

THE CRYSTAL AND MOLECULAR STRUCTURE
OF PTERIDINE

with additional paper on

THE STRUCTURE OF CELLULOSE FIBRES

THESIS

PRESENTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE

UNIVERSITY OF GLASGOW

BY

THOMAS A. HAMOR, B.Sc.

University of Glasgow

August, 1956.

ProQuest Number: 13849061

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13849061

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Preface

This thesis describes research work carried out during a three year period in the Chemistry Department of the University of Glasgow.

I wish to express my gratitude to my supervisor, Professor J.M. Robertson, for suggesting the topics of research, and for his guidance and constant encouragement. I would also like to thank Dr.I.M. Dawson of this Department for his assistance and much helpful discussion in connection with the work on cellulose.

In conclusion, I am indebted to The Gourock Ropework Company, for providing samples of cellulose fibres and for a maintenance allowance.

Summary

The main part of the work described in this thesis is concerned with the determination of the crystal and molecular structure of pteridine by X-ray crystallographic methods. The unit cell dimensions and space group of 'substance P', a dipteridyl derivative of unknown structure were also obtained. An additional paper deals with an investigation of the structure of the cellulose fibre sisal with the electron microscope.

Pteridine crystallises in the orthorhombic non-centrosymmetrical space group $Pna2_1 (C_{2v}^9)$, with four molecules per unit cell. The unit cell dimensions are $a = 24.70$, $b = 3.79$, $c = 6.32 \text{ \AA}$. The crystal and molecular structure has been determined with coordinate standard deviations of about 0.03 \AA from a study of Fourier and Fourier difference syntheses projecting on (010) and (001). The molecule is planar to within 0.06 \AA and the bond lengths vary between 1.28 and 1.41 \AA for carbon-nitrogen bonds and between 1.35 and 1.42 \AA for carbon-carbon bonds. The experimental bond length measurements are compared with molecular orbital calculations and are found to agree to within the limits of the experimental errors.

The investigation was hampered by the fact that owing

to the nature of the crystals only less than half of the possible reflections could be observed with $\text{CuK}\alpha$ radiation. Additional difficulties were caused by the non-centrosymmetrical nature of the one projection which gives good resolution and also by the difficulty of distinguishing between carbon and nitrogen atoms.

The peak heights in the final electron density projection on (010) confirm the position of the nitrogen atoms, and the hydrogen atoms are also resolved.

The crystals of 'substance P' dihydrochloride dihydrate are monoclinic, space group Ia (C_2^4), with four molecules per unit cell. The unit cell dimensions are $a = 12.12$, $b = 11.21$, $c = 13.46 \text{ \AA}$; $\beta = 114^\circ$.

Specimens of sisal fibres prepared by degradation and thin sectioning techniques have been examined with the electron microscope. From the degradations particles of diameter down to about 90 \AA , and length $500\text{-}5000 \text{ \AA}$ were obtained. Sectioning was not successful unless non-cellulosic components were first removed, which, however, may have caused some degradation of the cellulose at the same time. Sections of the purified material show concentric layers of width about 1400 \AA and length up to 3 microns. The particles of the degradation have appeared only in sections of such material.

TABLE OF CONTENTS

	Page
Preface	i
Summary	ii
Introduction	1
 <u>THE CRYSTAL AND MOLECULAR STRUCTURE OF PTERIDINE</u>	
I INTRODUCTION	4
II CRYSTAL DATA	12
III EXPERIMENTAL	13
IV STRUCTURE DETERMINATION	17
(a) Choice of Space Group	17
(b) General Considerations	19
(c) Investigation of the (010) Projection	20
(d) Investigation of the (001) Projection	30
(e) Refinement of the (001) Projection	34
(f) Refinement of the (010) Projection	38
(g) Investigation of the (100) Projection	43
V RESULTS	44
(a) Coordinates and Molecular Dimensions	44
(b) Geometry of the Molecule	48
(c) Orientation and Intermolecular Distances	49
VI ACCURACY OF THE RESULTS	57
VII DISCUSSION	61
VIII FINAL STRUCTURE FACTORS	68

	Page
<u>INVESTIGATION OF THE STRUCTURE OF 'SUBSTANCE P'</u>	
I INTRODUCTION	76
II CRYSTAL DATA	77
III EXPERIMENTAL	78

APPENDICES

I The Refinement of Non-Centrosymmetrical Structures by Fourier Synthesis	80
II The Treatment of Accidentally Absent Reflections in Fourier Refinements	84

<u>REFERENCES</u>	86
-------------------	----

ADDITIONAL PAPER: INVESTIGATION OF THE STRUCTURE
OF CELLULOSE FIBRES

I INTRODUCTION	91
II EXPERIMENTAL	99
III RESULTS	104
IV DISCUSSION	106
V ELECTRON MICROGRAPHS	109

<u>REFERENCES</u>	116
-------------------	-----

INTRODUCTION

The discovery by von Laue in 1912 that X-rays are diffracted by crystals has led to the development of a powerful method of structure analysis. The final result of a complete X-ray crystallographic structure analysis of a substance is the determination of the positions of all the constituent atoms, relative to axes which are determined by the symmetry of the crystal.

In the field of inorganic chemistry this led to the discovery that in many cases the molecule of the chemical formula does not exist as such, the crystals consisting of a regular three-dimensional assemblage of ions. The size of the various ions can be measured, and it was found that the structure of a crystal is largely determined by the ratio of the numbers and by the ratio of the radii of the different ions present.

In organic crystals the fundamental entity is the molecule, and in such cases it is generally possible to determine the spatial positions of the atoms within the molecule, in a relative sense, by chemical methods and ordinary valence theory. Application of the X-ray method which, in general, is only possible after this has been achieved, then leads to quantitative measurements of the bond

lengths, valency angles and electron distribution in the molecule. Intermolecular distances are also obtained, and special forces like the hydrogen bond can be detected. Classical qualitative stereochemistry is thus given a metrical basis.

In a similar way, the discovery by Davisson, Germer and Kunsman, that electrons also are diffracted by crystals, led to the electron diffraction method of structure analysis which is, however, of rather more limited application. Electron waves, unlike X-rays, can be focussed by means of magnetic and electrostatic fields and this has been utilised in the electron microscope. The resolving power of this instrument is, at present, only about 20 \AA and it has been employed mainly in the study of the grosser features of crystal structure.

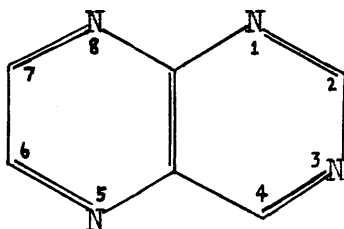
Other physical methods which may be used to determine molecular structures quantitatively include neutron diffraction and various types of spectroscopy, but these are also more limited in their application.

THE CRYSTAL AND MOLECULAR STRUCTURE

OF PTERIDINE

I INTRODUCTION

Pteridine is the parent substance of a number of naturally occurring compounds which were first discovered in 1889, when F.G. Hopkins¹ isolated a yellow and a white pigment from butterfly wings. At that time their structures were not known. However in 1940, Purrmann^{2,3} synthesized the two pigments, which had been named xanthopterin and leucopterin, and showed that they were derivatives of pyrimido [4:5-b] pyrazine. This compound, to which the name pteridine was given, remained unknown till 1948 when it was synthesized by Jones⁴.



Pteridine

A number of growth factors, which are derivatives of pteridine, have recently been discovered. These are composed of the units 2-amino-4-hydroxy-6-methylpteridine, p-aminobenzoic acid and a varying number of glutamic acid residues. They include pteroylglutamic acid, the fermentation L. casei factor, rhizopterin and the citrovorum factor. The three factors

which were originally known as folic acid, vitamin B_C and the liver L. casei factor were later shown to be identical with pteroylglutamic acid. It has been stated¹² that certain of these substances are essential for the process of cell division in mammals. As such, they would be of great importance in the study of growth and reproduction.

Pteridine may be considered as derived from naphthalene by substitution of nitrogen for carbon and hydrogen at the 1-, 3-, 5-, and 8-positions. In general this type of replacement in aromatic hydrocarbons leads to heterocyclic compounds which also show a high degree of aromatic character. Thus pyridine, pyrimidine, pyrazine and quinoline are stable aromatic-type substances.

This can be explained in terms of simple valence theory, since the electronic configuration of the heterocyclic compound is very similar to that of the corresponding aromatic hydrocarbon. However, owing to the greater electronegativity of the nitrogen atoms, there is a tendency for the π electrons to concentrate on them. A partial negative charge is therefore associated with each nitrogen atom, balanced by positive charges on the carbon atoms. In general this localisation of the π electrons is only slight and does not greatly interfere with the aromatic stabilisation.

In the course of an extensive study of the simpler natural and synthetic pteridines, Albert⁵⁻¹³ has found that they are unstable unless electron-releasing substituents, such as hydroxyl or amino groups, are present. Pteridine itself is highly unstable, being rapidly decomposed by dilute acid and alkali even at room temperature.⁷ In solution it is decomposed on exposure to day-light.⁵

Albert considers that the instability of pteridine, compared to the mono-, or diazanaphthalenes such as quinoline or quinazoline, is due to the greater degree of localisation of the π electrons, caused by the high nitrogen to carbon ratio of 4:6 in its ring system. The aromatic stabilisation to be expected from the presence of the π electrons is therefore much diminished, and it is suggested that pteridine may, in consequence, be non-planar. Substitution of the pteridine nucleus by electron-releasing groups counteracts the localisation of the π electrons and thus increases the aromatic stabilisation.^{7,12}

A similar case is the instability of 1:3:5-triazine compared to its triamino-, and trihydroxy- derivatives, melamine and cyanuric acid.

The object of this investigation was to determine the crystal and molecular structure of pteridine by X-ray diffraction methods.

No structural investigation by X-ray methods of a pteridine, or of any other azanaphthalene, has so far been reported. The related pyrimidine, pyrazine and purine ring systems have, however, been examined by X-ray and electron diffraction methods.

The X-ray work on the pyrimidines includes a three-dimensional determination of the structure of 4-amino-2:6-dichloropyrimidine¹⁴ and less accurate two-dimensional studies of certain other trisubstituted derivatives such as 2-amino-4:6-dichloropyrimidine¹⁵ and 4:6-dimethyl-2-hydroxypyrimidine dihydrate¹⁶. The pyrimidine ring was found to be planar, within the limits of the experimental errors, but it does not have the regular hexagonal shape of the benzene ring. For pyrimidine itself, Clews and Cochran¹⁴ have suggested lengths of 1.38 Å and 1.33 Å respectively, for the carbon-carbon and carbon-nitrogen bonds and valency angles of 130°, 114°, 124° and 114° for the NCN, CNC, NCC and CCC angles. These values were, however, deduced from the results of the investigations of the substituted pyrimidines and are, therefore, only approximate.

Only one structural investigation by X-ray methods of a pyrazine derivative has been reported. Cromer, Ihde and Ritter¹⁷, using three-dimensional data, found that the pyrazine ring in tetramethyl pyrazine is planar, the carbon-

carbon and the two crystallographically independent carbon-nitrogen bonds having lengths of 1.44 Å, 1.30 Å and 1.32 Å respectively. All the valency angles were found to be 120°, to within 2°.

Pyrazine itself has been examined in the gas phase by electron diffraction methods.¹⁸ Its structure was found to be similar to that of benzene, the atoms of the ring forming a planar hexagon. By assuming a length of 1.39 Å for the carbon-carbon bonds, the carbon-nitrogen bond lengths were measured as 1.35 Å.

Of substances containing the purine ring system the hydrated hydrochloride salts of adenine^{19,20} and guanine²¹ have been studied by X-ray diffraction methods, using two-dimensional data. In both substances, the purine ring was found to be planar, within the limits of the experimental errors. In these investigations, however, poor resolution in some of the projections prevented the accurate determination of all the atomic coordinates. In the case of adenine hydrochloride hemihydrate, accurate molecular dimensions were determined on the assumption of a planar ring system.²⁰ The carbon-nitrogen bond lengths were found to vary from 1.30 Å to 1.38 Å and the carbon-carbon bond lengths from 1.37 Å to 1.40 Å.

1:3:5- Triazine²² and its derivatives melamine²³ and

cyanuric acid²⁴ have also been examined by X-ray methods. All three molecules were found to be planar, in the case of 1:3:5- triazine an accurately planar structure being necessitated by symmetry considerations. In the investigation of 1:3:5- triazine three dimensional data were employed and the carbon-nitrogen bond was found to have a length of $1.319 \text{ \AA} \pm 0.005 \text{ \AA}$, the valency angles being 126.8° and 113.2° (both $\pm 0.4^\circ$) at carbon and nitrogen respectively. The melamine molecule has similar dimensions, but in cyanuric acid the relative magnitudes of the angles at carbon and nitrogen are reversed. The angles, averaged over chemically equivalent positions, are given as 115° for a carbon valency angle and 125° for a nitrogen valency angle.

The experimental work quoted above, and other work such as the X-ray investigation of δ -phenazine²⁵ and the examination of pyridine by microwave spectroscopy²⁶ and electron diffraction,¹⁸ indicates that the arrangement of the atoms in the six-membered nitrogen containing aromatic ring systems is in the form of a planar hexagon which, however, unlike the benzene ring is not regular. In general the carbon-nitrogen bonds are somewhat shorter than the carbon-carbon bonds and the valency angles at the nitrogen atoms are smaller than at the carbon atoms. The differences are, however, small and there is only a slight distortion of the regular hexagonal shape of the benzene ring.

These results were obtained mainly from investigations of ring systems containing a lower nitrogen to carbon ratio than the pteridine system, and except in the case of 1:3:5- triazine, electron-releasing substituents were present. Any deviations from a planar structure would be small, and would not have been detectable by the methods employed. In 1:3:5- triazine the nitrogen to carbon ratio is higher than in pteridine and no substituents are present. The arrangement of the nitrogen and carbon atoms is however different, and unlike pteridine the molecule consists of a single ring. The non-planar structure suggested for pteridine,^{7,12} although unlikely, is therefore not necessarily excluded. It is of interest that in the recent very accurate X-ray investigation of acridine,²⁷ the molecule was found to be non-planar, with the atoms situated at distances of up to 0.051 Å from the mean molecular plane. The carbon-carbon bonds in acridine vary from 1.364 Å to 1.439 Å and the carbon-nitrogen bond lengths are 1.342 Å and 1.347 Å.

It was hoped that the present investigation would provide reliable measurements of the shape and dimensions of the pteridine molecule. This would settle the question of its planarity and also, by comparison with the accurately determined dimensions of naphthalene,^{28,29} give a quantitative measurement of the effect on the dimensions of an aromatic

system, of substituting nitrogen for carbon and hydrogen at a relatively large number of positions. Comparison of the general arrangement of the molecules in the two crystals would indicate the effect of this substitution on packing. Bond lengths have been calculated for pteridine by the method of molecular orbitals³⁰ and it was hoped to compare these with the experimental values.

The information derived from this investigation would also be of value in the study of the structures of the naturally occurring, biologically active pteridines, which are themselves rather complex for analysis by X-ray methods.

II CRYSTAL DATA

Pteridine, $C_6H_4N_4$. Molecular weight, 132.1.

Melting point, 139°C.

Density calculated, 1.483gm./cm³ Density found, 1.487gm./cm³

The crystal is orthorhombic with

$$a = 24.70 \pm 0.05 \text{ \AA},$$

$$b = 3.79 \pm 0.01 \text{ \AA},$$

$$c = 6.32 \pm 0.01 \text{ \AA}.$$

Volume of the unit cell, 591.6 Å³

Absent spectra, (Ok l) when $k + l$ is odd.

