

A CORRELATION CURVE FOR CARBON-NITROGEN BONDS.

A Molecular Orbital Study of Pteridine and a Correlation Curve for Carbon-Nitrogen Bonds.
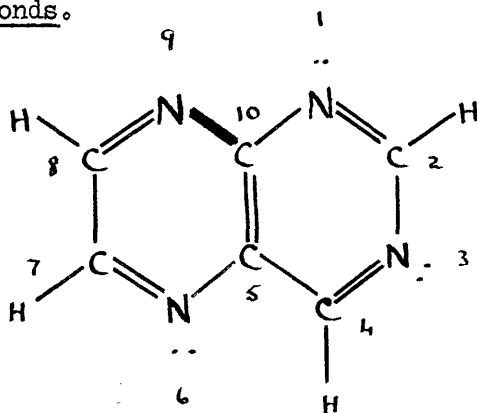
1. Introduction

There are a number of experimental techniques which can be employed in measuring bond lengths and each one of them shows that bonds of a similar type between any two given atoms are more or less constant in length.

The transition in bond type is gradual in character and many organic molecules can be assigned bond orders intermediate between single and double bonds. The lengths of these bonds can be estimated experimentally by any of the usual methods and so a relationship between bond orders and bond lengths can be obtained. Hence, if a bond length between two atoms is known the bond order can be predicted.

Conversely, if the bond order is known, then the bond lengths can be estimated.

In this work, the attempt is made to calculate these bond orders for pteridine(which is formulated as above) and so predict the bond lengths to be expected.

It is the electrons which form the bonds in a molecule and so one obvious way of estimating bond orders is to estimate the

distribution of electrons in the molecule.

In pteridine, for each carbon and nitrogen atom, there are three sp^2 hybridised orbitals lying in one plane at 120° to each other. Two of these overlap with sp^2 orbitals of nearest neighbours of the condensed ring system giving rise to the axially symmetrical σ bonds between carbon and nitrogen atoms, whilst the third, in the case of carbon atoms, is linked up with the 1s orbital of a hydrogen atom, and in the case of nitrogen atoms it contains a lone pair of ~~free~~ electrons.

The remaining electron of each atom lies in the $2p_z$ orbital which is at right angles to the planar framework of the sp^2 orbitals. When $2p_z$ orbitals overlap a π -bond is formed.

The σ -electrons are usually regarded as being effectively localised whereas the π -electrons are regarded as being the property of the molecule as a whole.

Hence, the problem of calculating the distribution of electrons in the pteridine molecule now degenerates to a consideration of the distribution of the π -electrons over the molecule.

2. The Secular Equations of Pteridine.

Electrons are particles which show some of the properties of waves and in order to correlate the particulate idea of an electron with its wave-like properties it is postulated that the probability of finding an electron in space is defined by a wave function ψ where $\psi^2 d\tau$ represents the probability of finding an electron in the volume element $d\tau$ of space.

There are several ways of attempting to find this electron

distribution and in the linear combination of atomic orbitals
molecular orbital approach the π-electrons are considered to be
under the influence of all the atoms of the condensed ring system.
The molecular orbitals are then supposed to be given by

$$\psi_j = \sum_{r=1}^{10} (c_{r,j} \psi_r)$$

where $c_{r,j}$ are constants representing the contribution which the
atomic orbitals ψ_r make towards the molecular orbital ψ_j and the
normalising condition $\sum (c_{r,j})^2 = 1$ is obeyed. $c_{r,j}^2$ then
represents the contribution of π-electron j to the electron density
on atom r.

In pteridine there are ten atoms contributing to the molecular
orbital and so this linear combination of atomic orbitals leads,
in the well-known manner, to ten secular equations, the first of
which is

$$c_1(H_{1,1} - E_j S_{1,1}) + c_2(H_{1,2} - E_j S_{1,2}) + \ldots\ldots + c_{10}(H_{1,10} - E_j S_{1,10}) = 0$$

where E_j is the energy of the jth π-electron,

$H_{rs} = \int \psi_r \mathcal{H} \psi_s d\tau$ and is known as the "Exchange Integral" or
"Resonance Integral" since it is a measure of the stability
which the system attains through exchange of an electron between atoms
r and s,

$H_{rr} = \int \psi_r \mathcal{H} \psi_r d\tau$ is the "Coulomb Integral" and is the energy
which the π-electron in question would have if it were confined to
the atom r instead of resonating round the ring of the pteridine
molecule,

$S_{rs} = \int \psi_r \psi_s d\tau$ and is the "Overlap Integral". It is a measure
of the extent of the overlap of the atomic orbitals r and s,

$S_{rr} = \int \psi^2 d\tau$ is the "Normalisation Integral" and since $\psi^2 d\tau$
represents the probability of finding an electron in volume $d\tau$ of
space $S_{rr} = 1$,

now let $H_{rs} = \beta$ when r and s are neighbouring atoms no matter whether r and s are carbon or nitrogen atoms.   When r and s are not nearest neighbours H_{rs} is assumed to be so small that it can be neglected.

The value of the "Overlap Integral" S_{rs} is only about 0.25 and is neglected since by doing so the arithmetic involved is appreciably shortened.

Let the energy which the π -electron would have if it were confined to a carbon atom, be α (i.e. $H_{rr} = \alpha$ for carbon atoms) and let $(\alpha - E_j)$ be some multiple of β - say $x_j \beta$.

