Acta Cryst. (1954). **7**, 808

Crystal and Molecular Structure of *p*-Amino-Salicylic Acid*

By F. Bertinotti, G. Giacomello and A. M. Liquori

Istituto di Chimica Farmaceutica, University of Rome, Italy

(*Received* 6 *May* 1954)

The unit cell of *p*-amino-salicylic acid is monoclinic with: $a = 7{\cdot}28$, $b = 3{\cdot}82$, $c = 25{\cdot}33$ Å, $\beta = 103°$, $Z = 4$. The space group is $P2_1/c$. The structure has been established by the use of Patterson projections and refined by two-dimensional Fourier syntheses. Finally a 'difference Fourier' has been computed to confirm the positions assumed for the hydrogen atoms.

Intermolecular and intramolecular hydrogen bridges have been found. The bond-length variations in the benzene ring and in the C–O and C–N bonds are discussed in terms of resonance between several structures.

Introduction

p-Amino-salicylic acid (P.A.S.) is known to be a biologically active substance and has therefore received much attention in recent years. Like salicylic acid, this substance has several physical properties which suggest that the molecules are associated by hydrogen bridges; for this reason we decided to investigate its crystal structure by X-ray diffraction.

A preliminary account of this investigation has been given elsewhere (Coccia, Giacomello & Liquori, 1951).

Determination of the structure

(1) *Crystallographic data*

The unit cell of *p*-amino-salicylic acid is monoclinic with the following dimensions:

$$a = 7{\cdot}28,\ b = 3{\cdot}82,\ c = 25{\cdot}33\ \text{Å},\ \beta = 103°;\quad V = 686\ \text{Å}^3\,.$$

The space group is $P2_1/c$ and there are four molecules in the unit cell in general positions.

These crystallographic data indicated from the start that the molecules were packed in a very favourable way. The long c axis and the very short b axis clearly suggested that the molecules do not overlap, at least on the projections on (010) and (100), and are only slightly tilted with respect to the a–c plane.

(2) *Intensity measurements*

Intensities recorded on Weissenberg photographs were estimated by visual comparison with standard scales. The multiple-film technique was used to correlate the strong and weak reflexions, the extreme range being 3500 to 1. The measurements were carried out by two independent observers, and the values averaged.

By comparing both the deviations from these averages, and the differences between the intensities of reflexions common to different sets, the error was estimated to be between 5% and 10%.

The experimental intensities were corrected only for Lorentz and polarization factors and were then related to the absolute scale by the method of Wilson (1942), which was separately applied to the three zones.

(3) *Molecular orientation on* (010)

A Patterson projection on (010) confirmed the above expectations and allowed us to establish the molecular orientation on this plane with the help of a model. The vector distribution of the Patterson map was found to be consistent with three different possibilities of orientation of the projected molecules. Structure factors were calculated for the three alternatives.

The best agreement was obtained for an orientation in which the molecular axis (connecting the carbon atom of the carboxyl group with the nitrogen of the amino group) forms an angle of 55° with the c axis. Using the phase constants of these structure factors, a Fourier synthesis was computed with 60 amplitudes. This gave a very sharp resolution of all the atoms.

The reliability factor R, calculated from the coordinates derived from the Fourier projection, was at this stage 0·40. The use of a thermal factor determined by a graphical plotting of $\log F_o/F_c$ against $(\sin\theta\,.\,\lambda^{-1})^2$ reduced this value to 0·22.

(4) *Molecular orientation on* (100)

Owing to the small b axis and to the slight tilting of the molecules with respect to the a–c plane, the Patterson projection on (100) did not lead to a straightforward interpretation as in the case of the projection on (010). Therefore the intensity distribution in the whole reciprocal lattice was investigated in order to find some intense reflexion which could indicate the approximate orientation of the molecules.

A very strong reflexion with index $(12\bar{3})$ was found on the Weissenberg photograph of the second layer. The intensity of this reflexion was related to the absolute scale by comparison with $0kl$ reflexions present

---

* This research has been supported by the 'Consiglio Nazionale delle Ricerche'.

on the same film. The absolute intensity of this reflexion, corrected for Lorentz and polarization factors was found to be close to the square of the structure factor calculated on the assumption that the whole molecule lay in the $(12\bar{3})$ plane. By solving eleven equations of the type

$$x_i+2y_i-3z_i = 1$$

the y coordinates of the eleven atoms contained in one molecule were derived from the corresponding x and z coordinates determined from the a–c Fourier projection. Structure factors were then calculated for the $0kl$ reflexions and found to agree surprisingly well with the observed values.