(h0 l) when h is odd.

Space group, Pna2₁ (C_{2v}⁹) or Pnam (D_{2h}¹⁶).

Structure analysis indicates Pna2₁.

Four molecules per unit cell.

No molecular symmetry required.

Absorption coefficient for X-rays, CuK α radiation, $\mu = 9.74\text{cm.}^{-1}$

Total number of electrons per unit cell = F(000) = 272.

III EXPERIMENTAL

Preparation of the Crystals.

The sample of pteridine used in this investigation was provided by Professor A. Albert⁵. The crystals, prepared by vacuum sublimation at 125-130°C/20mm., were in the form of very fine rectangular plates. The thickness of these was only about 0.02-0.03mm., and attempts to obtain better crystals by varying the experimental conditions were not successful. Owing to the very great solubility of pteridine in most solvents, no suitable crystals could be obtained by crystallisation from solution.

The plates are developed on the (100) face with the b and c axes lying in the plane of the plate parallel to its edges.

Stability.

Pteridine was found to be stable to X-rays and to the atmosphere. It is however slightly volatile, a crystal of the dimensions used in the investigation volatilising completely in one to three weeks depending on the temperature of the laboratory. At 0°C the life-time of a crystal was greatly increased, and by keeping the temperature of the laboratory as low as possible (0-5°C), the necessary X-ray diffraction pictures could be taken before a large part of

the crystal had volatilised. No attempts were made to preserve the crystals by sealing them inside a capsule as this would have introduced errors due to the absorption of the X-rays.

Determination of Crystal Data.

Copper $K\alpha$ radiation, $\lambda = 1.542 \text{ \AA}$, was employed, the $K\beta$ radiation being eliminated by means of a nickel filter. Rotation, oscillation and equatorial layer line moving film photographs were taken about each of the unit cell axes. In addition, a first layer line moving film about b was taken to record the general reflections of form (hll) . The axial lengths were measured from the rotation photographs. The oscillation photographs taken about the cell axes showed symmetry, indicating that the crystal system was orthorhombic. This was also evident from the moving film photographs, from which the systematic absences were determined.

The density of the crystals was determined by flotation in a mixture of pet-ether and n-propyl iodide containing a trace of detergent.

Measurement of the Intensities of the X-ray Reflections.

The moving film photographs were employed to record the $(0kl)$, $(h0l)$ and $(hk0)$ reflections which were used in the structure determination. The intensities were estimated visually on a relative scale by the multiple film technique.³¹

For the (h0l) and (hk0) zones, the crystals employed had cross-sections perpendicular to the rotation axes of about 0.5 x 0.02mm., and absorption corrections were worked out graphically by measuring the length of the path through the centre of the crystal for each reflection.³² The correction factor varied from 1.02 to 1.31. For the (Ok ℓ) zone the crystal cross-section was more uniform and absorption corrections were not applied. The intensities were corrected for Lorentz and polarisation factors, and the resulting structure amplitudes, $|F_o|$, scaled by comparison with the calculated values, are given in Table 8. In Table I are shown the dimensions of the crystals employed and also the number of reflections observed in each zone at the commencement of the investigation.

Table I.

Zone of reflections	Cross-section of crystal perpendicular to rotation axis	Reflections observed	Reflections possible for CuK α radiation	Range of intensities
(Ok ℓ)	0.70 x 0.60 mm.	9	18	670:1
(h0l)	0.50 x 0.03 mm.	50	110	2700:1
(hk0)	0.50 x 0.02 mm.	33	120	2000:1

An additional 16 (hk0) reflections were later recorded by means of a very long exposure to unfiltered radiation, making a total of 49 (hk0) reflections observed. The range of intensities was thereby extended to 5000:1.

The extreme thinness of the crystals was a source of considerable difficulty. Even with the very long X-ray exposures of up to 55 hours which were employed, only about 45% of the possible reflections could be recorded. There was also a tendency for the crystals to bend when mounted on a glass fibre.

IV STRUCTURE DETERMINATION

(a) Choice of Space Group.

A statistical examination of the intensities of the reflections by means of the 'ratio' test³³ and the 'N(z)' test³⁴ failed to distinguish between the two possible space groups, $Pna2_1$ which is non-centrosymmetrical and $Pnam$ which is centrosymmetrical. The results obtained were intermediate between the two cases. A possible explanation for this is given later (see page 33). The $N(z)$ distribution for the $(h0l)$ reflections of pteridine is shown in Fig.1, and it may be seen that it is intermediate between the theoretical 'acentric' and 'centric' curves. Polarity tests were also inconclusive.

Since there are four molecules in the unit cell, the higher symmetry space group $Pnam$, requiring eight asymmetric units, would be possible only if the molecules possessed either a mirror plane or a centre of symmetry and occupied special positions. The pteridine molecule would, if planar, possess a mirror plane and might also achieve centrosymmetry by random packing. However, considerations of packing requirements show that it is not possible to fit the molecules into the unit cell on the basis of this space group.

The non-centrosymmetrical space group $Pna2_1$, requiring four asymmetric units was therefore assumed. No molecular symmetry need then be postulated and the four molecules occupy general positions in the unit cell.

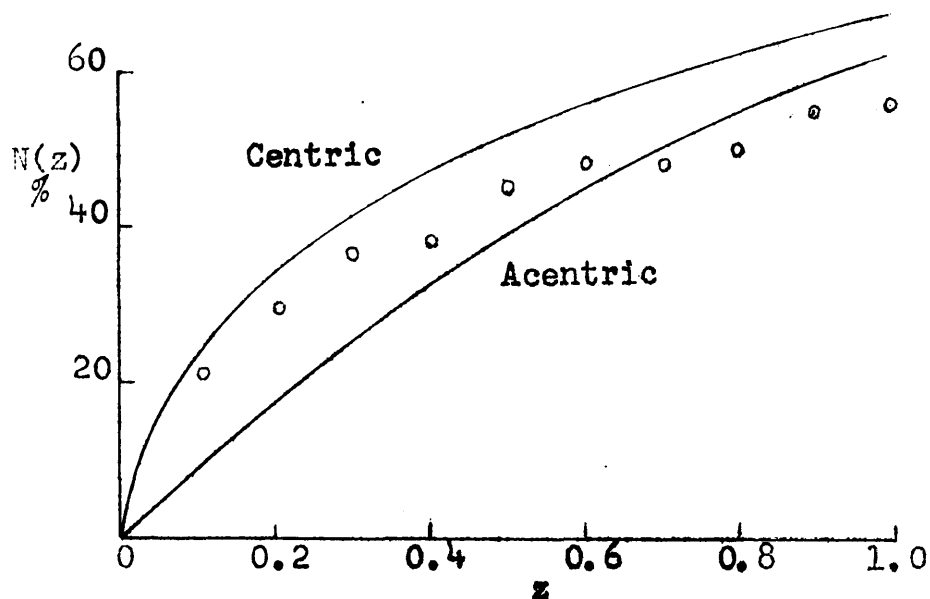


Fig. 1.

$N(z)$ Distribution for pteridine compared with theoretical curves for non-centrosymmetrical and symmetrical structures, where $N(z)$ is the fraction of the reflections whose intensities are equal to, or less than a fraction z of the local average intensity.

(b) General Considerations.

In order to obtain a trial structure it was necessary to postulate dimensions for the pteridine molecule. No attempt was made to distinguish between carbon and nitrogen or to take into account any variations in the bond lengths. It was assumed that the molecule consists of identical atoms (carbon) positioned at the vertices of two regular coplanar hexagons with one side in common. Taking into consideration the dimensions of the naphthalene molecule,^{28,29} a bond length of 1.40 Å was assumed. The hydrogen atoms were not considered at this stage. The postulated model, unlike the actual molecule, thus contains a centre of symmetry and mirror planes.

Although space group $Pna2_1$ does not possess a centre of symmetry, the (001) projection, being perpendicular to the twofold axis, is effectively centrosymmetrical. However, owing to the shortness of the b axis the molecules must lie very nearly in the (010) plane and this projection of the structure should be well resolved. The resolution in the (001) projection would be much inferior and it was therefore decided to consider first the non-centred (010) projection. The symmetry elements, coordinates of equivalent positions, and structure factor and electron density formulae for space group $Pna2_1$ are given in International

Tables³⁵. The origin of coordinates is on the twofold screw axis parallel to c , the origin of the z coordinates being therefore arbitrary.

(c) Investigation of the (010) Projection.

Examination of the $(h0l)$ zone of reflections showed that the axial reflections $(h00)$ were present only when $h = 4n$, the highest order reflection observed being 24,0,0. The symmetry of the space group however permits all the $(h00)$ reflections with $h = 2n$ to be present. The most probable interpretation of this quartering is that the four molecules in the unit cell are spaced equally along the a axis and that the molecules are symmetrical, or at least nearly so, with respect to the x -coordinates. This is obviously in agreement with the centrosymmetrical trial model which had been postulated.

The x coordinate of the centre of the reference molecule is therefore, to a first approximation, $a/8$. The origin of the z coordinates being arbitrary, the z coordinate of the centre of this molecule was taken to be zero. The trial structure in this projection thus consists of a centrosymmetrical model with its centre at the point $x = a/8, z = 0$.

In space group $Pna2_1$, the calculated structure factor

for a reflection ($h0l$) is given, in terms of its amplitude $|F|$ and phase constant α , by

$$F(h0l) = (A^2 + B^2)^{\frac{1}{2}},$$

$$\alpha(h0l) = \tan^{-1} B/A,$$

where, when l is even,

$$A = 4 \sum f \cos 2\pi hx \cos 2\pi lz,$$

$$B = 4 \sum f \cos 2\pi hx \sin 2\pi lz,$$

and when l is odd,

$$A = -4 \sum f \sin 2\pi hx \sin 2\pi lz,$$

$$B = 4 \sum f \sin 2\pi hx \cos 2\pi lz.$$

In these equations x and z are the atomic coordinates expressed as fractions of the corresponding cell axes and f represents the appropriate atomic scattering factor. The summation is taken over all the atoms in the asymmetric unit, which in this case is one molecule. However, for the special case of the centrosymmetrical triad structure which has been postulated, the expressions given above simplify further.

When l is even,

$$A = 8 \sum f \cos 2\pi hx \cos 2\pi lz$$

$$B = 0$$

$$\text{when } h = 4n,$$

and

$$A = 0$$

$$B = 8 \sum f \cos 2\pi h x \sin 2\pi l z \quad \text{when } h = 4n + 2,$$

and when l is odd,

$$A = -8 \sum f \sin 2\pi h x \sin 2\pi l z$$

$$B = 0 \quad \text{when } h = 4n,$$

and

$$A = 0$$

$$B = 8 \sum f \sin 2\pi h x \cos 2\pi l z \quad \text{when } h = 4n + 2,$$

the summation being taken over half the molecule. It follows that the calculated phase constants are restricted to multiples of 90° . Thus by postulating a centrosymmetrical model with its centre at the position $(a/8, 0)$, a centrosymmetrical projection has, in effect, been obtained.

The problem of determining the orientation of the model in the unit cell was also simplified by the shortness of the projection axis, $b = 3.79 \text{ \AA}$, since if the normal van der Waals distance of 3.4 \AA is to be maintained between molecules separated by a b translation, the maximum possible tilt of the model out of the (010) plane is 26° . The permitted orientations are thus limited.

For the purpose of this investigation, the ring atoms

of the molecule were numbered 1, 2, 3, 4, 5, 1', 2', 3', 4', and 5', such that atoms 1 and 1', 2 and 2', 3 and 3' etc., are related by the pseudo-centre.

Structure factors were calculated for various orientations of the trial model. It was eventually found that the best agreement between the observed and calculated structure factors, F_o and F_c , resulted when the long axis of the molecule (L) was at an angle of about 45° to the a axis in the (010) plane and the normal to the molecular plane (N) was inclined at about 26° to b in (100). This structure gave a discrepancy, expressed as $\sum |F_o - F_c| / \sum |F_o|$, of 21.8% for the 50 observed (h0l) structure factors. The atomic scattering factors for these preliminary calculations were derived from the empirical scattering curve used in the investigation of the naphthalene structure²⁸.

The observed structure factors and the calculated phase constants were now employed to evaluate a Fourier synthesis (F_o1) projecting on (010).

In space group $Pna2_1$, the projected electron density at points (xz) is given by

$$\rho(xz) = 1/ac \left\{ \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} |F(h0l)|^{\ell \cdot 2n} \cos 2\pi hx \cos [2\pi lz - \alpha(h0l)] - \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} |F(h0l)|^{\ell \cdot 2n+1} \sin 2\pi hx \sin [2\pi lz - \alpha(h0l)] \right\} ,$$

where x and z are expressed as fractions of the cell axes. In this special case, however, the phase constants are restricted to multiples of 90° and the electron density is given by the simplified series:

$$\rho(xz) = 1, - \sum_{l=2n+1, h=4n}^{\infty} \sum_{l=2n, h=4n}^{\infty} A(h0l) \sin 2\pi hx \sin 2\pi lz + \sum_{l=2n+1, h=4n+2}^{\infty} \sum_{l=2n, h=4n+2}^{\infty} B(h0l) \sin 2\pi hx \cos 2\pi lz \}.$$

$$- \sum_{l=2n+1, h=4n}^{\infty} \sum_{l=2n+1, h=4n+2}^{\infty} A(h0l) \sin 2\pi hx \sin 2\pi lz + \sum_{l=2n+1, h=4n+2}^{\infty} \sum_{l=2n+1, h=4n+2}^{\infty} B(h0l) \sin 2\pi hx \cos 2\pi lz \}.$$

The summation was carried out over one quarter of the unit cell, $a/4$ by c , the axial subdivisions being $a/120 = 0.2058 \text{ \AA}$ and $c/30 = 0.2107 \text{ \AA}$.

This synthesis which contained 46 structure factors gave good resolution of all the ring atoms. The electron density distribution was, however, centrosymmetrical about $(a/8, 0)$. The Fourier peaks representing atoms 1, 2, 3, 4 and 5 were merely reflections across this point of the peaks representing atoms 1', 2', 3', 4', and 5'. Each of the five independent peaks may thus be considered as representing the average of the contributions to the electron density of the two atoms, really independent, which are related by the pseudo-centre.