Let $(H_{rr} - E_j)$ = some other multiple of β - say $(x - h)\beta$ when r = nitrogen(n)

i.e. $$\frac{H_{nn} - E_j}{\beta} = \frac{\alpha - E_j}{\beta} - \frac{\alpha - H_{nn}}{\beta} = x - h$$

Therefore h depends on the value of $\alpha - H_{nn}$ which, in turn, depends on the difference in the eletronegativity of carbon and nitrogen. In this case h is assumed to be $-\frac{1}{2}$

i.e. $$H_{nn} - E_j = (x + \tfrac{1}{2})\beta$$

Slightly different values for the exchange and coulomb integrals of carbon-nitrogen bonds and nitrogen atoms respectively have been quoted in the literature (46, 47, 48) on the grounds that these give approximately the correct dipole moments for hetrocyclic molecules containing nitrogen: in this work, these figures are rounded off to ease the computation involved.

Hence, on substitution, and elimination of β , the secular equations now reduce to

(1) $\quad c_1\left(x + \tfrac{1}{2}\right) + c_2 + c_{10} = 0$

(2) $\quad c_1 + c_2 x + c_3 = 0$

(3) $\quad c_2 + c_3\left(x + \tfrac{1}{2}\right) + c_4 = 0$

(4) $\quad c_3 + c_4 x + c_5 = 0$

(5) $\quad c_4 + c_5 x + c_6 + c_{10} = 0$

(6) $\quad c_5 + c_6\left(x + \tfrac{1}{2}\right) + c_7 = 0$

(7) $\quad c_6 + c_7 x + c_8 = 0$

(8) $\quad c_7 + c_8 x + c_9 = 0$

(9) $\quad c_8 + c_9\left(x + \tfrac{1}{2}\right) + c_{10} = 0$

(10) $\quad c_1 + c_5 + c_9 + c_{10} x = 0$

3. <u>The Permitted Energy Levels for the π-electrons of Pteridine.</u>

The secular equations give rise to the 10th order secular determinant Δ .

$$
\Delta = \begin{vmatrix}
(x+\frac{1}{2}) & 1 & & & & & & & & 1 \\
1 & x & 1 & & & & & & & \\
& 1 & (x+\frac{1}{2}) & 1 & & & & & & \\
& & 1 & x & 1 & & & & & \\
& & & 1 & x & 1 & & & & 1 \\
& & & & 1 & (x+\frac{1}{2}) & 1 & & & \\
& & & & & 1 & x & 1 & & \\
& & & & & & 1 & x & 1 & \\
& & & & & & & 1 & (x+\frac{1}{2}) & 1 \\
1 & & & & 1 & & & & 1 & x
\end{vmatrix} = 0
$$

(10)

Now let $y = 2x$ and get

$$
\Delta = \begin{vmatrix}
(y+1) & 2 & & & & & & & & 2 \\
2 & y & 2 & & & & & & & \\
& 2 & (y+1) & 2 & & & & & & \\
& & 2 & y & 2 & & & & & \\
& & & 2 & y & 2 & & & & 2 \\
& & & & 2 & (y+1) & 2 & & & \\
& & & & & 2 & y & 2 & & \\
& & & & & & 2 & y & 2 & \\
& & & & & & & 2 & (y+1) & 2 \\
2 & & & & 2 & & & & 2 & y
\end{vmatrix} = 0
$$

(10)

These determinants are of the variety known as continuants of the form

$$P_n = \begin{vmatrix} y & 2 \\ 2 & y & 2 \\ & 2 & y & 2 \\ & & & \ddots \\ & & & & 2 & y & 2 \\ & & & & & 2 & y \end{vmatrix}_{(n)} = 2^n \begin{vmatrix} x & 1 \\ 1 & x & 1 \\ & 1 & x & 1 \\ & & & \ddots \\ & & & 1 & x & 1 \\ & & & & 1 & x \end{vmatrix}_{(n)} = 2^n S_n$$

with a number of decorations added.

A number of very useful properties of determinants of this type have been worked out by Goodwin and Vand(49) and by making use of these relations, Δ can be expressed in terms of unitary continuants S_n in the form

$$\Delta = 1024S_{10} + 2048S_9 - 512S_8 - 1536S_7 - 960S_6 - 2048S_5 - 6272S_4$$
$$- 6528S_3 - 3456S_2 - 3200S_1 - 4288$$
$$= 0$$

In order to solve this equation and find the roots within the range $-2 < x < +2$, Δ is now converted into a series of multiple cosines by making the substitution $x = 2\cos\theta$ which gives rise ($\leqslant 0$) to the formula
$$S_n = 2\left[\cos n\theta + \cos(n-2)\theta + \cos(n-4)\theta + \ldots\ldots\ldots + (\cos\theta \text{ or } \tfrac{1}{2})\right]$$
according as n is odd or even.
This gives

$$\Delta = 2048\cos10\theta + 4096\cos9\theta + 1024\cos8\theta + 1024\cos7\theta - 896\cos6\theta$$
$$- 3072\cos5\theta - 13440\cos4\theta - 16128\cos3\theta - 20352\cos2\theta$$
$$- 22528\cos\theta - 14464$$
$$= 0$$

Approximate roots of the equation $\Delta = 0$ can be obtained by plotting the function $\Delta = f(\theta)$ at intervals of $6°$ and finding where $f(\theta) = 0$. This can readily be carried out by means of Beevers – Lipson strips and once these approximate roots have been obtained, more accurate values can be calculated by making use of Taylor's theorem in the form

$$\Delta(x+h) = \Delta(x) + \frac{h}{1!}\Delta'(x) + \frac{h^2}{2!}\Delta''(x) + \ldots\ldots = 0$$

where $\Delta(x)$ is the value of Δ using the approximate root, x and h is the correction to be applied.