A projection on (100) was then computed with almost all the observed amplitudes. Although the electron-density map showed a very satisfying background, only three atoms of the molecules were fully resolved. The coordinates of these three atoms were used to calculate the coefficients HKL of the molecular plane by means of the following set of equations:

$$\begin{aligned} Hx_1+Ky_1+Lz_1 &= 1\ , \\ Hx_2+Ky_2+Lz_2 &= 1\ , \\ Hx_3+Ky_3+Lz_3 &= 1\ . \end{aligned}$$

This gave $H = 1{\cdot}015$, $K = 1{\cdot}941$, $L = -2{\cdot}985$, in good agreement with the Miller indices $(12\bar{3})$.

By the refinement of the coordinates these coefficients of the molecular plane were further changed to the final values $H = 1{\cdot}010$, $K = 1{\cdot}949$, $L = -3{\cdot}018$.

(5) *Refinement of the structure*

A refining process was carried out mainly for the projection on (010) and only indirectly for the projections on (100) and (001) by

(*a*) computation of successive Fourier syntheses in which a progressive number of amplitudes were included;

(*b*) exact location of the atomic coordinates by interpolating nine values of the electron density according to a Gaussian function (Shoemaker, Donohue, Schomaker & Corey, 1950);

(*c*) scaling the F_o values at every stage of the calculations and accurate determination of the temperature factor;

(*d*) back-shift corrections by computing 'calculated Fourier syntheses' according to Booth (1946);

(*e*) inclusion of the contribution of the hydrogens in the calculation of the structure factors.

The process proved to be very efficient since the accuracy in the atomic coordinates improved continuously, as indicated by the decrease of R at every stage.

Four Fourier syntheses were computed with the F_o's and three with the F_c's on (010).

The number of terms included was increased from 60 to 140, while R was reduced to 0·15 after correction for isotropic thermal vibrations and inclusion of the contribution of the hydrogens. The final density projection is shown in Fig. 1.

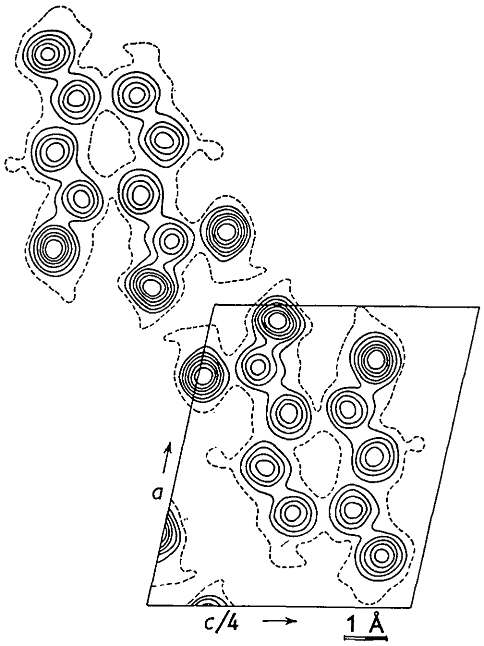


Fig. 1. Electron density projected on (010). Contours are drawn at intervals of 1 e.Å^{-2}, the first being broken.

At this stage it was recognized that no further progress could be achieved. This can probably be attributed to the anisotropy of the thermal vibration or of the atomic scattering. The density map shows in fact that the shape of the peaks is rather elliptical and this effect is particularly marked for the oxygen atoms.

It may be worth while to point out that the direction of the elongation of the oxygen peaks is nearly at right angles to the bond direction, just as found by Cochran (1953) in salicylic acid.

A smaller number of Fourier syntheses was calculated on (100) since the refinement of this projection

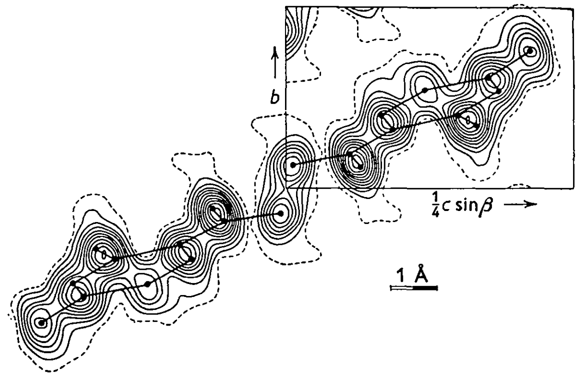


Fig. 2. Electron density projected on (100). Contours are drawn at intervals of 1 e.Å^{-2}, the 2 e.Å^{-2} line being broken.

was partially taken care of by the refinement calculations on the (010) projection. In this way we arrived at a value of $R = 0{\cdot}17$ for this projection. The final projection on (100) is given in Fig. 2.