Before the projection could be refined it was necessary to eliminate this centre of symmetry and separate atoms 1 and 1', 2 and 2', 3 and 3' etc. One way of achieving

this would be by shifting individual atoms or the whole molecule in such a way that the atomic coordinates would no longer be centrosymmetrical about the position $(a/8,0)$. This procedure would, however, be rather arbitrary and was not adopted. Instead, it was decided to insert into the trial model the nitrogen and hydrogen atoms which are present in the pteridine molecule. By distinguishing between carbon and nitrogen and by taking into account the hydrogen contributions the centre of symmetry is eliminated without necessitating any arbitrary shifts of the atomic positions. There are, however, now four possible orientations of the asymmetric model in the unit cell, as shown in Fig.2. These involve identical coordinates for the ten ring atoms and differ only in the positions occupied by the nitrogen and hydrogen atoms. In the four structures which are designated A, B, C, and D, the nitrogen atoms occupy the following positions:-

Structure	A	1	3	1'	4'
Structure	B	2	4	1'	4'
Structure	C	1	4	1'	3'
Structure	D	1	4	2'	4'

Corresponding to these arrangements of the nitrogen atoms, hydrogen atoms are linked to the following positions

occupied by carbon atoms:-

Structure	A	2	4	2'	3'
Structure	B	1	3	2'	3'
Structure	C	2	3	2'	4'
Structure	D	2	3	1'	3'

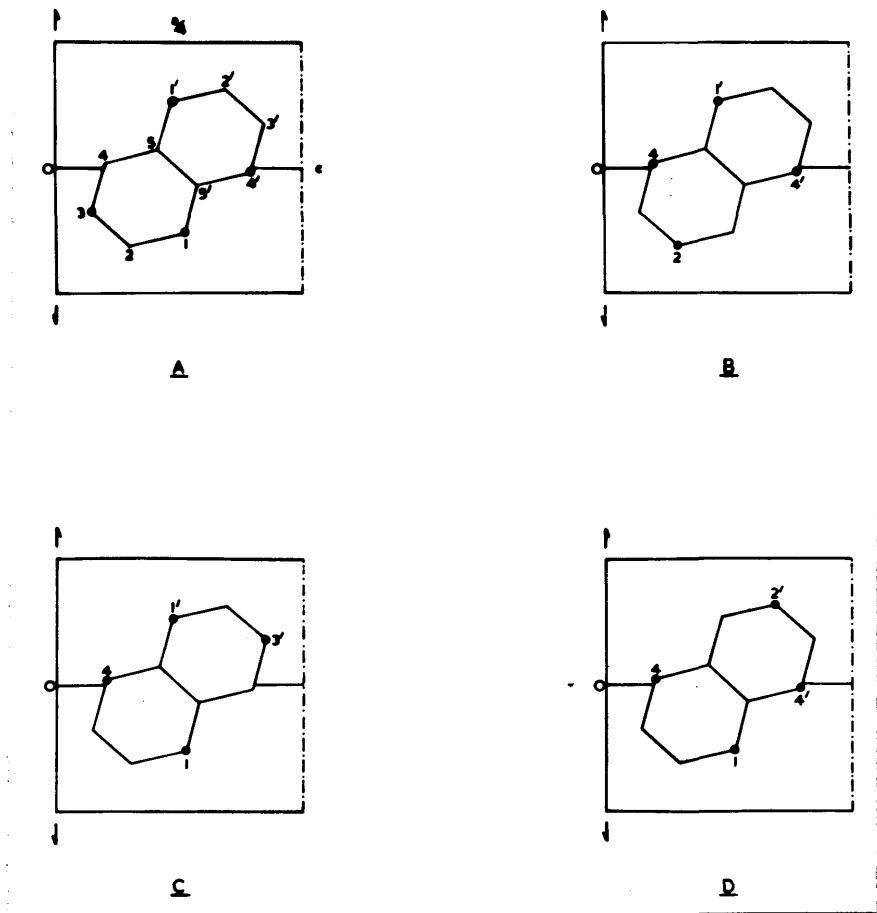


Fig. 2.

Projection of one quarter of the unit cell on (010), showing the four possible orientations of the asymmetric model in relation to the symmetry elements.

In the diagrams • represents nitrogen.

The structures A, B, C, and D are chemically identical and merely represent four orientations of the same molecule in the unit cell. As shown in Fig.2, they are related to each other by rotations of 180° about axes passing through the molecular centre. These structures are however crystallographically different. The choice of the correct structure depends only on the difference in the scattering power for X-rays of carbon and nitrogen, and on the hydrogen contributions, since the positions of the ten ring atoms are the same in each case. It was therefore a matter of some difficulty to choose the correct structure and this constituted one of the main problems of the investigation.

An additional difficulty was encountered when attempts were made to distinguish between these structures using the $(h0l)$ structure factors. In the (010) projection the 2_1 screw axis and the n glide plane cannot be distinguished, and the origin of coordinates is therefore not uniquely determined by considering this projection alone. Thus, since structure A with respect to 2_1 as origin is the same as structure C with respect to n as origin, they cannot be distinguished on the basis of the (010) projection. These structures give identical values for $|F(h0l)|$. Similarly structures B and D are indistinguishable in this projection.

This is the case only for the symmetrically shaped

model at the pseudo-special position $(a/8,0)$, and a distinction could have been achieved by arbitrarily shifting atomic coordinates. This procedure, which would have been rather unsatisfactory, was in fact unnecessary, since in the (001) projection all these structures can be distinguished.

Within the limitations stated above, information regarding the positions of the nitrogen and hydrogen atoms can, however, be obtained from the (010) projection data.

It was found that in the (010) Fourier synthesis, F_{01} , based on the symmetrical trial model, the electron density peak values were

Atoms 1 and 1'	6.2 $e/\text{\AA}^2$
Atoms 2 and 2'	4.8 $e/\text{\AA}^2$
Atoms 3 and 3'	5.3 $e/\text{\AA}^2$
Atoms 4 and 4'	5.5 $e/\text{\AA}^2$
Atoms 5 and 5'	5.3 $e/\text{\AA}^2$

It would seem probable that the highest peak represents two nitrogen atoms, and the lowest, two carbon atoms. If this is actually the case, then atoms 1 and 1' are nitrogen, and atoms 2 and 2' are carbon, corresponding to structures A or C.

Also, when the $(h0l)$ structure factors were calculated for the four cases, the discrepancies were 18.3% for

structures A and C, and 21.7% for B and D.

The evidence obtainable from the (010) projection is thus in favour of structures A or C.

In the calculation of the structure factors, the hydrogen contributions were included. It was assumed that the hydrogen atoms are situated radially at a distance of 1.09 Å from the carbon atoms, the carbon-hydrogen bond length being based on the values found in the naphthalene²⁸ and anthracene³⁶ structures. The atomic scattering factors for nitrogen and hydrogen were taken as 7/6 and 1/6 respectively, of the corresponding values for carbon. When the structure factors were later recalculated using theoretical carbon, nitrogen and hydrogen scattering curves, it was found that the discrepancy figures given above had not altered appreciably.

(d) Investigation of the (001) Projection.

On the basis of the trial structure obtained by consideration of the (010) projection, y coordinates for the atoms were calculated with respect to the centre of the model as origin. The y coordinate of the centre of the molecule, unlike the z coordinate, is, however, not arbitrary and the position of the molecule along the b axis has to be determined.

Examination of the (hk0) zone of reflections showed that the reflections (h10) were absent when $h = 4n$, and the reflections (h20) were comparatively very weak when $h = 4n+2$. It was not found possible to make direct use of these observations, but they seemed to indicate that the molecule might, to a first approximation, be in some pseudo-special position also in the y direction.

In space group $Pna2_1$, the (001) projection is centrosymmetrical and the calculated structure factor for a reflection (hk0) is given in terms of the fractional atomic coordinates x and y , by

$$F(hk0) = 4\sum f \cos 2\pi hx \cos 2\pi ky, \quad \text{when } h+k \text{ is even}$$

and

$$F(hk0) = -4\sum f \sin 2\pi hx \sin 2\pi ky, \quad \text{when } h+k \text{ is odd.}$$

The summation is taken over all the atoms in the asymmetric unit (one molecule in this case).

Structure factors were calculated for various positions of the symmetrical trial model along the b axis, the best agreement between the observed and calculated structure factors being obtained when the centre of the model had the y coordinate $b/4$. The discrepancy for the 33 observed (hk0) structure factors was 21.5%. As additional evidence for the correctness of this structure, it was found that a value of

zero was obtained for the calculated structure factors for the reflections (h10) when $h=4n$, and (h20) when $h=4n+2$. The atomic scattering factors for these calculations were derived, as before, from the scattering curve used in the investigation of naphthalene²⁸.

On the basis of this symmetrical trial structure a Fourier projection on (001), F_02 , was evaluated. This synthesis contained 29 structure factors.

The (001) projection being centrosymmetrical and the phase constants accordingly restricted to 0° or 180° , the

$$\rho(xy) = 1/ab \left\{ \sum_{h+k=2n} F(hk0) \cos 2\pi hx \cos 2\pi ky - \sum_{h+k=2n+1} F(hk0) \sin 2\pi hx \sin 2\pi ky \right\}$$

$$\rho(xy) = 1/ab \left\{ \sum_{h+k=2n} F(hk0) \cos 2\pi hx \cos 2\pi ky - \sum_{h+k=2n+1} F(hk0) \sin 2\pi hx \sin 2\pi ky \right\}$$

where x and y are expressed as fractions of the cell axes.

The summation was carried out over one quarter of the unit cell, $a/4$ by b , the axial subdivisions being $a/120=0.2058 \text{ \AA}$ and $b/15=0.2527 \text{ \AA}$.

In this projection of the structure, the resolution is poor and only the atoms 2 and 2' were resolved. The remaining atoms overlapped in pairs. It was however found that the electron density peak values on these two atoms were approximately the same. This would seem to indicate that atoms 2 and 2' are of the same type, which again corresponds to structures A or C.

The structure which has so far been derived, consisting of a centrosymmetrical model with its centre at $(a/8, b/4, 0)$ appears to be basically correct. It gives a discrepancy of about 22% for 78 $(h0l)$ and $(hk0)$ structure factors, and a preliminary consideration of intermolecular approach distances showed no abnormally close contacts.

That the centrosymmetry of the model and its special position do not hold strictly, is shown by the presence of certain reflections, such as 620 and $10,2,0$, and also 410 and 810 , which were recorded later by means of a very long exposure. As mentioned above, these reflections calculate zero for the symmetrical structure.

The restriction on the phases of the $(h0l)$ reflections is therefore not due to symmetry requirements, as is the case for the $(hk0)$ reflections. The values of 0° , 90° , 180° and 270° found for $\alpha(h0l)$ merely represent a first approximation to the true phase constants which are permitted to have any value between 0° and 360° . The phase constants which were finally obtained are shown in Table 8. It may be seen, that although the $(h0l)$ reflections have general phases, the majority of these are in fact multiples of 90° , to within about 10° .

The statistical tests which had been employed to distinguish between the possible space groups (see page 17)

depended on the fact that in the non-centrosymmetrical space group $Pna2_1$, the $(h0l)$ structure factors are complex quantities, whereas in the centrosymmetrical space group $Pnam$, they are real. The close approximation of the (010) projection to centrosymmetry may therefore account for the failure of the statistical tests in this case.

(e) Refinement of the (001) Projection.

In the (001) projection of the structure, the atoms of the molecule are rather close together and the electron density projection F_02 which had been evaluated, resolved only atoms 2 and 2'. Detailed refinement was therefore not possible by means of ordinary Fourier syntheses and use was accordingly made of the Fourier difference synthesis, by means of which unresolved projections may be refined to a somewhat higher degree.

Starting with the symmetrical trial structure, two successive (001) difference syntheses, D1 and D2, were evaluated. The discrepancy for the 33 $(hk0)$ structure factors was thereby reduced from 21.5% to 15.1%.

In these syntheses, since all the ring atoms were treated as being identical carbon atoms and the hydrogen contributions were omitted, the resulting electron density distribution showed the difference between the observed

electron density and that calculated for a model consisting only of carbon atoms. The nitrogen and hydrogen atoms should therefore appear as electron density peaks. These would, however, be obscured to some extent because of errors in the atomic coordinates and also because of the thermal motion of the atoms which had not been allowed for at this stage.

The expected electron density peaks did not in fact appear, but it was found that both in D1 and D2 the electron density was positive at atoms 1, 4, 1', 3' and 4' and negative at the other ring atoms. Since it is unlikely that a nitrogen atom would be in a region of negative electron density, the structures A, B, and D are excluded. The hydrogen atoms, however, could not be located.

Further information concerning the positions of the nitrogen and hydrogen atoms was obtained by calculating structure factors for the four possible orientations on the basis of the atomic coordinates derived from D2. The discrepancy figures obtained were,

Structure A	15.2%,
Structure B	14.9%,
Structure C	13.9%,
Structure D	17.3%.

The lowest discrepancy is thus given by structure C, which is in agreement with the results previously obtained both from this and the (010) projection.

As a final test, the five (001) difference syntheses, D3 - D7, corresponding to the all-carbon model and the structures A, B, C and D were evaluated. Consideration of the electron densities at the atomic positions in these syntheses showed quite conclusively that atoms 1, 4, 1' and 3' are nitrogens, corresponding to structure C.

In the calculation of these structure factors the theoretical atomic scattering curves for carbon, nitrogen and hydrogen given by McWeeny³⁷ were employed, modified by a temperature factor $\exp[-B(\sin\theta/\lambda)^2]$, where $B = 4.8 \text{ \AA}^2$. This value for B was obtained from a comparison of the observed (hk0) structure factors with those calculated for structure C on the basis of the coordinates derived from D2. As stated above, the discrepancy at this stage was 13.9%. The validity of this temperature factor was shown by the results of subsequent difference syntheses, and these scattering curves were used in all further work.

Structure C having thus been established, the refinement process was now continued with the inclusion of the nitrogen and hydrogen atoms. The atomic positions were shifted according to the difference synthesis corresponding

to structure C, and three further difference syntheses, D8 - D10, then reduced the discrepancy for the 33 (hk0) structure factors to 9.4%. The coordinate shifts indicated by D10 were very small, and on the basis of the experimental data available, the projection seemed to have been refined to the limit.

At this stage, however, another 16 (hk0) reflections were recorded by means of a very long exposure to unfiltered copper radiation. Structure factors were calculated for the additional reflections, and it was found that the discrepancy for the (hk0) zone had risen to 13.1%. The atomic coordinates were now refined further by means of two successive difference syntheses, D11 and D12. The shifts indicated by D12 were all quite small and the resulting coordinates, which are corrected for termination of series errors³⁸, were taken as the final atomic coordinates for the (001) projection. The final discrepancy for the 49 (hk0) structure factors was 12.2%.

During the refinement of this projection it had been noticed that a small number of low angle reflections which were too weak to be observed, calculated comparatively high, and terms corresponding to these were included in the last two difference syntheses (D11 and D12). The observed structure amplitudes for these vanishingly small reflections were taken as half the estimated maximum value. A justification of this

procedure is given in Appendix II.

The final electron density projection on (001), F_03 , containing 46 observed structure factors is shown in Fig.4. Each of the four pairs of unresolved atoms consists of one carbon and one nitrogen atom. In every case the electron density is greater at nitrogen than at carbon and the peak densities on the resolved atoms 2 and 2' are again approximately the same. This seems to confirm the validity of the orientation of the molecule corresponding to structure C.

It is of interest that although quite large coordinate shifts were involved in the refinement of this projection (up to 0.14 \AA in x and 0.16 \AA in y), there had been no alteration in the phase constants of the 29 largest structure factors which had been determined on the basis of the symmetrical trial model. The initial (001) Fourier projection, F_02 , which contained only these 29 structure factors is, in fact, very similar to F_03 , the projection finally obtained.

(f) Refinement of the (010) Projection.

The previous work in this and in the (001) projection had established quite conclusively that the atoms 1, 4, 1', and 3' are nitrogens and that the hydrogen atoms are linked to the carbon atoms 2, 3, 2', and 4', corresponding to

structure C. By inserting the nitrogen and hydrogen atoms into the symmetrical trial model in this way, an asymmetric model was obtained. The pseudo-centrosymmetry of the (010) projection was thus eliminated and it was now possible to proceed with the refinement of this projection of the structure.

The (h0l) structure factors for structure C, which have unrestricted phases, had been calculated previously (see page 29) and showed a discrepancy of 18.3%. Using these calculated phase constants and the observed amplitudes, a Fourier synthesis (F_04) projecting on (010) was evaluated. Unlike F_01 , this synthesis, which contained 48 terms, was not centrosymmetrical, and independent electron density peaks were obtained for all the atoms in the molecule.

The coordinates of the electron density maxima were determined by means of two different analytical methods, a simple treatment due to Booth,³⁹ and a more accurate but laborious least squares treatment due to Shoemaker et al.⁴⁰ Almost identical results were obtained and in all subsequent work the Booth treatment was used to locate the positions of the electron density maxima of the ring atoms. The positions of the hydrogen atoms were calculated on the assumption that they are situated radially at a distance of 1.09 Å from the carbon atoms.