The derivatives $\frac{dS_n}{dx}$ can be obtained from the formula

$$\frac{dS_n}{dx} = nS_{n-1} + (n-2)S_{n-3} + (n-4)S_{n-5} + \ldots\ldots\ldots\ldots + (1 \text{ or } 0)$$

the final term depending on whether n is odd or even.

Successive approximations of this type lead to exact values for the roots x of the secular equations and roots which lie outside the range $-2 < x < +2$ can then be obtained from the theory of equations. These exact values for x are

x = + 2.15775; +1.45460; +1.07655; +0.90997; +0.33221;

  −0.91874; −1.10302; −1.58342; −1.84220; −2.48377

But $E_j = \alpha - x\beta$ and so the permitted energy levels for the π —electrons of pteridine are, in ascending order of magnitude (since β is negative).

j	E_j	j	E_j
(1)	α +2.48377 β	(6)	α −0.33221 β
(2)	α +1.84220 β	(7)	α −0.90997 β
(3)	α +1.58342 β	(8)	α −1.07655 β
(4)	α +1.10302 β	(9)	α −1.45460 β
(5)	α +0.91874 β	(10)	α −2.15775 β

Table (X)

Pteridine has 10 π —electrons and each energy level can hold two electrons (with opposite spins). Hence, to obtain the energy level for the molecule whilst in its ground state the electrons are filled in two at a time, starting with the lowest level and working upwards,

until each of the elctrons has been accounted for.

In this way, the permitted energy levels for the π-electrons of pteridine when the molecule is in the ground state are shown to be

i	E_i	Number of electrons in each level
(1)	$\alpha +2.48377\ \beta$	2
(2)	$\alpha +1.84220\ \beta$	2
(3)	$\alpha +1.58342\ \beta$	2
(4)	$\alpha +1.10302\ \beta$	2
(5)	$\alpha +0.91874\ \beta$	2

4. The Secular Coefficients for each Energy Level.

By means of a process of successive elimination, the secular equations on p.(103) can be reduced to

(10) $c_1 + c_5 + c_9 + xc_{10} = 0$

(3) $2c_2 + (2x+1)c_3 + 2c_4 = 0$

(4) $c_3 + xc_4 + c_5 = 0$

(5) $c_4 + xc_5 + c_6 + c_{10} = 0$

(6) $2c_5 + (2x+1)c_6 + 2c_7 = 0$

(7) $c_6 + xc_7 + c_8 = 0$

(8) $c_7 + xc_8 + c_9 = 0$

(9) $2c_8 + (2x+1)c_9 + 2c_{10} = 0$

(1) $(8x^7 + 12x^6 - 34x^5 - 43x^4 + 42x^3 + 35x^2 - 12x)c_9$

$+ (8x^6 + 8x^5 - 30x^4 - 20x^3 + 46x^2 + 16x - 24)c_{10}$

$= 0$

(2) $(8x^8 + 12x^7 - 50x^6 - 63x^5 + 106x^4 + 98x^3 - 84x^2 - 43x + 20) c_9$

$+ (8x^7 + 8x^6 - 46x^5 - 32x^4 + 92x^3 + 36x^2 - 54x - 8)c_{10}$

$= 0$

and all of the coefficients can, by substituting the correct values of x, be expressed in terms of c_{10}.

The normalising condition $\sum (c_{r,j})^2 = 1$ can then be employed and and hence c_{10} estimated. In this way, absolute values of the coefficients for each energy level can be obtained.

These values were found to be

Energy Level (E_j)	$c_{1,j}$	$c_{2,j}$	$c_{3,j}$	$c_{4,j}$	$c_{5,j}$
$\alpha +2.48377\,\beta$	+0.34532	+ 0.24379	+0.26013	+0.27224	+0.41606
$\alpha +1.84220\,\beta$	−0.29466	−0.42432	−0.48705	−0.22939	+0.06446
$\alpha +1.58342\,\beta$	+0.37720	+0.08588	−0.24089	−0.34687	−0.30835
$\alpha +1.10302\,\beta$	+0.21515	−0.22189	−0.45989	−0.05543	+0.39874
$\alpha +0.91874\,\beta$	−0.46632	−0.34388	+0.15041	+0.40686	+0.22339

Energy Level (E_j)	$c_{6,j}$	$c_{7,j}$	$c_{8,j}$	$c_{9,j}$	$c_{10,j}$
$\alpha +2.48377\,\beta$	+0.31985	+0.21844	+0.22272	+0.33473	+0.44131
$\alpha +1.84220\,\beta$	+0.31929	+0.36410	+0.35145	+0.28334	+0.02885
$\alpha +1.58342\,\beta$	−0.46401	−0.19436	+0.15625	+0.44202	+0.32263
$\alpha +1.10302\,\beta$	+0.14365	−0.31212	−0.48793	−0.22607	+0.35160
$\alpha +0.91874\,\beta$	−0.35026	−0.37004	+0.01029	+0.37950	+0.14863

5. The Bond Orders of Pteridine.

The mobile bond order p_{rs} between atoms r and s is defined as

$$p_{rs} = \sum n_j c_{r,j} c_{s,j}$$

where n_j is the number of electrons in energy level E_j (2 in this case) $c_{r,j}$ and $c_{s,j}$ are the coefficients of the secular equations for atoms r and s for the jth level, and the summation is taken over all the energy levels,

i.e. $$p_{rs} = 2 \sum c_{r,j} c_{s,j}$$

Use of this equation gives the following values for the mobile bond orders of pteridine.