Only one Fourier synthesis was computed for the projection on (001) but it appeared impossible to draw any conclusion from it. The reliability factor R for this zone was therefore computed only at the last stage of the structure determination; a value of 0·23 was obtained.

In Table 1* approximately 300 values of F_c are compared with the F_o's for the projections on (100), (010) and (001).

(6) *The hydrogen atoms*

A characteristic feature of the projection on (010), which was not eliminated by the refinement, consisted in a regular deformation of the 1 e.Å^{-2} line in regions where the hydrogen atoms were expected to appear. Such an effect was not present in the map calculated with the F_c's omitting the hydrogen contributions. The contribution of the hydrogens to the structure factors was therefore taken into consideration.

The coordinates were calculated by assuming usual bond angles and bond distances, and the atomic scattering factor for hydrogen reported in the *International Tables* was used. A reduction in the value of R from 0·18 to 0·14 was observed (only F_{h0l}'s with $2 \sin \theta \leqslant 1$ were considered this time).

These results encouraged the computation of a 'Fourier synthesis' using the differences $(F_o - F_c)$ as amplitudes, F_o's being the observed structure factors on an absolute scale and F_c's the structure factors calculated from the final coordinates without the contribution of the hydrogens and corrected for the isotropic thermal factor. The resulting map is shown in Fig. 3.

Apart from fluctuations in the electron density mainly caused by anisotropic effects, the difference map contains regions of about 0·6 e.Å^{-2} which can be attributed with certainty to the hydrogen atoms. An electron density of 0·3 e.Å^{-2} is found on some of the C–C bonds. This may be taken as a qualitative indication that the electron density in the bonds is higher than would be expected from the overlapping of the spherical atomic clouds, as has been already shown by Brill (1950).

Unfortunately, a quantitative interpretation of such an effect cannot be attempted, since the mean standard deviation (m.s.d.) of the electron density was estimated to be about 0·2 e.Å^{-2} (see paragraph 9). It is, however, to be noticed that these positive values of the electron density on the bonds are compensated by negative values in the centre of the benzene ring and of the chelation rings. A similar distribution of the electron density has been found by Cochran for salicylic acid, although this author was able to get a much more detailed and precise map as a result of a more refined experimental technique in the measurements of the intensities.

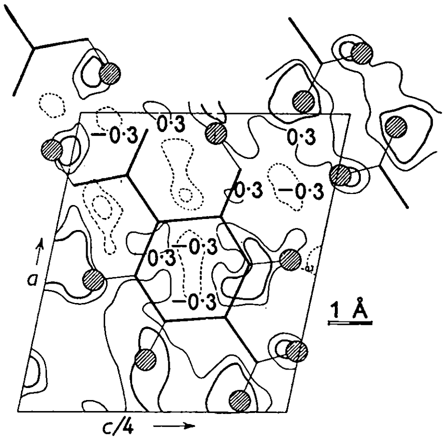


Fig. 3. $(F_o - F_c)$ Fourier projection on (010). Contributions from the hydrogen atoms have not been included. The probable positions of hydrogen atoms are indicated by shaded circles. Contour lines are drawn at intervals of ±0·3 e.Å^{-2}, the +0·6 e.Å^{-2} line being heavy and the negative contours being broken.

(7) *Arrangement of the molecules in the lattice, and intermolecular distances*

A drawing of the molecular arrangement and the intermolecular distances is given in Fig. 4.

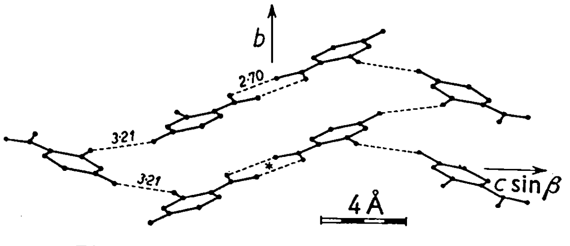


Fig. 4. The structure viewed along [100].

Dymeric molecules, held together by hydrogen bridges between carboxyl groups, are present in the crystal lattice of p-amino-salicylic acid.

The other intermolecular distances are of the van der Waals type, except the distances between the phenolic hydroxyl and the amino group and between two animo groups which seem to indicate the presence of very weak O · · · N and N · · · N hydrogen bonds.