The structure derived in this way from the Fourier

synthesis F_0 gave a discrepancy of 11.3% for the 50 ($h0l$) structure factors.

The further refinement of the (010) projection was carried out by means of a method described by Donohue.⁴¹ Two Fourier syntheses are employed at each stage, one containing the observed amplitudes as coefficients (F_0 synthesis), and the other, the corresponding calculated amplitudes (F_c synthesis), the phase constants being the same for both syntheses.

The atomic positions derived from the F_0 synthesis are first corrected for termination of series errors by application of the back-shift rule of Booth.⁴² Since, however, the projection is non-centrosymmetrical, the coordinate shifts which are indicated, after this correction has been applied, are then increased by a factor, n , in order to obtain a more rapid refinement of the structure.^{40,43,44} This procedure leads to the equation

$$x_2 = x_1 + n(x_0 - x_c),$$

where x_2 is the refined coordinate of an atom,

x_1 is the coordinate of the atom in the previous structure,

x_0 is the coordinate of the corresponding electron density peak in the F_0 synthesis,

x_c is the coordinate of the same electron density peak in the F_c synthesis,
 n is a constant.

Provided that the shifts involved are not too large, this treatment allows for simultaneous refinement of a structure and correction for termination of series errors (see Appendix I.).

The value of the constant n may vary from two for a purely asymmetric structure with all structure factors having general phase constants, to unity for the centrosymmetrical case. In any particular case its value depends on the relative numbers of the two types of structure factors which are contained in the Fourier syntheses.

Of the 50 observed ($h0l$) structure factors, the five of type ($h00$) have restricted phases and to be on the safe side, the possibly rather low value of 1.5 was taken for n . This assumption seemed, however, to be justified by the rapid convergence of the refinement process.

The above cycle was carried out twice. In the first, 49 structure factors were employed in the evaluation of each (010) Fourier synthesis (F_{05} and F_{c1}) and at this stage the discrepancy figure was 10.5%. In the second cycle, the Fourier syntheses (F_{06} and F_{c2}) contained terms corresponding

to all the observed structure factors. The shifts indicated by these were small and the resulting coordinates were taken as the final atomic coordinates for the (010) projection. The final discrepancy for the 50 ($h0l$) structure factors was 9.1%.

The final Fourier projection on (010), F_06 , is shown in Fig.5. It gives good resolution of all the ring atoms, and three of the four hydrogen atoms appear as bulges in the one electron per \AA^2 contour line. The nitrogen atoms are marked, and it may be seen that the electron density peak values on these are higher than on the other atoms. These results seem to confirm that the nitrogen and hydrogen atoms have been located correctly. It must, however, be remembered that the positions of the nitrogen and hydrogen atoms had been assumed in the calculation of the phase constants, the majority of which have general values, and that a Fourier synthesis evaluated with such phases will tend to reproduce the structure from which they were derived.

It is of interest that the electron density peak values on the outer atoms of the molecule are lower than on those nearer the centre. This effect was also noticed in the naphthalene and anthracene structures and was suggested to be due to a thermal oscillation of the molecule about its centre.^{28,36}

(g) Investigation of the (100) Projection.

Since only nine (0kl) reflections could be recorded no attempt was made to refine this projection of the structure. Using the y and z coordinates derived from the investigations of the (001) and (010) projections, structure factors were calculated for these reflections, a discrepancy of 17.8% being obtained.

V RESULTS

The atoms of the molecule are numbered in the same way as before, ie., as in Fig. 2, Structure C, and all subsequent diagrams. The conventional numbering, given in Ring Index⁴⁵ and used in the recent literature, is shown on page 4.

(a) Coordinates and Molecular Dimensions.

The coordinates of the atoms of the molecule, referred to the unit cell axes with origin on the twofold screw axis, are given in Table 2. The coordinates of the ring atoms were obtained from the investigations of the (010) and (001) projections, the x coordinates being the mean of the values derived from the two projections. The positions of the hydrogen atoms were calculated on the assumption that they are situated radially at a distance of 1.09 Å from the carbon atoms. The centre of the molecule was calculated from the coordinates of the ring atoms.

The observed bond lengths and valency angles, derived from the atomic coordinates given in Table 2, are shown in Fig. 3, and collected in Tables 3 and 4. The bond lengths are compared with those calculated by means of wave mechanics using a molecular orbital approximation.³⁰ These calculated

bond lengths, which are based on the assumption of a strictly planar molecule, are shown in Fig.3 and Table 3.

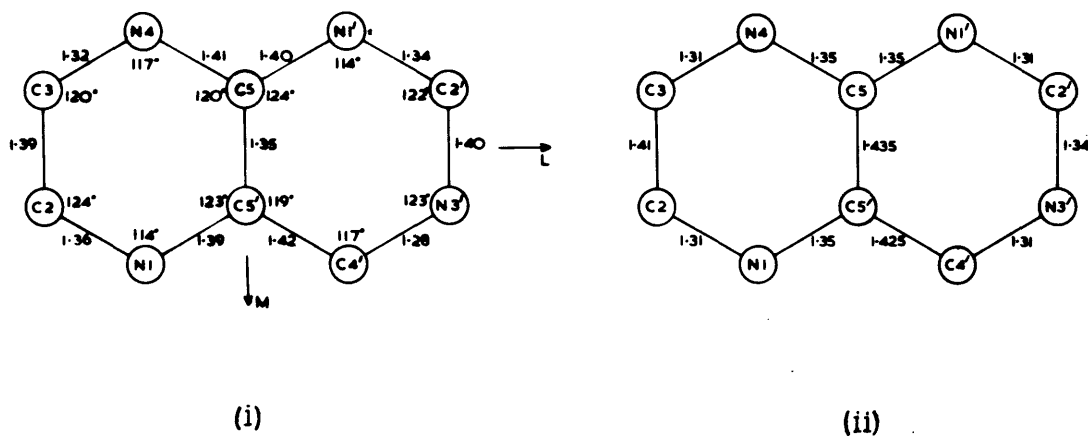


Fig. 3.

(i) Observed dimensions of the pteridine molecule.

(ii) Calculated bond lengths.

Table 2.

Final atomic coordinates.

<u>Atom</u>	x(Å)	y(Å)	z(Å)	x/a	y/b	z/c
N(1)	3.32	0.13	-1.63	0.1344	0.034	-0.258
C(2)	2.00	0.11	-1.96	0.0810	0.029	-0.310
C(3)	0.98	0.52	-1.11	0.0397	0.137	-0.176
N(4)	1.26	1.16	0.01	0.0510	0.306	0.002
C(5)	2.61	1.18	0.42	0.1057	0.311	0.066
N(1')	2.86	1.83	1.63	0.1158	0.483	0.258
C(2')	4.16	1.92	1.94	0.1684	0.507	0.307
N(3')	5.16	1.35	1.15	0.2089	0.356	0.182
C(4')	4.93	0.82	0.01	0.1996	0.216	0.002
C(5')	3.58	0.74	-0.41	0.1449	0.195	-0.065
H(2)	1.76	-0.30	-2.89	0.0713	-0.079	-0.457
H(3)	-0.03	0.44	-1.39	-0.0012	0.116	-0.220
H(2')	4.39	2.38	2.86	0.1777	0.628	0.452
H(4')	5.74	0.43	-0.62	0.2324	0.113	-0.098
Molecular centre	3.09	0.98	0.005	0.1251	0.259	0.001

Table 3.

Observed and calculated bond lengths.

<u>Bond</u>	<u>Length (obs.)</u>	<u>Length (calc.)</u>	<u>Discrepancy</u>
N(1) - C(2)	1.36 Å	1.31 Å	0.05 Å
C(2) - C(3)	1.39	1.41	0.02
C(3) - N(4)	1.32	1.31	0.01
N(4) - C(5)	1.41	1.35	0.06
C(5) - N(1')	1.40	1.35	0.05
N(1') - C(2')	1.34	1.31	0.03
C(2') - N(3')	1.40	1.34	0.06
N(3') - C(4')	1.28	1.31	0.03
C(4') - C(5')	1.42	1.425	0.005
C(5') - N(1)	1.39	1.35	0.04
C(5') - C(5)	1.35	1.435	0.085

Root mean square discrepancy = 0.047 Å

Table 4.

Observed valency angles.

$N(1) - C(2) - C(3) = 123.9^\circ$	$N(1') - C(2') - N(3') = 122.4^\circ$
$C(2) - C(3) - N(4) = 120.4^\circ$	$C(2') - N(3') - C(4') = 123.0^\circ$
$C(3) - N(4) - C(5) = 117.2^\circ$	$N(3') - C(4') - C(5') = 117.3^\circ$
$N(4) - C(5) - C(5') = 120.3^\circ$	$C(4') - C(5') - C(5) = 119.0^\circ$
$C(5') - C(5) - N(1') = 123.7^\circ$	$C(5) - C(5') - N(1) = 123.2^\circ$
$C(5) - N(1') - C(2') = 113.9^\circ$	$C(5') - N(1) - C(2) = 113.6^\circ$

(b) Geometry of the Molecule.

If the atoms of the molecule are strictly coplanar then their coordinates x , y , and z should satisfy an equation of the form

$$x + qy + rz + s = 0.$$

The coefficients q , r , and s were determined by the method of least squares from the coordinates of the ring atoms, giving as the equation of the mean plane of the molecule

$$x + 11.919y - 6.082z - 14.701 = 0.$$

The displacements of the ring atoms from this plane are given in Table 5.

Table 5.

Displacements from the mean plane.

<u>Atom</u>	<u>Displacement</u>	<u>Atom</u>	<u>Displacement</u>
N(1)	0.01 Å	N(1')	0.00 Å
C(2)	0.04	C(2')	0.04
C(3)	-0.06	N(3')	-0.03
N(4)	0.02	C(4')	0.00
C(5)	-0.04	C(5')	0.01

Root mean square displacement = 0.031 Å

(c) Orientation and Intermolecular Distances.

Molecular axes, OL, OM and ON were chosen, the origin O being the centre of the molecule. The axes OL and OM lie in the mean plane of the molecule and are directed as shown in Fig.3. The axis ON is normal to this plane. The orientation of the molecule in the unit cell is then expressed in terms of the angles which these molecular axes make with the crystallographic axes a, b and c. These angles are given in Table 6.

The intermolecular approach distances of less than 3.75 Å between the reference molecule, coordinates (x,y,z), and the neighbouring molecules derived from it by operation

of the screw axis, coordinates $(\bar{x}, \bar{y}, \frac{1}{2}+z)$, and by operation of the n glide plane, coordinates $(\frac{1}{2}-x, \frac{1}{2}+y, \frac{1}{2}+z)$, and by unit translations along b and c are given in Table 7. The closer contacts are indicated in Figs. 6 and 7, which show the general arrangement of the molecules in the crystal viewed in projection along the b and c axes.

Table 6.

Orientation of the molecule.

LOa = 46.9°,	cos LOa = 0.6838
LOb = 73.3°,	cos LOb = 0.2875
LOc = 47.9°,	cos LOc = 0.6707
MOa = 135.6°,	cos MOa = -0.7142
MOb = 68.6°,	cos MOb = 0.3644
MOc = 53.3°,	cos MOc = 0.5976
NOa = 85.7°,	cos NOa = 0.0745
NOb = 27.3°,	cos NOb = 0.8882
NOc = 117.0°,	cos NOc = -0.4532

Table 7.Intermolecular distances (\AA).Molecule (x,y,z)..... Molecule ($\bar{x},\bar{y},\frac{1}{2}+z$).

N(4) C(3) = 3.46

N(4) C(2) = 3.70

Molecule (x,y,z)..... Molecule ($\bar{x},1-y,\frac{1}{2}+z$).

N(4) C(3) = 3.69

Molecule (x,y,z)..... Molecule ($\frac{1}{2}-x,\frac{1}{2}+y,\frac{1}{2}+z$).

N(3') C(4') = 3.33

C(2') C(4') = 3.58

Molecule (x,y,z)..... Molecule (x,1+y,z).

N(4) C(3) = 3.36

N(1') C(5) = 3.37

C(2') C(4') = 3.40

C(5') N(1) = 3.42

N(3') C(4') = 3.46

C(2') N(3') = 3.46

N(4) C(2) = 3.46

N(1') C(5') = 3.46

C(5) N(1) = 3.49

C(2') C(5') = 3.56

C(3) C(2) = 3.63

C(5) C(2) = 3.66

Distance between mean molecular planes = 3.36

Molecule (x,y,z) Molecule (x,y,l+z).

N(1') C(2) = 3.34

C(2') N(1) = 3.39

N(1') N(1) = 3.53

C(2') C(2) = 3.71

Molecule (x,y,z) Molecule (x,l+y,l+z).

C(2') N(1) = 3.50

N(1') C(2) = 3.53

N(1') N(1) = 3.73

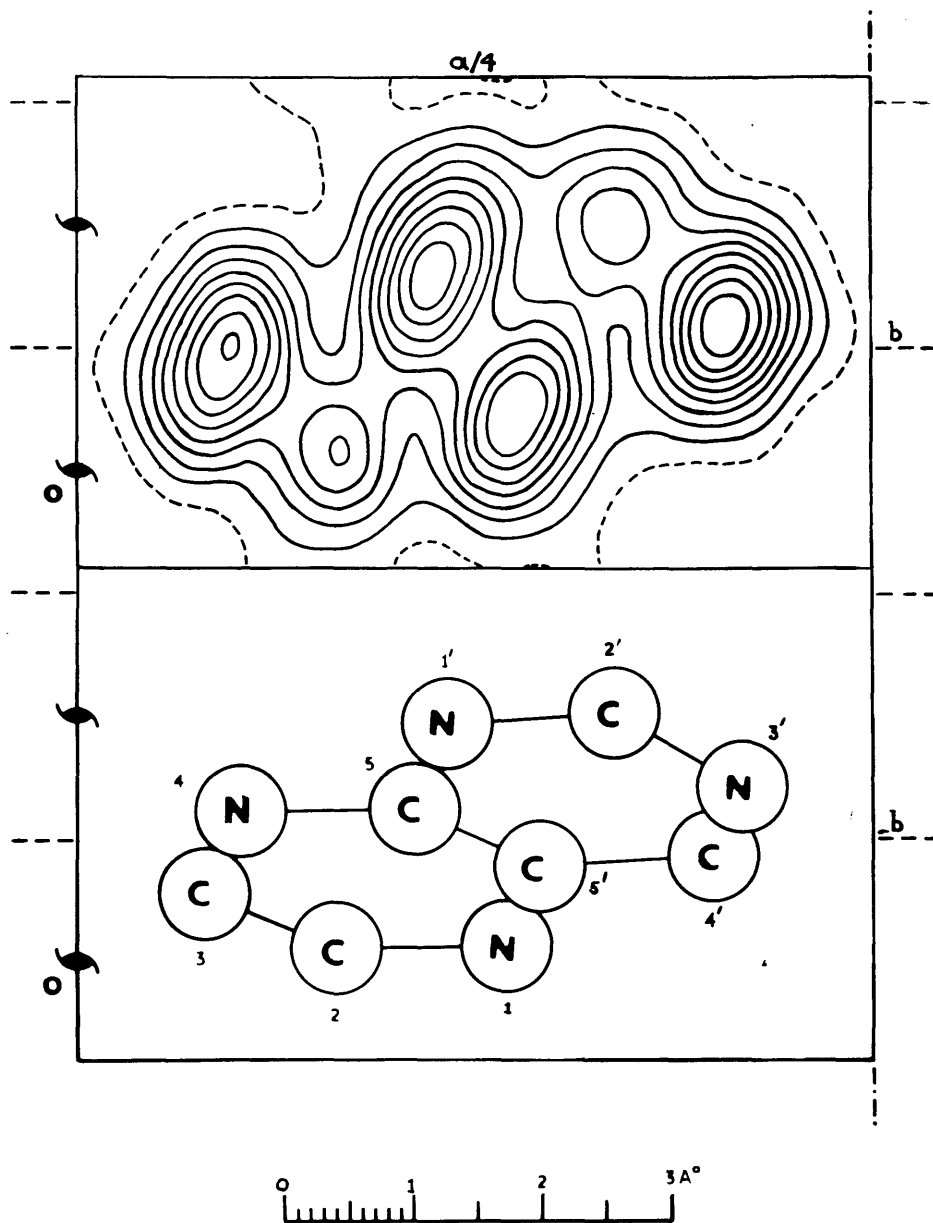


Fig. 4.