Bond r-s	1-2	1-10	2-3	3-4	4-5	5-6
p_{rs}	0.709	0.558	0.599	0.706	0.549	0.552

Bond r-s	5-10	6-7	7-8	8-9	9-10
p_{rs}	0.519	0.722	0.590	0.715	0.551

6. π -electron Distributions(q_r) in the Ground State of Pteridine.

$(c_{r,j})^2$ represents the contribution of electron j to the π -electron density at atom r. Hence, the total π -electron density at atom r is given by

$$q_r = \sum_j n(c_{r,j})^2$$

the summation being taken over all the occupied energy levels and n is the number of electrons in each level. This results in the following π -electron densities for pteridine.

r	1	2	3	4	5	6	7	8	9	10
q_r	1.224	0.829	1.194	0.831	0.962	1.126	0.905	0.872	1.166	0.891

Carbon atoms 2 and 4 are therefore, most deficient in electrons and any nucleophilic substitution should occur at these points [Chalvet and Sandorfy (47) by algebraic superposition of results for compounds containing fewer nitrogen atoms in different positions find C_4 the more deficient.]

7.   The Effect of Variation of the Values Assigned to the

     Arbitrary Parameters on the Calculations.

In the process of obtaining the bond orders of the pteridine
molecule in the ground state, a number of arbitrary assumptions have been
made about the values to be assigned to the parameters H_{rr}, H_{rs}, S_{rs},
These have all been expressed in terms of α and β (the Coulomb integral
for carbon and the Exchange integral for carbon-carbon bonds respectively)
or else have been neglected in the calculations.

It is of interest to know what effect varying the values assigned
to these parameters has on the final results obtained for the
bond orders and π-electron distribution of pteridine.

H_{nn} is usually assigned a value of $(\alpha + 0.6\beta)$ since this
reproduces correctly the dipole moments of simple hetrocyclics
containing nitrogen $(47, 48)$.   A value of $H_{nn} = \alpha + \frac{1}{2}\beta$ has been used
in this work on pteridine.   A comparison of calculations carried
out by the author on pyridine using $H_{nn} = \alpha + \frac{1}{2}\beta$ with those of
Chalvet and Sandorfy (47) using $H_{nn} = \alpha + 0.6\beta$ shows that this minor
variation results in differences in π-electron distributions and
mobile bond orders of 0.015e and 0.002 units respectively - differences
of no importance.

The Exchange integral for a carbon-nitrogen bond will be slightly
greater than that of a carbon-carbon bond, although in this work the
two Exchange integrals have been assumed to be equal.   To determine
the effect of varying H_{C-N} on the calculations, they have been
reperformed using an extreme - and erroneous - value of H_{C-N} equals $1\frac{1}{2}\beta$.
This makes a serious difference in the permitted energy levels of as
much as 0.8β although the average differences which result in
the π- eletron distributions for carbon atoms is only 0.03e and for

nitrogen is 0.05e. This variation in the Exchange integral for carbon-nitrogen bonds also results in small differences on the average of 0.03 units in mobile bond order.

Recently, (51) , it has been pointed out that under certain circumstances nitrogen may be slightly more electropositive than carbon. Under these circumstances H_{NN} and H_{CN} will be less than α and β respectively. However, nitrogen will only be less eletronegative than carbon if, in some way, a high negative charge is resident on the atom. This situation seems to obtain in some five-membered ring systems and consequently, doubts as to the values to be assigned to H_{NN} and H_{CN} must arise in such cases. However, in pteridine the nitrogen atoms are undoubtedly more electronegative than carbon and so the values of H_{NN} and H_{CN} which have been used are those appropriate to this problem.

The Overlap integral S_{rs} has been nelgected in this work. The justification for this is to be found in a paper by D.W. Davis (52) in which the effect of overlap in molecular orbital calculations of hetrocyclic molecules is examined. Davis finds that, by including overlap in the calculations, the change produced in the π -electron distributions is extremely small,(of the order of 0.001 - 0.003e) and so also is the effect on the mobile bond orders(of the order of 0.002 - 0.007 bond units).

Hence, as far as varying the values assigned to the parameters H_{rr}, H_{rs}, S_{rs} is concerned, the effect on the π -electron distributions and mobile bond orders is negligible although the permitted energy levels are more sensitive to such changes.

8. **A Correlation Curve for Carbon-Nitrogen Bonds and Bond Lengths of Pteridine.**

So far, a discussion of the study of the pteridine molecule by means of molecular orbital theory has been given up to the point of estimation of mobile bond orders and it has been shown that these bond orders ~~dervied~~ *derived* by this method should be very reliable.