(8) *Bond distances and bond angles*

Bond distances were calculated from the final coordinates listed in Table 2 and bond angles were

---

* *Editorial note.*—Table 1 has been withdrawn and is deposited as Document No. 4299 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., U.S.A. A copy may be secured by citing the Document number and by remitting $2·50 for photoprints, or $1·75 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

Table 2. *Final atomic parameters*

Atom	x/a	y/b	z/c
C_1	0·6343	0·329	0·0932
C_2	0·6473	0·407	0·1485
C_3	0·4960	0·543	0·1855
C_4	0·3204	0·617	0·1742
C_5	0·3020	0·542	0·1197
C_6	0·4547	0·404	0·0821
C_7	0·7892	0·190	0·0554
O_I	0·8130	0·346	0·1642
O_{II}	0·9412	0·125	0·0547
O_{III}	0·7547	0·131	0·0060
N	0·1665	0·754	0·2112

determined graphically. The values obtained are reported in Fig. 5.

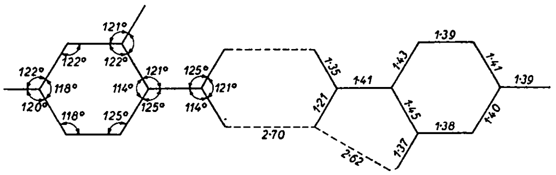


Fig. 5. Measured bond lengths and bond angles.

(9) *Accuracy in the structure determination*

The m.s.d. in the x_r and z_r coordinates were estimated by using the following relation given by Cruickshank (1949):

$$\sigma(x_{jr}) = \frac{2\pi}{a_j A}\,[\sum_2 h_j^2 \Delta F_j^2]^{\frac{1}{2}}/2\overline{p\varrho}_o\,.$$

For ΔF_j the values $4(F_o-F_c)$ and for $2\overline{p\varrho}_o$ the curvatures of the peaks in the a–c projections, averaged for atoms of the same type, were substituted. The curvatures were determined graphically by assuming a Gaussian distribution of the electron density within 0·4 Å from the centre of the peaks (Burbank, 1953).

Table 3. *Electron density and curvature of Fourier peaks in the projection on* (010)

Atom	ϱ_o (e.Å^{-2})	$-\frac{d^2\varrho_o}{dr^2}$ (e.Å^{-4})
C_1	6·84	52
C_2	6·71	52
C_3	6·93	54
C_4	6·71	52
C_5	6·45	45
C_6	6·60	49
C_7	6·77	61
O_I	9·24	70
O_{II}	9·58	69
O_{III}	9·39	71
N	7·83	60

ϱ_o, carbon $= 6{\cdot}72 \pm 0{\cdot}11$ e.Å^{-2}, ϱ_o, oxygen $= 9{\cdot}40 \pm 0{\cdot}12$ e.Å^{-2}.

The values of the curvatures are reported together with the values of the electron density at the center of the peaks in Table 3.

The m.s.d. in the x and z coordinates were

$\sigma(x) \cong \sigma(z) = 0{\cdot}017$ Å for the carbon atoms,
$\sigma(x) \cong \sigma(z) = 0{\cdot}012$ Å for the oxygen atoms,
and $\sigma(x) \cong \sigma(z) = 0{\cdot}014$ Å for the nitrogen atom.

Since the y coordinates were obtained indirectly from the x and z coordinates, neglecting the error in the orientation of the molecular plane, the m.s.d. in the bond distances was calculated from the following relation:

$$\sigma(d) = \frac{d}{d_{a-c}}\,\overline{\sigma}(x)\,\{2(1+\cos\beta)\}^{\frac{1}{2}}\,,$$

where d is the bond distance and d_{a-c} is its projection on the a–c plane.

The mean value of $\sigma(d)$ was 0·02 Å.

The σ values are to be considered overestimated since it was assumed that the differences F_o-F_c are to be attributed only to errors in the atomic coordinates, while, as discussed above, there are clear indications that anisotropic effects contribute also to this disagreement.

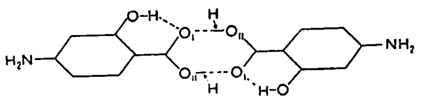
The mean standard deviation in the electron density on (010) was estimated from the following relation (Cochran, 1948):

$$\sigma(\varrho) = (\overline{\Delta\varrho^2})^{\frac{1}{2}} = \frac{2}{A}\cdot 4\sigma(F_o)(\tfrac{1}{2}N)^{\frac{1}{2}}\,.$$

$\sigma(F_o)$ was estimated by comparing the measured intensities of $00l$ reflexions of two different equatorial Weissenberg photographs and a value of 0·36 was obtained ($F(000) = 80$). $\sigma(\varrho_o)$ therefore comes out to be about 0·2 e.Å^{-2}.