Electron density and diagrammatic projections along the c axis on (001) showing one molecule. Each contour line represents a density increment of one electron per \AA^2 , the one-electron line being dotted. (F03)

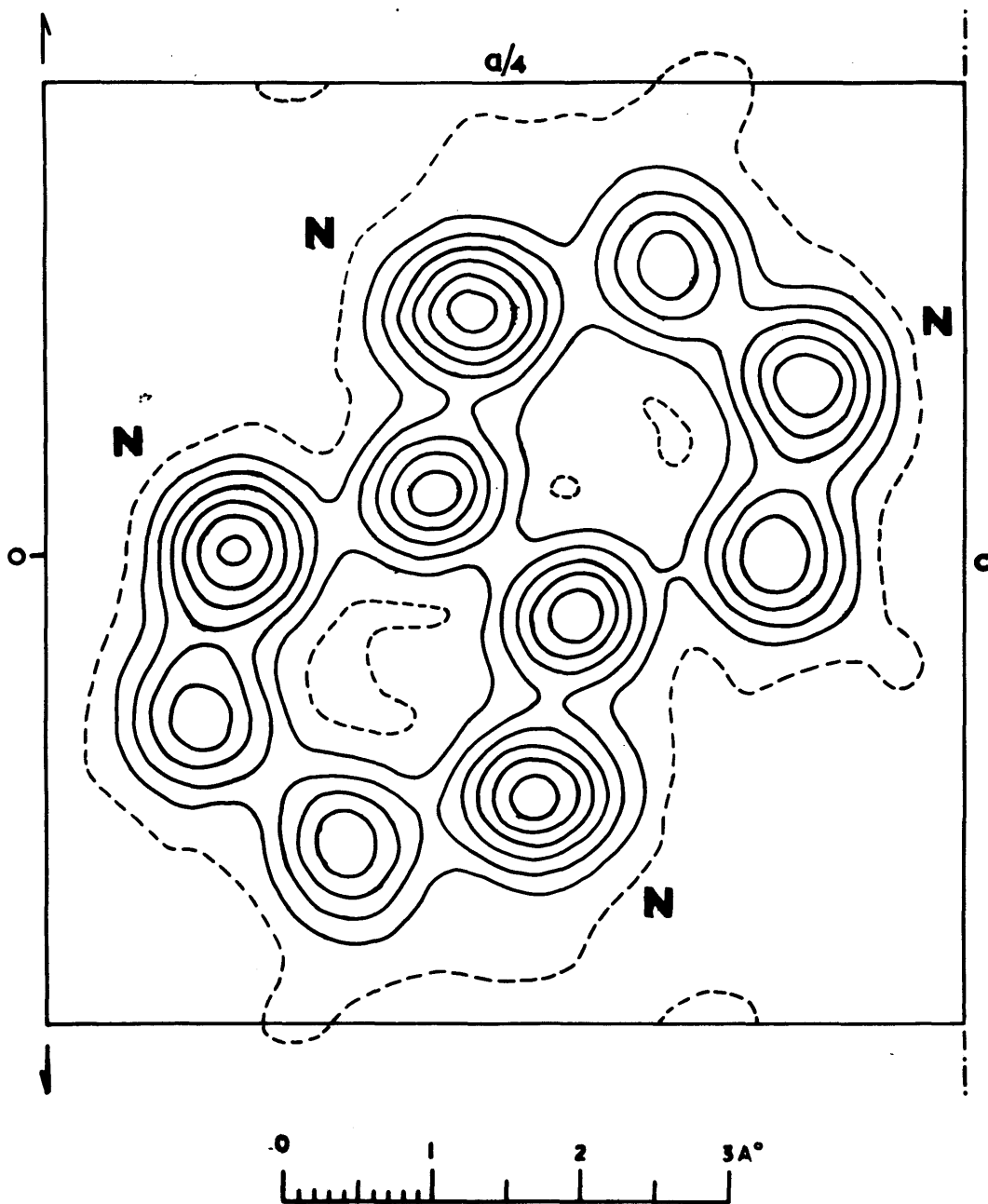


Fig. 5.

Electron density projection along the b axis on (010) showing one molecule. Contour scale as in Fig. 4. (F_o6)

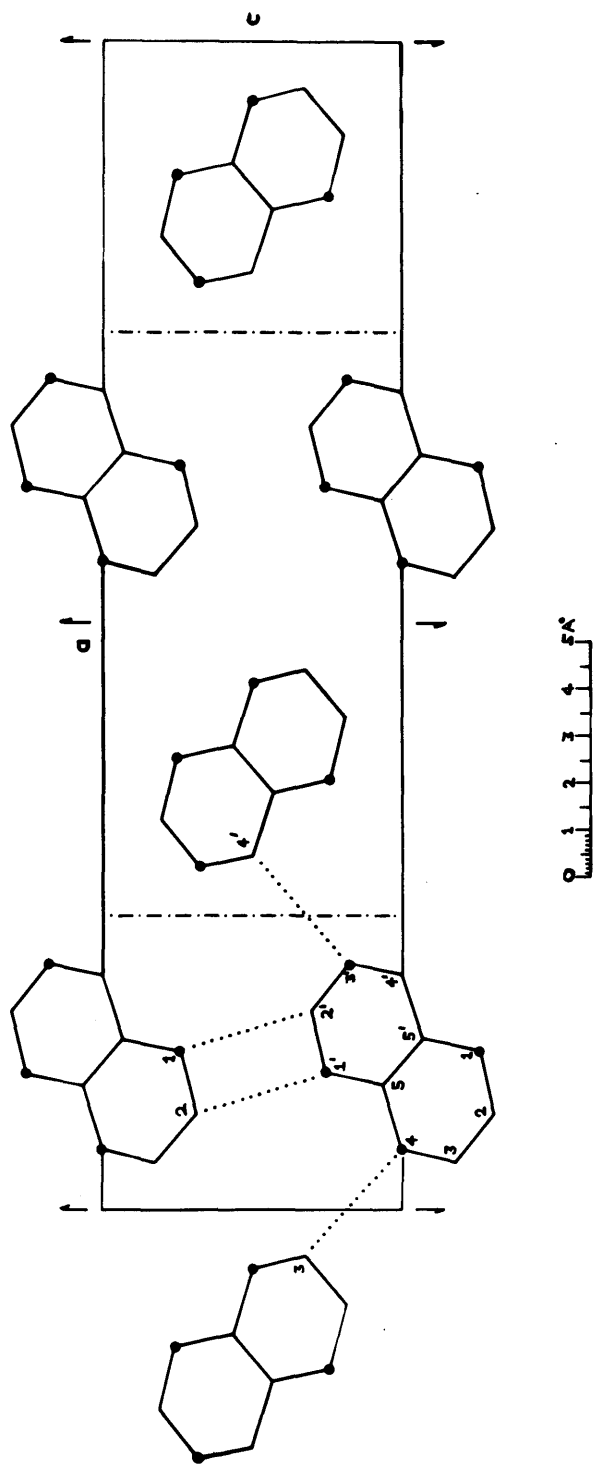


Fig. 6.

General arrangement of the molecules in the (010) projection, showing closest intermolecular contacts. The nitrogen atoms are represented as in Fig.2.

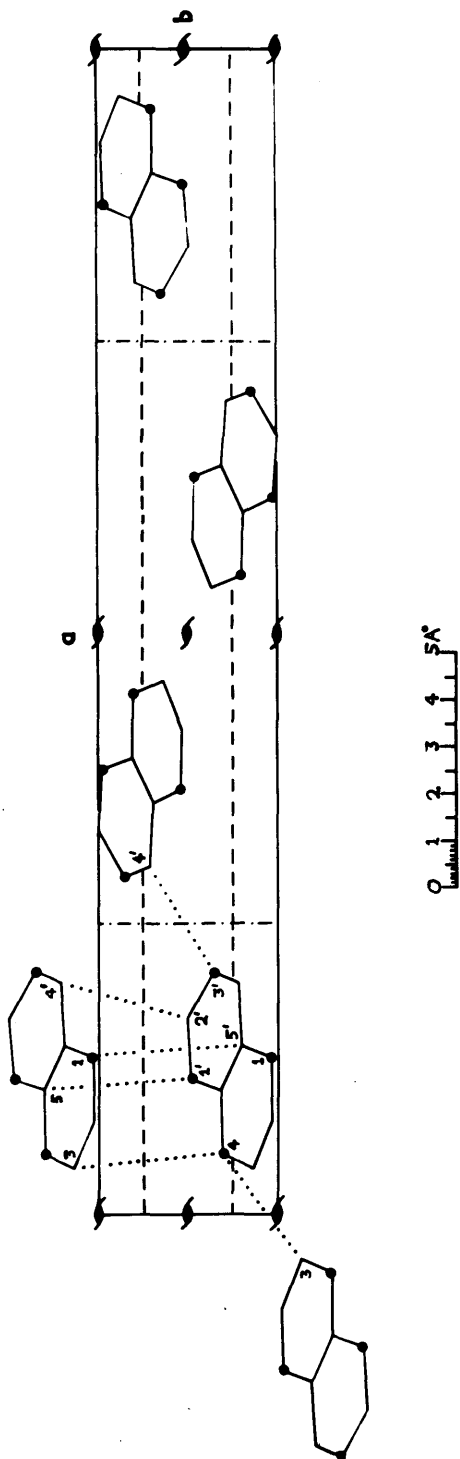


Fig. 7.

General arrangement of the molecules in the (001) projection, showing closest intermolecular contacts. The nitrogen atoms are represented as in Fig. 2.

VI ACCURACY OF THE RESULTS

The standard deviations of the atomic coordinates, $\sigma(x)$, $\sigma(y)$, and $\sigma(z)$ were estimated by the method of Cruickshank⁴⁶. According to this treatment, for an atomic coordinate derived from a well resolved projection, after correction for termination of series errors,

$$\sigma(x) = \frac{1}{A} \cdot \frac{2\pi}{a} \left[\sum h^2 (\Delta F)^2 \right]^{\frac{1}{2}} / \left| \frac{\delta^2 \rho}{\delta x^2} \right|,$$

where A is the area of the cell in projection,

ΔF is the error in the observed structure amplitudes,

$\delta^2 \rho / \delta x^2$ is the curvature of the electron density peak at the centre of the atom,

\sum denotes summation over all observed structure factors.

ΔF was taken as $||F_o| - |F_c||$, which gives an estimate of the combined experimental and residual termination of series errors due to the imperfections of the postulated model. The values which were used for $|F_o|$ and $|F_c|$ are given in Table 8.

The central curvature was determined by assuming that the electron density in the vicinity of an atomic centre can be represented by the expression⁴²

$$\rho = \rho_0 \exp[-pr^2],$$

where ρ is the electron density at a distance r from the centre of the atom,

ρ_0 is the density at the centre of the atom,
 p is a constant.

It follows that at the centre of the peak,

$$\delta^2 \rho / \delta x^2 = -2p\rho_0.$$

The final electron density projection on (010), Fo6, was used to obtain graphs of $\ln \rho$ against r^2 for the various atomic peaks. From these, the average values of p and ρ_0 were derived, giving $-39.7e./\text{\AA}^4$ and $-47.2e./\text{\AA}^4$ as the central curvatures for carbon and nitrogen respectively.

The standard deviations obtained in this way were increased by a factor $[u/(u-v)]^{\frac{1}{2}}$ to allow for the rather small number, u , of independent observed structure factors compared with the number, v , of unknown parameters.⁴⁷

Also, it has been shown by Cruickshank⁴⁸ that for a non-centrosymmetrical projection, the standard deviations given by the formula must be increased by a factor of two to allow for errors in the phase constants. More generally, a factor, n , of value intermediate between one and two is applicable,⁴⁰ n being the same constant which was used in

the refinement process (see page 40).

In the refinement of the (010) projection n had been taken as 1.5, but this value may have been rather low since the majority of the $(h0l)$ structure factors have general phase constants. Accordingly, to be on the safe side, the values of $\sigma(x)$ and $\sigma(z)$ obtained from the (010) projection data were increased by a factor of two.

The y coordinates were derived from the largely unresolved (001) projection for which Cruickshank's treatment is not strictly applicable, and the values of $\sigma(y)$ which were obtained, were increased by a factor of 1.25. There is, however, no theoretical basis for this procedure and it is possible that the effect of the lack of resolution on the accuracy of the y coordinates may have been underestimated.

The coordinate standard deviations finally obtained are

$$\begin{array}{lll} \text{Carbon:} & \sigma(x) = 0.022\text{\AA}; & \sigma(y) = 0.03\text{\AA}; & \sigma(z) = 0.023\text{\AA}. \\ \text{Nitrogen:} & \sigma(x) = 0.018\text{\AA}; & \sigma(y) = 0.025\text{\AA}; & \sigma(z) = 0.019\text{\AA}. \end{array}$$

The standard deviation of a carbon-carbon bond is then 0.035 Å and of a carbon-nitrogen bond 0.033 Å. The standard deviation of a valency angle, calculated by the formula

given by Cruickshank and Robertson⁴⁹ is about 2.3° , and the electron density standard deviations are $0.42\text{e}/\text{\AA}^2$ for the (010) projection and $0.27\text{e}/\text{\AA}^2$ for the (001) projection.

VII DISCUSSION

The errors in the observed results having been estimated, and since the probability distribution for these follows approximately the Gaussian error law,⁴⁶ the observed results may be compared with the theoretical or postulated values by means of statistical significance tests. According to the treatment of Cruickshank and Robertson,⁴⁹ the significance of a difference, Δ , between an observed value with estimated standard deviation σ and a theoretical value, which is assumed to be the true value with zero standard deviation, is estimated in terms of the probability, P , that such a difference, or one greater, could be due to random errors alone. It is assumed that if

$P \geq 0.05,$	Δ is not significant.
$0.05 > P > 0.01,$	Δ is possibly significant.
$0.01 \geq P,$	Δ is significant.

When the number of degrees of freedom, $(u-v)$, is large, σ may be considered as an accurate estimate of the standard deviation and the values of $t = \Delta/\sigma$ corresponding to the above values of P are

$P = 0.05,$	$t = 1.960,$
$P = 0.01,$	$t = 2.576.$

If, however, the number of degrees of freedom is small (< 30), σ is an uncertain estimate of the standard deviation, and the values of t for the different significance levels are greater.⁴⁷

The bond lengths observed in pteridine together with those calculated by the method of molecular orbitals³⁰ are shown in Fig.3 and Table 3. The largest discrepancy between the observed and theoretical values, 0.085 \AA , occurs in the case of the central bond, C(5)-C(5'), where the observed length is too small. Here $t = 2.43$, and this difference is therefore possibly significant. For all the other bonds the discrepancies are not significant. The root mean square discrepancy, 0.047 \AA , is also within the limits of experimental error. Thus, owing to the comparatively large standard deviations in the observed lengths, the quite considerable differences between the observed and calculated bond lengths cannot be regarded as real.