In order to convert bond orders into bond lengths a correlation curve is necessary. For the carbon-carbon bonds this curve is well established (≤3) but the position for carbon-nitrogen bonds is much less satisfactory. The curve of Cox and Jeffrey (≤4) has often been used in this connection but is unsatisfactory since the intermediate points (pyridine and melamine) are in error. Since Cox and Jeffrey published their paper, however, a number of compounds containing accurately measured C-N bonds have been reported. These, and the results for them are listed below.
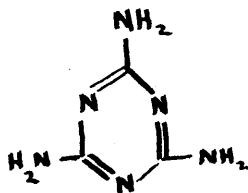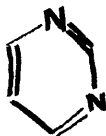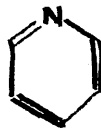

I   II   III

IV   V
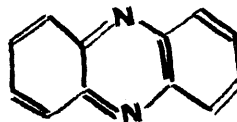
Compound	p_{C-N}	Bond Lengths	Method of Measurement
Melamine I(ring C-N)	0.563(55)	1.343 ± 0.05A.	X-ray (56)
(side C-N)	0.520(55)	1.353 ± 0.05A.	
s-Triazine II	0.654(55)	1.319 ± 0.005A.	X-ray (57)
α-Phenazine III	0.603(58)	1.345 ± 0.009A.	X-ray (59)
Pyrimidine IV	0.656(6c)	1.33A.	X-ray (61)
Pyridine V	0.654(6c)	1.340A.	Micro-wave (62)

The extremities of the single bond – double bond range
are not very satisfactory.   Cox and Jeffrey quote a pure
single bond length of 1.475A. but in the aliphatic amines
which they quote as the source of this measurement the
hybridisation is sp^3 whereas sp^2 hybridisation is involved
in condensed ring compounds.   Coulson(43) has considered the
variation of length in carbon-carbon bonds as the hybridisation
changes and has shown that the length of 1.54A. for the
tetrahedral C-C becomes 1.50A. in the trigonal hybrid.   It is,
therefore, suggested that a similar shortening is to be
expected for the C-N bond.   Hence, 1.435A. has been taken for
the sp^2 carbon-nitrogen bond of zero order.

For the pure bond of π-order 1, Cox and Jeffrey select a
length of 1.28A. and the covalent radii values of Stevenson and
Schomaker (44) with the electro-negativity correction according
tb Gordy (45) give 1.27A.

In building up the bond order – bond length curve in this
work, however, another method of determining the point for the
pure bond of mobile order 1, has been employed.   The six
intermediate points listed above, mean that any reasonable curve
must pass through or very near to the point (0.600,1.336).

Fig XX1.

Bond Order — Bond Length Curves
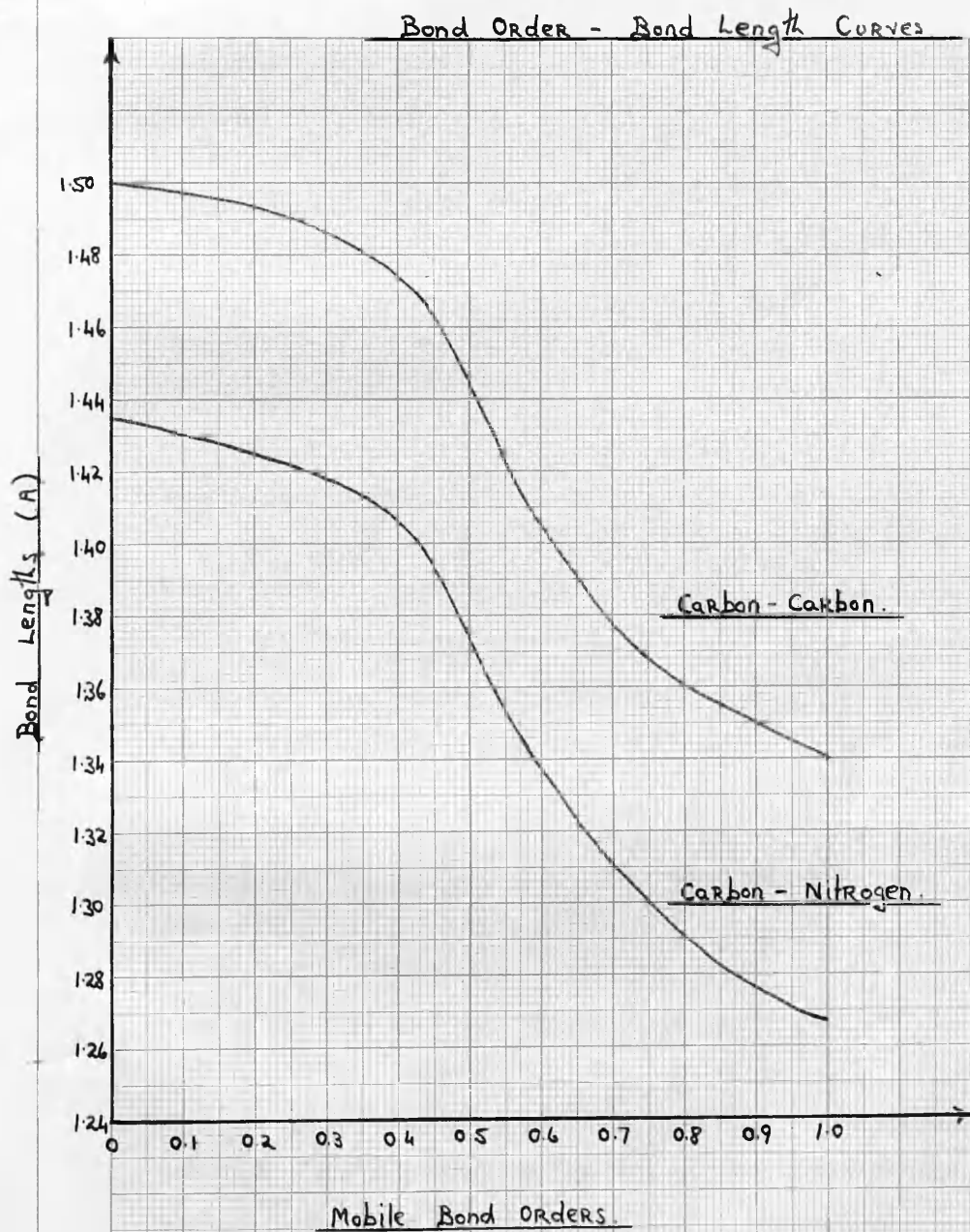
This point has, therefore, been taken as a datum for developing the carbon-nitrogen bond order - bond length curve. In the carbon-carbon correlation curve ($\mathcal{S}3$) the bond lengths corresponding to orders 0 and 0.6 are 1.50A. and 1.406A. The carbon-nitrogen curve has been obtained by displacing and proportionally compressing the carbon-carbon curve so that it passes through (0,1.435) and (0.600,1.336). The curve shown in fig.(XXI) results and passes through the points (0.1,1.431), (0.2,1.425), (0.3,1.418), (0.4,1.407), (0.5,1.375), (0.6,1.336), (0.7,1.312), (0.8,1,291), (0.9,1.277) and (1.0,1.267). The last of these gives the length of the pure double bond and agrees closely with the values assigned to this bond length by the authors already mentioned.