Discussion of the structure

The packing of the molecules in the unit cell, the intermolecular distances, the bond distances and the positions of the hydrogen atoms are all consistent with the following scheme for the hydrogen bond system in *p*-amino-salicylic acid:



The different bond distances found for C–O_I and C–O_{II} and the positions of the hydrogens show clearly that these groups are not equivalent. The C–O_I and C–O_{II} correspond respectively to the C–O group and to the C–OH group.

O_I forms two hydrogen bonds, an intramolecular one with the phenolic OH and an intermolecular one with the carboxylic OH of a molecule related by a centre of symmetry.

The intermolecular bridge is responsible for the relatively high melting point and low solubility of *p*-amino-salicylic acid, while the intramolecular hydrogen bridge might well explain the fact that the dipole moment of *p*-amino-benzoic acid does not change appreciably when the OH group of the *p*-

amino-salicylic acid is introduced (Iannelli, Giordano Orsini & Daniele, 1953).

In contrast to the results obtained for aliphatic amino acids, a hydrogen bridge between the carboxyl groups is preferred in this case to that between an amino group and the carboxyl group. This can be explained by the fact that the NH_2 group is conjugated with the benzene ring and therefore zwitterions cannot be formed. Thus the molecules are uncharged in the crystal.

Assuming that the four formulas shown in Fig. 6 contribute to the state of an uncharged molecule of *p*-amino-salicylic acid, the best agreement between measured and calculated distances was obtained assuming the following contributions:

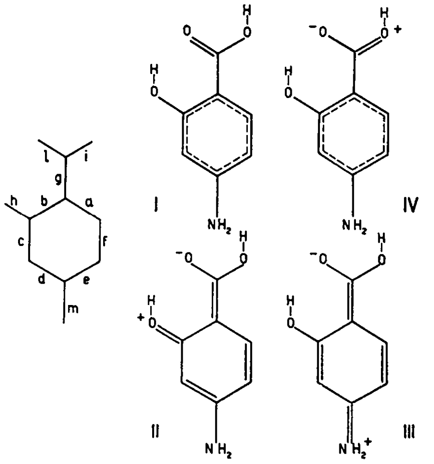


Fig. 6.

(I) 58%, (II) 9%, (III) 20%, (IV) 13% (Table 4).

The relation:

$$R = R_1-(R_1-R_2)\frac{3x}{2x-1}$$

was used with $R_1 = 1{\cdot}54$ Å, $R_2 = 1{\cdot}34$ Å for C–C, $R_1 = 1{\cdot}42$ Å, $R_2 = 1{\cdot}20$ Å for C–O and $R_1 = 1{\cdot}51$ Å, $R_2 = 1{\cdot}24$ Å for C–N (Pauling, 1948).

Table 4. *Comparison between measured and calculated bond lengths*

Bond	Measured distances (Å)	Percentage double bond	Calculated distances (Å)
a	1·43	28	1·42
b	1·45	21	1·42
c	1·38	55	1·38
d	1·40	43	1·40
e	1·44	25	1·42
f	1·39	50	1·37
g	1·41	39	1·43
h	1·37	9	1·37
i	1·35	13	1·35
l	1·21	87	1·24
m	1·39	21	1·39

The Kekulé formulas account for only about 60% of the total contribution, and this may explain several properties of this substance such as the high rate of decarboxylation, the low pK value, corresponding to the ionization of the amino group, and the marked variation of the absorption spectrum with the pH (Liquori & Ripamonti, unpublished).

We are indebted to Prof. R. Brill, guest of our Institute, for many helpful discussions. This work was carried on while one of us (F.B.) was on a research fellowship granted by the 'Fondazione Donegani'.

References

Booth, A. D. (1946). *Proc. Roy. Soc.* A, **188**, 77.
Brill, R. (1950). *Acta Cryst.* **3**, 333.
Burbank, R. D. (1953). *Acta Cryst.* **6**, 55.
Coccia, S., Giacomello, G. & Liquori, A. M. (1951). *Ric. Sci.* **21**, 205.
Cochran, W. (1948). *Acta Cryst.* **1**, 54.
Cochran, W. (1953). *Acta Cryst.* **6**, 260.
Cruickshank, D. W. J. (1949). *Acta Cryst.* **2**, 65.
Iannelli, L., Giordano Orsini, P. & Daniele, G. (1953). *R. C. Accad. Napoli*, **20**, 1.
Pauling, L. (1948). *The Nature of the Chemical Bond*. Ithaca: Cornell University Press.
Shoemaker, D. P., Donohue, J., Schomaker, V. & Corey, R. B. (1950). *J. Amer. Chem. Soc.* **72**, 2328.
Wilson, A. J. C. (1942). *Nature, Lond.* **150**, 152.