It is, however, of interest that in the case of the central bond and in six of the other ten bonds, the observed values deviate from the calculated values in the direction of the symmetrical Kekule structure. Similar though smaller discrepancies are found also in the case of naphthalene where much more accurate bond lengths are available.^{28,29} Here, the largest deviation also occurs in the central bond,

where the observed length is significantly less than the calculated value.^{29,50} Although there does not seem to be any theoretical basis for this, it would appear that in pteridine, as has been suggested for naphthalene,⁵⁰ the symmetrical Kekule structure should, in some way, be given greater weight in the calculations.

The observed length of 1.35 Å for the central bond in pteridine is, however, only 0.01 Å greater than the accepted value for a carbon-carbon pure double bond. It seems unlikely that such a bond should occur in the molecule and at least part of the discrepancy between the observed and calculated lengths is, in this case, obviously due to an error in the observed value.

The shortest observed bond length occurs in the pyrimidine ring, N(3')-C(4') = 1.28 Å. This would correspond very nearly to a carbon-nitrogen pure double bond which has, however, never been measured directly, but whose length has been estimated by different workers as 1.24 Å,⁵¹ 1.265 Å,⁵² 1.27 Å,⁵³ or 1.28 Å.^{22,54} It does seem therefore, that the observed length is rather too low also in this case. A very short bond in the pyrimidine ring would, however, account for the formation of 2-aminopyrazine-3-aldehyde when pteridine is treated with dilute acid,⁷ since ring fissure would tend to take place at such a bond. This would also account for the

general instability of pteridine.^{5,7}

The observed valency angles are shown in Table 4 and Fig.3. Three of the four nitrogen valency angles are less than 120° , for two of these, $C(5)-N(1')-C(2') = 113.9^\circ$ and $C(5')-N(1)-C(2) = 113.6^\circ$, the difference being significant. The fourth nitrogen valency angle, which is 123.0° , is not significantly greater than 120° . The angles at the carbon atoms are generally greater than 120° but in no case is the difference significant. The bond angles observed in pteridine thus show variations similar to those found in other nitrogen containing aromatic ring systems.

The displacements of the ring atoms from the calculated mean molecular plane are given in Table 5. These are measured almost entirely by the y coordinates and the appropriate coordinate standard deviation is given by $\sigma(y)$. The y coordinates were determined from the (001) projection on the basis of only 49 independent structure factors. The number of degrees of freedom is therefore 29, and the values of t for the different levels of significance are

$$P = 0.05, \quad t = 2.045,$$

$$P = 0.01, \quad t = 2.756.$$

The maximum displacement of a carbon atom from the mean plane is 0.06 \AA , for which $t = 0.06/0.03 = 2.00$. This displacement

cannot therefore be regarded as even possibly significant. The maximum displacement of a nitrogen atom is 0.03 \AA , for which $t = 0.03/0.025 = 1.20$ and this is not significant. The root mean square displacement is also not significant.

It may therefore be concluded that within the limits of the experimental errors, the pteridine molecule is planar. It cannot, however, be claimed that the question of its planarity has been settled, since, owing to the comparatively large errors in the experimental measurements, deviations of up to 0.06 \AA from a strictly planar structure would be difficult to detect. A further investigation of pteridine using three-dimensional data would obviously be desirable. In connection with this, it is, however, of interest that in the α -phenazine structure,²⁵ which was shown to be planar to within 0.014 \AA in a three-dimensional investigation, deviations of up to 0.06 \AA from a planar structure were found on the basis of two-dimensional data.

The arrangement of the molecules in the crystal and the shorter intermolecular distances are shown in Figs.6 and 7. As in the case of α -phenazine, the crystal structure appears to consist of stacks of parallel molecules separated by unit translations along the b axis, the perpendicular distance between the mean molecular planes being 3.36 \AA . Some of the atoms in these molecules overlie each other closely,

and many of the closest contacts (minimum 3.36 Å) occur between such molecules.

The shortest intermolecular approach distance, 3.33 Å, is between the reference molecule and the one derived from it by operation of the n glide plane. Another very close contact of 3.34 Å occurs between molecules separated by a translation c . The shortest distance of approach between atoms in molecules related by the twofold screw axis is 3.46 Å.

It may be noted that all the intermolecular contacts of less than 3.40 Å involve a carbon atom of one molecule and a nitrogen atom of another, the shortest non-bonded carbon.....carbon distance being 3.40 Å. None of the intermolecular approach distances are therefore less than the sum of the van der Waals radii of the atoms concerned. The interplanar distance of 3.36 Å is similar to that found in graphite (3.35-3.40 Å), benzperylene⁵⁵ (3.38 Å) and coronene⁵⁶ (3.40 Å), but is rather less than the corresponding distances in ovalene⁵⁷, acridine²⁷ and α -phenazine²⁵ which are 3.45, 3.47 and 3.49 Å respectively.

Comparison with the naphthalene structure,^{28,58} shows that the packing of the molecules in pteridine is much closer. The volume occupied by a pteridine molecule is 148 Å³ compared to 181 Å³ for naphthalene, the corresponding

crystal densities being 1.48 and 1.15 gm./cm³. The minimum intermolecular distance of 3.60 Å in naphthalene may be compared with the many considerably shorter contacts in the pteridine structure.

The more efficient packing found in pteridine may be partly due to attractive forces between the carbon and nitrogen atoms of different molecules as shown by the large number of short carbon.....nitrogen intermolecular distances. However, the fact that the pteridine molecule contains four hydrogen atoms fewer than the naphthalene molecule must also be of importance in this respect. The 'parallel disc' type of packing⁵⁹ shown by pteridine is, however, basically different from the packing found in naphthalene, and no direct comparison can be made.

VIII FINAL STRUCTURE FACTORS

The final structure factors, calculated from the atomic coordinates given in Table 2, are shown in Table 8 under $|F_c|$ and α together with the observed structure amplitudes, $|F_o|$, which have been scaled by comparison with the calculated values. In the calculation of these structure factors, the theoretical scattering curves of McWeeny³⁷ for carbon, nitrogen and hydrogen were employed, the value of the constant B in the temperature factor being, as before, 4.8 \AA^2 . The final discrepancy, expressed as $\sum||F_o| - |F_c|| / \sum|F_o|$, was 11.1% for the 101 observed (0kl), (h0l) and (hk0) structure factors. The discrepancies for the separate zones were 17.8%, 9.1%, and 12.2% respectively.

Table 8.

Observed and calculated structure amplitudes and
phase constants.

<u>hkl</u>	<u>2sinθ</u>	<u> F_o </u>	<u> F_c </u>	<u>δ°</u>
000	-	-	272.0	-
200	0.125	< 0.6	0.2	180
400	0.249	73.5	63.1	180
600	0.374	< 1.1	0.0	-
800	0.499	11.7	11.2	180
10,0,0	0.624	< 1.6	0.3	0
12,0,0	0.748	41.0	37.7	180
14,0,0	0.873	< 2.0	3.1	0
16,0,0	0.998	< 2.3	1.1	180
18,0,0	1.122	< 2.5	1.1	0
20,0,0	1.247	9.0	7.9	0
22,0,0	1.372	< 3.8	0.9	180
24,0,0	1.496	4.9	6.2	0
020	0.813	< 1.8	5.2	0
040	1.626	< 2.3	4.2	0
002	0.487	27.2	25.1	180
004	0.975	29.4	26.2	357
006	1.462	< 3.8	3.4	14
011	0.474	64.2	75.4	87
013	0.836	13.0	10.5	93

Table 8. (contd.)

<u>hkl</u>	<u>2sinθ</u>	<u> F_o </u>	<u> F_c </u>	<u>α°</u>
015	1.285	8.7	7.7	91
022	0.948	36.4	40.8	181
024	1.269	4.8	3.1	13
031	1.243	5.9	7.1	296
033	1.421	9.0	5.9	256
201	0.277	44.9	53.9	93
401	0.353	41.5	45.7	2
601	0.449	32.5	28.9	96
801	0.558	27.9	27.7	1
10,0,1	0.671	21.5	20.8	92
12,0,1	0.787	24.1	24.1	186
14,0,1	0.908	19.2	18.3	277
16,0,1	1.027	14.5	15.7	355
18,0,1	1.153	16.6	14.7	276
20,0,1	1.271	18.1	20.3	179
22,0,1	1.395	9.3	8.7	90
202	0.504	4.4	6.4	307
402	0.548	20.9	23.6	356
602	0.616	27.7	28.8	263
802	0.698	24.0	23.8	188
10,0,2	0.793	27.2	27.9	102
12,0,2	0.895	< 2.9	1.9	286
14,0,2	1.003	6.6	7.2	104

Table 8. (contd.)

<u>hkl</u>	<u>2sinθ</u>	<u> F_o </u>	<u> F_c </u>	<u>α</u>
16,0,2	1.111	<3.4	0.8	21
18,0,2	1.227	<3.6	1.1	326
20,0,2	1.341	13.4	13.3	2
22,0,2	1.458	3.7	3.5	282
24,0,2	1.577	7.3	7.9	179
203	0.742	24.3	25.1	87
403	0.775	15.4	13.8	157
603	0.822	6.7	6.9	285
803	0.887	4.7	4.7	38
10,0,3	0.963	14.4	14.3	273
204	0.985	10.2	9.9	82
404	1.008	11.8	11.0	183
604	1.046	<3.3	2.7	84
804	1.098	<3.4	1.7	40
10,0,4	1.159	<3.5	1.9	242
12,0,4	1.230	9.4	10.3	175
14,0,4	1.313	16.6	16.7	272
16,0,4	1.396	6.7	6.8	3
18,0,4	1.492	7.9	8.0	86
205	1.227	5.8	3.5	262
405	1.246	10.9	10.3	181
605	1.278	14.3	16.3	92
805	1.319	6.0	5.4	356

Table 8. (contd.)

<u>hkl</u>	<u>2sinθ</u>	<u> F_o </u>	<u> F_c </u>	<u>α</u>
10,0,5	1.374	< 3.9	4.9	90
12,0,5	1.433	9.9	8.2	181
14,0,5	1.504	7.6	8.6	271
16,0,5	1.577	7.8	7.7	5
18,0,5	1.662	3.5	4.6	270
206	1.468	< 3.8	0.7	45
406	1.485	11.4	10.5	3
606	1.511	6.3	5.0	268
806	1.547	12.9	12.3	183
10,0,6	1.593	4.4	3.8	93
110	0.410	< 3.0	8.4	0
210	0.424	69.9	73.8	180
310	0.446	23.3	22.9	0
410	0.477	1.2	2.0	180
510	0.512	11.8	12.2	180
610	0.550	28.9	28.3	180
710	0.597	1.4	5.6	0
810	0.644	1.8	3.2	0
910	0.689	12.7	9.2	180
10,1,0	0.740	20.1	18.3	180
11,1,0	0.791	10.3	11.0	180
12,1,0	0.852	< 1.8	2.8	0
13,1,0	0.901	15.0	15.1	0

Table 8. (contd.)

<u>hkl</u>	<u>2sinθ</u>	<u> F_o </u>	<u> F_c </u>	<u>δ</u>
14,1,0	0.958	18.1	16.2	0
15,1,0	1.022	9.6	11.8	0
16,1,0	1.078	<2.1	1.7	180
17,1,0	1.127	15.3	14.2	180
18,1,0	1.186	16.2	14.8	0
19,1,0	1.244	14.3	15.0	180
20,1,0	1.310	<2.5	0.1	0
21,1,0	1.360	5.7	8.7	0
22,1,0	1.420	4.8	4.7	180
23,1,0	1.492	3.0	3.6	0
30,1,0	1.916	2.1	2.7	180
120	0.814	<1.8	1.4	0
220	0.820	<1.8	0.6	0
320	0.830	6.1	3.9	180
420	0.847	19.0	21.0	180
520	0.867	8.4	7.7	0
620	0.892	3.0	0.1	0
720	0.919	7.1	6.9	180
820	0.950	12.7	12.7	0
920	0.984	11.3	11.5	180
10,2,0	1.021	4.1	0.0	-
11,2,0	1.060	7.5	7.4	0
12,2,0	1.100	9.0	10.2	0

Table 8. (contd.)

<u>hkl</u>	<u>$2\sin\theta$</u>	<u>F_o</u>	<u>F_c</u>	<u>α°</u>
13,2,0	1.154	< 2.3	1.8	180
14,2,0	1.200	< 2.3	0.3	180
15,2,0	1.240	2.7	2.9	180
16,2,0	1.288	2.8	3.3	0
17,2,0	1.329	6.1	5.4	180
18,2,0	1.380	< 2.5	0.3	0
19,2,0	1.438	4.1	5.2	0
20,2,0	1.481	9.8	12.0	180
21,2,0	1.531	7.3	6.2	0
22,2,0	1.600	< 2.3	0.0	-
23,2,0	1.654	< 2.3	0.0	-
24,2,0	1.705	3.2	4.4	0
130	1.223	< 2.3	2.1	0
230	1.227	< 2.3	3.4	0
330	1.235	< 2.3	3.5	0
430	1.246	3.0	2.4	180
530	1.260	3.5	4.1	180
630	1.277	3.0	2.1	0
730	1.297	6.0	8.9	180
830	1.319	3.9	1.0	0
930	1.344	3.2	3.4	0
22,3,0	1.837	2.5	3.3	0

INVESTIGATION OF THE STRUCTURE OF

' SUBSTANCE P '

I INTRODUCTION

It was found that when 6-hydroxypteridine is dissolved in sodium hydroxide solution a small quantity of a dipteridyl ¹¹ derivative is formed. This was named 'substance P' by Albert who showed that it is formed in 25% yield when 7:8-dihydro-6-hydroxypteridine and 6-hydroxypteridine are allowed to react in dilute sodium hydroxide solution. Under these conditions, neither of the starting materials will give rise to 'substance P' alone, and it has been suggested that the NH group of 7:8-dihydro-6-hydroxypteridine has added across the 7:8-double bond of 6-hydroxypteridine. According to this mechanism, 'substance P' is 7:8:7':8'-tetrahydro-6:6'-dihydroxy-7:8'-dipteridyl. On the basis of spectroscopic evidence the corresponding 8:8'-dipteridyl which is a symmetrical hydrazine was considered to be less likely. Other possible structures for 'substance P' are the 6:6'- and the 7:7'-dipteridyl which are also symmetrical.⁶⁰

The present investigation was undertaken at the suggestion of Professor Albert in order to determine the structure of 'substance P'. The free base is microcrystalline and crystals of the dihydrochloride dihydrate obtained by crystallisation from hydrochloric acid solution were employed.

II CRYSTAL DATA

'Substance P' dihydrochloride dihydrate, $C_{12}H_{10}O_2N_8 \cdot 2HCl \cdot 2H_2O$.

Molecular weight, 407,2.

Decomposes below melting point at about 280°C.

Density calculated 1.620gm./cm.³ Density found, 1.597gm./cm.³

The crystal is monoclinic with

$$a = 12.12 \pm 0.04 \text{ \AA},$$

$$b = 11.21 \pm 0.02 \text{ \AA},$$

$$c = 13.46 \pm 0.03 \text{ \AA},$$

$$\beta = 114.2^\circ \pm 0.5^\circ$$

Volume of the unit cell, 1668.1 Å³

Absent spectra, (hkl) when $h + k + l$ is odd.

(hol) when h or l is odd.

(Ok0) when k is odd.

Space group $Ia(C_S^4)$ or $I2/a(C_{2h}^6)$.

Polarity tests indicate Ia.

Four molecules per unit cell.

No molecular symmetry required.