The justification for this procedure is that the carbon-carbon curve has been found to be very reliable and it is expected that the variation in carbon-nitrogen bond length with bond order would follow a path similar to the carbon-carbon curve.

It is difficult to decide just how accurate bond lengths obtained from the curve throughout its whole length will be, since this obviously depends on the accuracy with which the reference points on the curve have been obtained. However, in the case of pteridine the effective region of the curve to be used lies within the range in C-N mobile bond orders from 0.73 - 0.55. This range lies very near to the datum point (0.6,1.336) and so it is believed that the error associated with bond lengths derived from this curve will not be greater than $\pm$ 0.02A.

On making use of the C-C and C-N curves discussed above the bond lengths obtained for the ground state of the

pteridine molecule are found to be
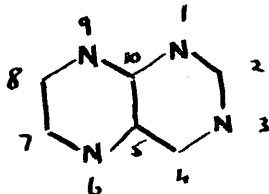
Bond r-s	Derived Length	Observed Length	Δ
1-2	1.310A.	1.34A.	0.03A.
1-10	1.349	1.40	0.05
2-3	1.337	1.40	0.06
3-4	1.311	1.28	-0.03
4-5	1.425	1.42	-0.01
5-6	1.351	1.39	0.04
5-10	1.435	1.35	-0.09
6-7	1.307	1.36	0.05
7-8	1.410	1.39	-0.02
8-9	1.309	1.32	0.01
9-10	1.352	1.41	0.06

This table compares the derived lengths with those obtained
in an X-ray crystallographic investigation of pteridine by
Hamor and Robertson (66) .   In it, Δ signifies measured
length minus derived length.   The results give a mean value
$|\Delta|$ mean = 0.04.   The observed lengths are nearly always
greater than the derived lengths and this would appear, at
first sight, to indicate that the correlation curves should be raised
slightly.   However, the standard deviation in observed bond
lengths has been quoted by Hamor and Robertson to be 0.03A.,
so the mean discrepancy is well within the limits of

experimental error.   The observed length for bond $C_5- C_{10}(1.35A)$ is obviously too short and it is of interest to note that if atoms 5 and 10 are equally displaced along the line joining them so that they become separated by the distance 1.435A. indicated by the calculations, then the discrepancies in the bonds $N_9- C_{10}$, $C_{10}- N_1$, $C_5- N_6$, $C_5- C_4$ are greatly reduced to 0.04, 0.02, 0.01, and —0.03A. respectively.

Again, much the shortest bond measured is $N_3- C_4(1.28A.)$, a length which would mean that bond $N_3- C_4$ was almost, if not quite, a pure double bond — a state of affairs which, according to molecular orbital theory is impossible in this type of compound.

Occasionally in the literature during comparison of observed and calculated bond lengths one comes across statements to the effect that discrepancies in the results obtained by the two methods of approach may be due to

(a) the effect of formal charges on bond lengths, or,

(b) the compounds under consideration are not planar, or,

(c) the bonds, although co-planar, are distorted from the simple sp^2 hybridised type which has been assumed in performing the cabulations.

It is true that all of these effects do alter bond lengths but these effects have been neglected in the calculations since their effects were believed to be small.   The justification for this procedure is in two papers published by Coulson on this subject.   In the first paper (67) Coulson deals with the effect of formal charge distribution on bond lengths and shows that an estimate of this effect can be obtained from the equation

$$\Delta r \;=\; \frac{q_1 q_2 e^2}{K k r^2}$$

where the bond length in the absence of any charges on the atoms is r and Δr is the change in r when charges q_1 and q_2 are placed on atoms 1 and 2. e is the electronic charge. K is the effective dielectric constant for the medium surrounding the bond and k is the force constant for the bond 1-2. In the case of pteridine the greatest formal charges are on N_1 and C_2 and are $-0.224e$ and $+0.171e$ respectively. If K is assumed to be about 2, $r = 1.31A$, $k = 6 \times 10^5$ dynes/cm. then Δr turns out to be $-0.005A$. i.e. a contraction of 0.005A in bond length occurs. Obviously, this as a source of discrepancy between observed and calculated bond lengths is of no importance.