Absorption coefficient for X-rays, $CuK\alpha$ radiation, $\mu = 39.4 \text{ cm}^{-1}$

Total number of electrons per unit cell = $F(000) = 840$.

III EXPERIMENTAL

The sample of the dihydrochloride of 'substance P' used in this investigation was obtained from Professor A. Albert¹¹. The large, but irregular shaped crystals are stable and unlike pteridine are not volatile at room temperature. Nickel filtered copper $K\alpha$ radiation ($\lambda=1.542 \text{ \AA}$) was employed in the determination of the crystal data. Rotation, oscillation and equatorial layer line moving film photographs about each of the unit cell axes and a first layer line moving film about b were taken. The axial lengths were measured from the rotation photographs and the β angle was estimated approximately from the moving film taken about the b axis.

The systematic absences, determined from the moving film photographs correspond to the two space groups $Ia(C_s^4)$ or $I2/a(C_{2h}^6)$. Application of pyroelectric tests⁶⁶, showed that the crystals are polar, thus excluding the centrosymmetrical space group.

The density of the crystals was determined by flotation in a mixture of carbon tetrachloride and benzene containing a trace of detergent.

APPENDICES

APPENDIX IThe Refinement of Non-Centrosymmetrical Structures by
Fourier Synthesis.

Since a crystal structure is periodic, it may be represented in terms of the electron density distribution by a Fourier series. In the usual notation,

$$\rho(xyz) = 1/v \sum_{\underline{h}} \sum_{\underline{k}} \sum_{\underline{l}} |F(hkl)| \cos [2\pi hx + 2\pi ky + 2\pi lz - \alpha(hkl)].$$

For centrosymmetrical structures, the possible values of the phase constant are restricted to 0° or 180° . If these values can be correctly determined, the resulting Fourier synthesis will give an exact representation of the structure, apart from errors due to the inaccuracies in the observed amplitudes and the finite termination of the series.

In the non-centrosymmetrical case, however, the phase constants have unrestricted values, and even slight errors in the atomic positions for which the structure factors are calculated will lead to errors in the phases as well as in the amplitudes. The Fourier synthesis based on these phases will therefore be in error and will tend to reproduce the structure from which the phases were derived. The refinement process is thus slow and in theory never reaches completion.

It has been shown by Cruickshank⁴³ that when the shifts indicated by such a Fourier synthesis are less than 0.10 \AA , they should be doubled in order to compensate for this effect and thus obtain more rapid refinement. More generally, if the structure does not have a centre of symmetry, but its projection on certain planes or lines is centred, a factor n is applicable.⁴⁰ The value of n in any given case depends on the relative numbers of real and complex structure factors which are being used, and may vary from two for the purely asymmetric case to unity for a centrosymmetrical structure. A proof of the validity of this 'n-shift' rule has been given by Cruickshank.⁴⁴

Owing to the experimental conditions, the Fourier series will, in general, be terminated when its coefficients are still appreciable and the theoretical requirement of a summation over an infinite number of terms is, therefore, not even approximately fulfilled. Systematic errors then occur in the electron density distribution as given by the synthesis, each electron density peak being displaced from its true position by diffraction ripples from the other peaks.

The following method of correcting atomic coordinates for termination of series errors, known as the 'back-shift' method, has been suggested by Booth^{42,63} and verified theoretically by Cochran.⁶⁴ 'Having calculated the structure

factors from the final atomic coordinates, a synthesis is computed using these calculated values as coefficients. Any terms not included in the original synthesis with experimental coefficients are similarly omitted in this new synthesis. The coordinates derived will, in general, deviate slightly from those used in calculating the F values; these deviations give the errors with reversed signs of the original coordinates and may if desired be applied as corrections.'

This simple procedure is generally applied, as such, only to centrosymmetrical structures. Its extension for simultaneously correcting for termination of series errors and refining structures with no centre of symmetry has been described by Donohue⁴¹.

Let the coordinates of an atom in a nearly correct trial structure be (x_1, y_1, z_1) . On the basis of this structure, structure factors are calculated and the resulting phase constants used to evaluate two Fourier syntheses, one containing the observed amplitudes as coefficients (F_o synthesis), and the other, the calculated amplitudes (F_c synthesis). If the coordinates of the corresponding electron density peak in the F_c synthesis are (x_c, y_c, z_c) , the error due to termination of series is given (for the x coordinate) by

$$x_c - x_1.$$

This is applied with reversed sign to the corresponding peak in the F_0 synthesis (coordinates x_0, y_0, z_0), giving

$$x_0 - (x_c - x_1)$$

as the coordinate of the peak, corrected for the effect of termination of series.

The shift indicated by the F_0 synthesis, after it has been corrected in this way by the back-shift method, is therefore

$$x_0 - (x_c - x_1) - x_1.$$

Application of the n -shift rule gives the required shift as

$$n[x_0 - (x_c - x_1) - x_1].$$

The refined coordinate x_2 is then given by

$$\begin{aligned} x_2 &= x_1 + n[x_0 - (x_c - x_1) - x_1] \\ &= x_1 + n(x_0 - x_c) \end{aligned}$$

and similarly for the y and z coordinates.

These coordinates are, however, subject to errors due to the approximations in the n -shift rule, and a single cycle will, in general, not give the true atomic positions. The method may be considered as providing a means of refining non-centrosymmetrical structures more rapidly.

APPENDIX IIThe Treatment of Accidentally Absent Reflections in
Fourier Refinements.

Although the possibility of including accidentally absent reflections in 'error syntheses' had been pointed out by Crowfoot et al,⁶⁵ in most structure analyses, absent reflections are not considered during the refinement process^{61,62} and structure factors corresponding to these are calculated only for the final atomic coordinates. The structure amplitudes of the unobserved reflections are estimated on the basis of the minimum value that it has been found possible to observe at the corresponding angle. This amplitude, $|F_o^{\max}|$, then measures the maximum possible value of the structure amplitude for the absent reflection. The general agreement between these amplitudes and the calculated values, which in this case implies that $|F_c|$ should be less than $|F_o^{\max}|$, is taken to demonstrate the validity of the structure, which has been derived only from the data actually observed. When, however, the number of observed reflections is small owing to the experimental conditions, such as the nature of the crystal, it would appear to be desirable to extend the available information by making use of vanishingly small reflections, at least in the final stages of the refinement process.

In the case of refinement by difference syntheses, the criterion for the use of absent reflections would be that the corresponding values taken for $(F_o - F_c)$ should be nearer to the true values than is obtained by omitting these terms, which implies that $|F_o| - |F_c| = 0$. If $|F_c|$ is considerably greater than $|F_o^{\max}|$, say by a factor of two, then it may be assumed that F_c defines the phase constant for the reflection with a considerable degree of certainty. It follows that by taking the observed amplitude of the vanishingly small reflection as $\frac{1}{2}|F_o^{\max}|$, the maximum possible error in the value of $(F_o - F_c)$ is equal to $\frac{1}{2}|F_o^{\max}|$. In general this would be small and it can be shown that if $|F_c| > 3/2|F_o^{\max}|$, this procedure will lead to a better value for $(F_o - F_c)$ than is obtained by making it equal to zero.

It thus appears, that if the calculated value of the structure amplitude for an unobserved reflection is more than twice as great as the estimated maximum value, then a term for such a reflection may be included in the difference synthesis as $(\frac{1}{2}F_o^{\max} - F_c)$.

The corresponding procedure in the ordinary Fourier method, using separate F_o and F_c syntheses, would be to include in the F_o synthesis the values $\frac{1}{2}F_o^{\max}$ for all those unobserved reflections where $|F_c| > 2|F_o^{\max}|$. The calculated structure factors corresponding to these would be included in the F_c synthesis.

REFERENCES

1. F.G. Hopkins, Nature, 1889, 40, 335.
2. R. Purrmann, Annalen, 1940, 546, 98.
3. R. Purrmann, Annalen, 1940, 544, 182.
4. W.G.M. Jones, Nature, 1948, 162, 524.
5. A. Albert, D.J. Brown and G. Cheeseman, J. Chem. Soc., 1951, p. 474.
6. A. Albert, D.J. Brown and G. Cheeseman, J. Chem. Soc., 1952, p. 1620.
7. A. Albert, D.J. Brown and G. Cheeseman, J. Chem. Soc., 1952, p. 4219.
8. A. Albert, and D.J. Brown, J. Chem. Soc., 1953, p. 74.
9. A. Albert, D.J. Brown and H.C.S. Wood, J. Chem. Soc., 1954. p. 3832.
10. S.F. Mason, J. Chem. Soc., 1955, p. 2336.
11. A. Albert, J. Chem. Soc., 1955, p. 2690.
12. A. Albert, Quart. Rev., 1952, 6, 197.
13. A. Albert, Special Publication Chem. Soc., No.3, 1955, p. 124.
14. C.J.B. Clews and W. Cochran, Acta Cryst., 1949, 2, 46.
15. C.J.B. Clews and W. Cochran, Acta Cryst., 1948, 1, 4.
16. G.J. Pitt, Acta Cryst., 1948, 1, 168.
17. D.T. Cromer, A.J. Ihde and H.L. Ritter, J. Amer. Chem. Soc., 1951, 73, 5587.

18. V. Schomaker and L. Pauling, J. Amer. Soc., 1939, 61, 1769.
19. J.M. Broomhead, Acta Cryst., 1948, 1, 324.
20. W. Cochran, Acta Cryst., 1951, 4, 81.
21. J.M. Broomhead, Acta Cryst., 1951, 4, 92.
22. P.J. Wheatley, Acta Cryst., 1955, 8, 224.
23. E.W. Hughes, J. Amer. Chem. Soc., 1941, 63, 1737.
24. E. Wiebenga, J. Amer. Chem. Soc., 1952, 74, 6156.
25. F.H. Herbststein and G.M.J. Schmidt, Acta Cryst., 1955, 8, 399, 406.
26. B.B. De More, W.S. Wilcox, J.H. Goldstein, J. Chem. Phys., 1954, 22, 876; B. Bak, L. Hansen and J. Rastrup-Andersen, J. Chem. Phys., 1954, 22, 2013.
27. D.C. Phillips, Acta Cryst., 1956, 9, 237.
28. S.C. Abrahams, J.M. Robertson and J.G. White, Acta Cryst., 1949, 2, 233, 238.
29. F.R. Ahmed and D.W.J. Cruickshank, Acta Cryst., 1952, 5, 852.
30. T.H. Goodwin and A. Porte, J.Chem. Soc., 1956, (forthcoming paper)
31. J.M. Robertson, J. Sci. Instr., 1943, 20, 175.
32. G. Albrecht, Rev. Sci. Instr., 1939, 10, 221.
33. A.J.C. Wilson, Acta Cryst., 1949, 2, 318.

34. E.R. Howells, D.C. Phillips and D. Rogers, *Acta Cryst.*, 1950, 3, 210.
35. International Tables for X-ray Crystallography, Birmingham: Kynoch Press, 1952, pp. 119, 392.
36. V.C. Sinclair, J.M. Robertson and A.M. Mathieson, *Acta Cryst.*, 1950, 3, 245, 251.
37. R. McWeeny, *Acta Cryst.*, 1951, 4, 513.
38. W. Cochran, *Acta Cryst.*, 1951, 4, 408.
39. A.D. Booth, *Fourier Technique in X-ray Organic Structure Analysis*, Cambridge: University Press, 1948, p. 62.
40. D.P. Shoemaker, J. Donohue, V. Schomaker and R.B. Corey, *J. Amer. Chem. Soc.*, 1950, 72, 2328.
41. J. Donohue, *J. Amer. Chem. Soc.*, 1950, 72, 949.
42. A.D. Booth, *Proc. Roy. Soc.*, 1946, A188, 77.
43. D.W.J. Cruickshank, *Acta Cryst.*, 1950, 3, 10.
44. D.W.J. Cruickshank, *Acta Cryst.*, 1952, 5, 511.
45. Ring Index, Reinhold Publishing Corporation, New York, 1940.
46. D.W.J. Cruickshank, *Acta Cryst.*, 1949, 2, 65.
47. D.W.J. Cruickshank, *Acta Cryst.*, 1949, 2, 154.
48. D.W.J. Cruickshank, *Acta Cryst.*, 1950, 3, 72.
49. D.W.J. Cruickshank, and A.P. Robertson, *Acta Cryst.*, 1953, 6, 698.

50. C.A. Coulson, R. Daudel and J.M. Robertson,
Proc. Roy. Soc., 1951, A207, 306.
51. P. Vaughan and J. Donohue, Acta Cryst., 1952, 5, 530.
52. L. Pauling, The Nature of the Chemical Bond, Ithaca,
N.Y.: Cornell University Press 1940.
53. W. Gordy, J.Chem. Phys., 1947, 15, 81.
54. E.G. Cox and G.A. Jeffrey, Proc. Roy. Soc., 1951,
A207, 110.
55. J.G. White, J. Chem. Soc., 1948, p. 1398.
56. J.M. Robertson and J.G. White, J. Chem. Soc., 1945, p.607.
57. D.M. Donaldson and J.M. Robertson, Proc. Roy. Soc.,
1953, A220, 157.
58. J.M. Robertson, Proc. Roy. Soc., 1933, A142, 674.
59. D. Hodgkin, Ann. Rep. Progr. Chem., 1950, 47, 432.
60. A. Albert, Private communication.
61. G.A. Jeffrey and D.W.J. Cruickshank, Quart. Rev., 1953, 7, 336.
62. H. Lipson and W. Cochran, The Crystalline State,
Vol.III, London: G. Bell and Sons, 1953, p.307.
63. A.D. Booth, Reference 39, p.99.
64. W. Cochran, Acta Cryst., 1948, 1, 138.
65. D. Crowfoot, C.W. Bunn, B.W. Rogers-Low and
A. Turner-Jones, The Chemistry of Penicillin,
Princeton University Press, 1949, p. 310.
66. J.M. Robertson, J.Chem. Soc., 1935, p. 615.

ADDITIONAL PAPER

INVESTIGATION OF THE STRUCTURE OF
CELLULOSE FIBRES

I INTRODUCTION

The chemical work of Haworth, Irvine, Hirst and others during the years 1920 to 1930 showed that cellulose is a linear macromolecule consisting of a large number of β -glucopyranose units linked through 1:4- glucosidic bonds. The molecular weight of cellulose has been estimated by various chemical and physical methods as about 500,000 corresponding to about 3,000 glucose residues.

Early X-ray investigations of cellulose fibres by Nishikawa and Ono¹ in 1913 and by Herzog and Jancke² in 1920 demonstrated that they are crystalline and that different fibres have identical crystalline structures, but that this crystallinity is discontinuous. Later X-ray studies by Meyer and co-workers^{3,4,5} determined the unit cell dimensions, and a structure was put forward in which the chain molecules lie on the twofold screw axis parallel to b of a monoclinic cell containing four glucose residues. Two glucose units account exactly for the fibre period of 10.3\AA and the chain length corresponding to a molecular weight of 500,000 is thus about $15,000\text{\AA}$.

The X-ray investigations had shown that the fibre is not a single crystal but that the regular three-dimensional arrangement of glucose units, as defined by the unit cell, is built up to form small crystalline regions, or crystallites,

separated by more amorphous material. In addition, non-cellulosic components are also always present. The dimensions of these crystalline regions have been extensively studied by X-ray and chemical methods and more recently, by means of the electron microscope.

X-ray measurements by Herzog⁶, Hengstenberg and Mark⁷, and Kratky⁸ appeared to show that the crystallites in ramie are approximately 50-100Å wide and at least 500-600Å long, but rather lower values for the widths were obtained by Heyn⁹ for a number of different fibres. However, owing to various uncertainties and assumptions, these values may represent only an order of magnitude.