Provided the molecule deviates only very slightly from planarity effects (b) and (c) can be considered together. The effect of variation in hybridisation on atomic radius is discussed by Coulson (63). The molecular orbital description of a hybrid bond for a carbon atom is defined by the expression

$$\psi(\lambda) \;=\; \frac{s + \lambda p}{\sqrt{(1 + \lambda^2)}}$$

where s and p denote carbon 2s and 2p orbitals respectively and the square root ensures normalisation.

In rough terms it might be said that the amounts of s and p atomic orbitals in the hybrid orbital $\psi(\lambda)$ are $\dfrac{1}{1+\lambda^2}$ and $\dfrac{\lambda^2}{1+\lambda^2}$. When the hybridisation ratio λ^2 is 1, an sp hybrid orbital results and when $\lambda^2 = 2$ or 3, sp^2 or sp^3 hybrids are obtained.

The variation of atomic radius for carbon atoms as the hybridisation changes from sp through sp^2 to sp^3 is known accurately from spectroscopic data and hence a correlation curve of atomic radius of carbon against λ can be obtained. Coulson has derived expressions relating to trigonometric functions of the angles between the bonds radiating from an atom. Hence if these bond angles are known, λ can be calculated and then by reference to the correlation curve the atomic radius for the carbon atom can be obtained. The difference of this radius from that of the atom when in the pure sp^2 hybrid state can then be estimated and thence the contribution of this atom to the shortening (or lengthening) of a bond can be obtained (it is assumed that a small variation from the pure sp^2 type in nitrogen atoms

produces a shortening or lengthening of the nitrogen radius by an amount which is the same as that found for carbon).

By arguing in this way it is readily shown that the order of correction due to effects (b) and (c) in pteridine is about 0.002A. Clearly these changes are insignificant.

A summary of the calculated bond orders, π-electron distributions and calculated bond lengths obtained for pteridine is shown in figure (XXII).

Fig XXII.
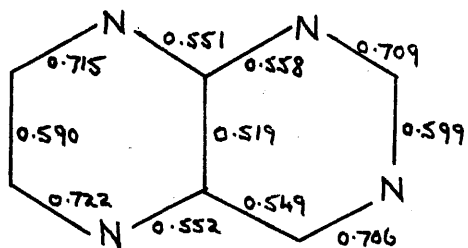
Mobile

Bond Orders.
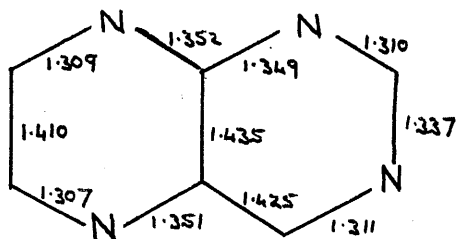
Bond Lengths.
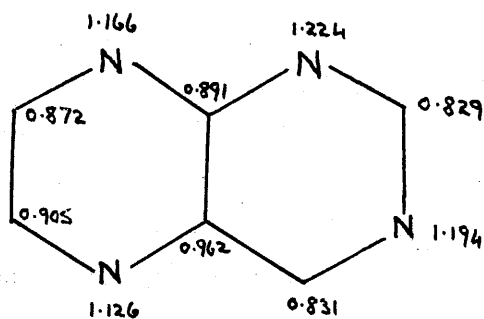
π - electron Distributions.

R E F E R E N C E S

1) M.B. Epstein and F.D. Rossini.

        Chem. Eng. News, 1948, _26_, 2959.

2) J. Cason and G. Sumrell.

        J. Biol. Chem., 1951, _192_, 405.

3) J. Cason, N.K. Freeman and G. Sumrell.

        J. Biol. Chem., 1951, _192_, 415.

4) J. Cason and C.F. Allen.

        J. Biol. Chem., 1953, _205_, 449.

5) J. Cason, N.L. Allinger and D.E. Williams.

        J. Org. Chem., 1953, _18_, 842.

6) J. Cason, N.L. Allinger and C.F. Allen.

        J. Org. Chem., 1953, _18_, 857.

7) S. Ställberg-Stenhagen.

        Arkiv. Kemi., 1954, _6_, 537.

8) J. Cason and M.J. Kalm.

        J. Org. Chem., 1954, _19_, 1836.

9) R. Adams and B.L. Van Duuren.

        J. Amer. Chem. Soc., 1953, _75_, 4631.

10) J.J. Sudborough and M.J.P. Davies.

        Trans. Chem. Soc., 1909, _95_, 976.

11) K. von Auwers and H. Wissebach.

        Ber. 1923, _56_, 715-733.

12) P. Pfeiffer.

        Z. Physik. Chem., 1904, _48_, 40.

13) A.S. Dreiding and R.J. Pratt.

        J. Amer. Chem. Soc., 1954, _76_, 1902-1904.

14) G. Knott.

        Proc. Phys. Soc. Lond., 1940, _52_, 229.

15)   D.M. Wrinch.

"Fourier Transforms and Structure Factors",

A.S.X.R.E.D. Monograph No.2.

16)   Haushofer.

Groth, "Chemische Krystallographie", III, 400.

17)   J.M. Robertson.

J. Sci. Instrum., 1943, 20, 175.

18)   G.W. Brindley and R.G. Wood.

Phil. Mag., 1929, 7, 616-619.

19)   D. Crowfoot, C.W. Bunn, B.W. Rogers-Low, and A.Turner-Jones.

"The Chemistry of Penicillin",

Princeton University Press, 1949, 310, 334.

20)   C. Finbak and N. Norman.

Acta Chem. Scand., 1948, 2, 813.

21)   A.D. Booth.

"Nature", Lond., 1948, 161, 765.

22)   W. Cochran.

Acta Cryst., 1951, 4, 81.

23)   R. McWeeny.

Acta Cryst., 1951, 4, 513.

24)   D.W.J. Cruickshank.

Acta Cryst., 1949, 2, 65.

25)   D.W.J. Cruickshank.

Acta Cryst., 1954, 7, 519.

26)   F.R. Ahmed and D.W.J. Cruickshank.

Acta Cryst., 1953, 6, 385.

27)   D.W.J. Cruickshank and A.P. Robertson.

Acta Cryst., 1953, 6, 698.

28)   W. Cochran.

Acta Cryst., 1953, 6, 260.

29) G.A. Sim, J.M. Robertson and T.H. Goodwin.

Acta Cryst., 1955, $\underline{8}$, 157.

30) L. Pauling.

"Nature of the Chemical Bond", Cornell University

Press, Ithaca, New York, 1950, p.189.

31) J.M. Robertson.

Proc. Roy. Soc., 1933, $\underline{142A}$, 671.

32) L.O. Brockway and J.M. Robertson.

J. Chem. Soc., 1939, 1331.

33) D.T. Cromer, A.J. Ihde and H.L. Ritter.

J. Amer. Chem. Soc., 1951, $\underline{73}$, 5590.

34) H.A. Levy and R.B. Corey.

J. Amer. Chem. Soc., 1941, $\underline{63}$, 2095.

35) L. Pauling.

"Nature of the Chemical Bond", Cornell University

Press, Ithaca, New York, 1950, p.189.

36) E. Mack.

J. Amer. Chem. Soc., 1932, $\underline{54}$, 2159.

37) Schimper.

Groth, "Chemische Krystallographie", $\underline{III}$, 400.

38) A.J.C. Wilson.

"Nature", Lond., 1942, $\underline{150}$, 152.

39) See R.E. Buckles, G.V. Mock and L. Locatell Jr.

Chem. Rev., 1955, $\underline{55}$, 659 and references therein.

40) E.D. Bergmann, D. Ginsburg, Y.Hirshberg, M.Mayot,

A. Pullman and B. Pullman.

Bull. Soc., Chim., 1951, $\underline{18}$, 697.

41) E. Harnik, F.H. Herbstein, G.M.J. Schmidt and F.L. Hirshfeld.

    J. Chem. Soc., 1954, 3293.

42) A.O. McIntosh, J.M.Robertson and V. Vand.

    "Nature", Lond., 1952, 169, 322.

43) C.P. Fenimore.

    Acta Cryst., 1948, 1, 295.

44) Groth, "Chemische Krystallographie", V, 431.

45) E. Harnik and G.M.J. Schmidt.

    J. Chem. Soc., 1954, 3295.

46) L.E. Orgel, T.L. Cottrell, W. Dick and L.E. Sutton.

    Trans. Far. Soc., 1951, 47(2), 113.

47) O. Chalvet and C. Sandorfy.

    Compt. rend., 1949, 228, 556.

48) P.Löwdin.

    J. Chem. Phys., 1951, 19, 1323.

49) T.H. Goodwin and V. Vand.

    J. Chem. Soc., 1955, 1683.

50) T.H. Goodwin and V. Vand.

    J. Chem. Soc., 1955, 1684.

51) I.M. Bassett, R.D. Brown and A. Penfold.

    Chem. and Ind.,1956, 892.

52) D.W. Davies.

    Trans. Far. Soc., 1955, 51(4), 449.

53) T.H. Goodwin and V. Vand.

    J. Chem. Soc., 1955, 1687.

54) E.G. Cox and G.A. Jeffrey.

    Proc. Roy. Soc., 1951, 207, 113.

55) T.H. Goodwin.

    Personal communication 1956.

56) E.W. Hughes.

J. Amer. Chem. Soc., 1941, 63, 1737.

57) P.J. Wheatley.

Acta Cryst., 1955, 8, 224.

58) Personal communication by Pullman quoted in ref. 59(p.410).

59) F.H. Herbstein and G.M.J. Schmidt.

Acta Cryst., 1955, 8, 399.

60) O. Chalvet and C. Sandorfy.

Compt. rend., 1949, 228, 556.

61) Predicted by C.J.B. Clews and W. Cochran(Acta Cryst.,1948,1,4)
on the basis of their and other data(G.J.Pitt,ibid., p.168;
c.f. C.J.B. Clews and W. Cochran, ibid., 1949, 2, 46).

62) B. Bak, L. Hansen and J. Rastrup-Andersen.

J. Chem.Phys., 1954, 22, 2013.

63) C.A. Coulson.

Victor Henri Memorial Volume, " Contribution à l'étude
de la structure moléculaire," Desoer, Liège, 1948,15.

64) D.P. Stevenson and V. Schomaker.

J. Amer. Chem. Soc., 1941, 63, 37.

65) W. Gordy.

J. Chem. Phys., 1947, 15, 81.

66) T. Hamor and J.M. Robertson.

Personal communication, 1956.

67) C.A. Coulson.

J. Phys. Chem., 1952, 56, 311.