Chemical methods are based on degree of polymerisation measurements after partial acid hydrolysis, it being assumed that only the amorphous regions have been hydrolysed, the crystallites remaining intact.¹⁰ This has been carried out, among others, by Nickerson and Habrle¹¹ and Morehead¹² who report lengths of about 1,200-1,500Å for the crystallites of native fibres. Ranby and Ribl¹³ however found that after vigorous hydrolysis, most of the cellulose has a degree of polymerisation of about 100, corresponding to a crystallite length of about 500Å.

Electron microscopic investigations of cellulose have been limited by the necessity of obtaining sufficiently thin specimens. To achieve this, the fibres have been broken down

by chemical and physical methods such as partial acid or alkali hydrolysis or by beating, grinding or ultrasonic irradiation. Alternately replicas of the surfaces and of sections of fibres have been employed.

A large number of investigations have been reported in which fibres disintegrated by such treatments have been examined in the electron microscope. The results of the earlier studies are in agreement that the fibres break down into very long fine fibrils but differ considerably in the actual dimensions measured. More recent investigations have shown that the apparently intact cell walls of both plant and bacterial celluloses seem to consist largely of very long "microfibrils" 250-400Å wide,¹⁴⁻¹⁷ while drastic dispersion yields smaller units of approximately the same dimensions as the crystallites postulated from X-ray measurements. These smaller units have recently been extensively investigated with the electron microscope. Morehead¹² found that cotton, wood and ramie on acid hydrolysis followed by ultrasonic irradiation yielded particles which have widths of 64, 45 and 40Å respectively and lengths in close agreement with those calculated from degree of polymerisation experiments (1,200-1,500Å). Ranby and Ribí¹³ examined cotton and wood cellulose disintegrated by washing with distilled water to a pH of about 3.5, after treatment with sulphuric

acid. Under these conditions a colloidal solution of cellulose is formed, which was found to consist of particles 50-100Å wide and 500-600Å long. Using similar methods, Mukherjee, Sikorski and Woods¹⁸ showed that ramie, cotton and jute disintegrated into particles which are 155, 130 and 85Å wide respectively and which vary in length from 100Å to 2,500Å. Very long strands, 50-100Å wide, have been observed by Hodge and Wardrop¹⁹ in sections of delignified wood which had been disintegrated in a Waring Blendor and by Ranby and Ribl^{13,20} in cotton and wood broken down by ultrasonic irradiation, and also in sections of various natural and synthetic fibres. It was further found that these particles are in the form of flat filaments, about 30Å thick, developed on the (101) face,¹⁸ and that the larger microfibrils have also a similar shape.²¹

According to Kratky⁸ and Frey-Wyssling²² the small particles, all of about the same dimensions obtained in these investigations from a number of different celluloses, are "elementary fibrils" which correspond to the crystallites postulated from the earlier X-ray studies. The microfibrils of about 250Å width are considered to be made up of these crystalline units, a width of about 70Å and a thickness of 30Å being postulated. It is assumed²² that the elementary fibrils are surrounded by imperfectly crystalline, or "paracrystalline", cellulose which is responsible for their

aggregation into the larger microfibrils, the crystalline core of which is, in turn, embedded in a cortex of paracrystalline cellulose. The cell wall is then built up of these microfibrils.

The crystallites or elementary fibrils are not sharply defined entities in their natural state.^{3,23,24} It is believed^{24,25,26} that they merely represent randomly distributed regions of good crystallisation in a coherent three-dimensional network in which there is a gradual transition from regions of high lattice order to regions of low lattice order where the molecular chains are only approximately parallel. A single cellulose molecule will, in general, pass through several crystalline and intermediate paracrystalline areas, thus holding the structure together.

Ordinary light microscopy has shown that most cellulose fibres are made up of many individual cells, which in the case of sisal have a diameter of about 24 microns.²⁷ These consist of a thin outer primary wall and a thicker secondary wall which contains the bulk of the crystalline cellulose. The secondary wall is built up of layers of width estimated as 0.1-0.5 microns, which are arranged in the form of concentric tubes. This configuration is apparently due to the alternation of layers of material of high and low cellulose content. The layers of cellulose are, in turn, built up of

fibrils, a few tenths of a micron in diameter, which are oriented in a spiral at an angle of up to about 45° to the main fibre axis.

The orientation of the crystallites with respect to the fibre axis may be determined by X-ray methods, and by comparing the results of these measurements with microscopic measurements of fibril orientation, it has been shown that the long axes of the crystallites (or elementary fibrils) are aligned parallel to the long axis of the fibril. It is thus possible to relate X-ray measurements of crystallite orientation to fibril orientation as observed in the light microscope. Since the elementary fibrils are presumably oriented parallel to the length of the microfibril, it would appear that the microfibrils are, themselves, oriented parallel to the much larger fibrils. This has been confirmed in the case of the marine alga *Valonia ventricosa* by Preston and co-workers, who found that the microfibrils observed in the intact cell wall with the electron microscope¹⁴ showed the same orientation as had previously been determined for the crystallites and the fibrils from X-ray and light microscopic studies.²⁸ By means of electron diffraction measurements it has been shown that the molecular chains lie parallel to the length of the microfibril.²⁹

Other electron microscopic investigations^{15,19} of this nature have involved the use of material disintegrated by

rather drastic procedures and the interpretation of the resulting micrographs is difficult. Comparatively little direct evidence is thus available concerning the arrangement of the microfibrils in the cell wall.

In this investigation the structure of the hard leaf fibre sisal has been studied, firstly, by applying the disintegration techniques developed by the earlier investigators, and secondly, by attempting to cut sections of the intact material sufficiently thin for direct examination in the electron microscope.

Sisal has not been investigated in the past to the same extent as other fibres, due, no doubt, to the high proportion of non-cellulosic components which it contains. However, Mühlethaler¹⁶ found that after removal of impurities and disintegration in a Waring Blendor, sisal shows microfibrils 250-400Å wide, similar to other fibres. The orientation of the crystallites has been measured by X-ray methods, a value of about 18° for the spiral angle being given.^{30,31.}

No electron micrographs of sections of a native cellulose fibre have apparently been published. Hodge and Wardrop¹⁹ have however used sections of delignified wood disintegrated in a blendor and also replicas of sections, while Ribl,²⁰ although not showing any micrographs, states that microtome sections have been examined.

It was hoped that the present investigation would con-

firm the presence of elementary fibrils in sisal and that their dimensions could be measured. Electron micrographs of sufficiently thin sections should reveal the organisation of the cell wall, from the elementary fibril, through the microfibril, to the coarser fibrils and layers visible in the light microscope. This would, naturally, also determine what degree of individuality these postulated units possess in the intact fibre.

II EXPERIMENTAL

The material employed throughout this investigation was sisal fibre supplied by the Gourock Rope Work Co. The micrographs were obtained with a Philips Metalix electron microscope.

(a) Investigation of Fibres Disintegrated by Acid Hydrolysis.

Following the procedure of Mukherjee et al.,¹⁸ the hydrolysis was carried out at room temperature with solutions of sulphuric acid of concentrations varying from 950 to 1,000 gm./litre, acting for periods of one to three weeks. The treated fibres were centrifuged and the supernatant acid removed. Further washings by centrifugation after adding distilled water were continued till neutrality was reached, the supernatant being rejected each time. As had been observed in previous investigations,^{13,18} when the pH of the supernatant approached a value of about 3 - 3.5, the fibres disintegrated, forming a solution of colloidal cellulose. Specimens for examination in the electron microscope were prepared by evaporating drops of the solution on Formvar covered mounts and shadowing with palladium-nickel at 15°.

When sisal is placed in sulphuric acid of such concentrations the solution remains colourless for 24 hours, but during the next 24 hours the solution and the fibres themselves become coloured a deep red, presumably due to the extraction

of some impurity in the fibre. Washing with distilled water did not suffice to remove this impurity, and the electron micrographs obtained from the examination of such solutions were not satisfactory.

This red colouration reaches a maximum after about two days and if the discoloured acid is then removed and replaced by fresh acid, no further discolouration takes place. This was carried out, the fresh acid being allowed to act for several days. Much more satisfactory results were obtained.

Since sisal has a relatively high content of non-cellulosic materials, particularly lignin, a large proportion of this part of the work was concerned with the problem of obtaining a clean sample of colloidal cellulose. Several different methods of achieving this were attempted. Finally a rigorous extraction process described by Preston³² was employed, in which pectins are removed by treatment, first, with a 50% solution of hydrochloric acid in alcohol and then with ammonium oxalate solution, hemicelluloses, by extraction with hot dilute sodium hydroxide solution and lignin, by treatment alternately with chlorine water and hot dilute sodium sulphite solution till no colouration results. A white, clean-looking product was obtained which gave no discolouration on sulphuric acid hydrolysis.

Plate 1 is a micrograph of a specimen prepared in this

way. It shows the characteristic cellulose particles similar to those obtained in previous investigations on other fibres. There does not appear to be any contamination by impurities.

(b) Investigation of Fibres Using a Thin-Sectioning Technique.

The main difficulty of cutting sections for examination in the electron microscope is the necessity of making these sufficiently thin. Ideally their thickness should not exceed 0.1 microns, and although this was not in fact achieved, a number of sections were obtained which gave quite interesting electron micrographs. I am indebted to Dr. J. Hossack of this Department for the cutting of the sections.

Since sisal is a stable material no fixation was required. The fibres were dehydrated in absolute alcohol and impregnated with monomeric n-butyl methacrylate containing 5% methyl methacrylate. The blocks, prepared by heat polymerisation at 60° C. were cut on a Spencer Rotary Microtome modified by the inclusion of a Pease and Baker adaptor, using glass knives. The sections were floated on aqueous alcohol and picked up on Formvar covered specimen mounts.

Staining of cellulose has not been very successful in the past and instead, the plastic was dissolved out with acetone and the specimens shadow-cast with palladium-nickel in the usual way. The removal of the plastic appears to be

necessary since electron micrographs of unextracted sections show extremely poor contrast.

Plate 2 which is a micrograph of an unshadowed specimen shows part of a slightly oblique cross-section of a fibre, parts of several of the individual cells being visible.

It was found that blocks prepared in this way from intact fibres did not cut well. This was probably due to the fact that the fibres, having a hard close texture, had not been impregnated sufficiently well with monomer. It was therefore decided to use material from which the non-cellulosic components had been removed. The purification was carried out in the same way as for the sulphuric acid degradations (see page 100). The cleaned material was placed in distilled water and shaken up for a few minutes, when it disintegrated into a loose pulp consisting of smaller fibres, presumably the individual cells of which sisal is composed.²⁷ This was prepared for cutting in the same way as before. The blocks thus contain a randomly oriented mass of small fibres which could be cut much better.

However, owing to the random orientation of the small fibres in the block, it was a matter of some difficulty to determine the angle at which a given fibre was being cut. Also, since the purification process is of such a drastic nature, some degradation of the cellulose and considerable swelling and distortion probable take place.

Plates 3 - 7 are electron micrographs of shadow-cast sections of purified sisal prepared in this way.

III RESULTS

Plate 1 is a micrograph of sisal disintegrated by treatment with 1,000 gm./litre sulphuric acid solution at 18° C. for 13 days after rigorous extraction of impurities.

The width of the smallest particles visible is about 90Å. A larger number have a width of about 170Å, while the majority of the particles have a width of around 300Å. Although intermediate values also occur it does appear that a large proportion of the particles obtained have one of these three widths. The thickness of the particles, measured from the length of the shadows, was found to vary considerably, the minimum being about 20Å. The lengths of the particles vary from approximately 500Å to 5,000Å.

Plate 2 is a micrograph of a section of an intact fibre, showing the individual cells. Some detail is visible in the cell wall, but its interpretation does not seem possible.

Plates 3 - 7 show sections of purified sisal fibres cut at various angles. Plate 3 is a micrograph of a longitudinal section. The approximately parallel layers which are visible have a width of about 1,400Å and appear to be broken up into lengths varying from about 3/4 to 2 microns. Also visible are some very fine long strands about 300Å wide and of length of the order of one micron. Plate 4 shows a cross-section. The diameter of the cell shown is about 20

microns. It can be seen that the layers are arranged in concentric rings, which are however not continuous. The width of the edge of the rings is about $1,400\text{\AA}$.

Plates 5 and 6 show the same oblique section, Plate 6 being a twofold enlargement of part of Plate 5. Here there is considerably greater disorder in the arrangement of the layers, due perhaps to a greater degree of degradation during the purification process. The width of the layers is rather variable, but the average seems to be again about $1,400\text{\AA}$. The lengths vary considerably, going up to about 3 microns.

In Plate 7, the layers have been cut obliquely and something of their internal structure is visible. A few particles of about 100\AA width are present. However, the majority of the units have a greater width, of the order of $300 - 500\text{\AA}$, with quite a large number having a width of about 300\AA . The particles seem to lie more or less perpendicular to the plane of the section, and their length appears to be equal to the height of the section. This would suggest that they are parts of a longer system.

IV DISCUSSION

The 90Å width of the smallest particles obtained in the sulphuric acid degradations agrees fairly well with the determination of crystallite size from X-ray measurements, and also with the results of other electron microscopic investigations. It would thus appear that these particles correspond to the postulated elementary fibrils of sisal. The measurements are, however, only approximate and smaller particles, if present, may not be resolvable by the microscope.

Attempts to cut sufficiently thin sections of intact fibres have not been successful, and in the micrographs obtained, very little detail is visible in the cell wall. Comparison with the micrographs of sections of purified material shows that a great deal of non-cellulosic matter is present.

On the assumption that the section is oblique, the diameter of the cell shown on Plate 2 appears to be approximately 9 microns, considerably less than the value obtained from light microscopic studies. The circumference of the cell, however, would correspond to a diameter of about 23 microns, in excellent agreement with the optical measurements, suggesting that it is really a cross-section, with the fibre

having been squashed during the cutting. The cell shown on Plate 4, which is, however, a micrograph of purified material, also has a diameter of about this length.

Much more interesting results have been obtained with the purified material. The layers of cellulose, whose width could merely be estimated as being of the order of 0.1-0.5 micron from light microscopic studies have been found, in actual fact, to be 1,400Å wide in sisal. In oblique and longitudinal section these layers appear to be broken up into lengths varying from about 3/4 to 3 microns. This may be due to degradation of the cellulose in the purification process, or to the resolution of the individual fibrils of which the layers are composed. The discontinuity of the layers seen in cross-section on Plate 4, could also be due to fibrils of about 1,400Å diameter lying roughly perpendicular to the plane of the section.

Plate 7 shows that the layers of 1,400Å width are built up of small particles, the majority of which appear to have approximately the same width (300 - 500Å) as the microfibrils discovered in the earlier electron microscopic studies.¹⁴⁻¹⁷ The full length of these particles cannot be estimated from this plate, but they appear to be at least as long as the section is thick (probably about 0.2 micron). On Plate 3 can be seen strands of about the same width which have lengths

of the order of one micron. However, since in every case, one end of these merges into the layers, this value also, represents only a lower limit for their length, in agreement with the indefinite length found for the microfibrils in the earlier investigations. It is of interest that particles which are 300Å wide constitute the main product of the chemical degradations (see Plate 1).

The particles of 100Å width also visible on Plate 7, may correspond to the elementary fibrils which have previously been found only in material which has undergone drastic disintegration. They could, however, represent edge-on views of the larger, ribbon-shaped microfibrils.

The results of this investigation are thus in agreement with the recent theories of Kratky⁸ and Frey-Wyssling.²² The postulated elementary fibrils and microfibrils have, however, been visible only in sections of material which may have undergone considerable degradation.

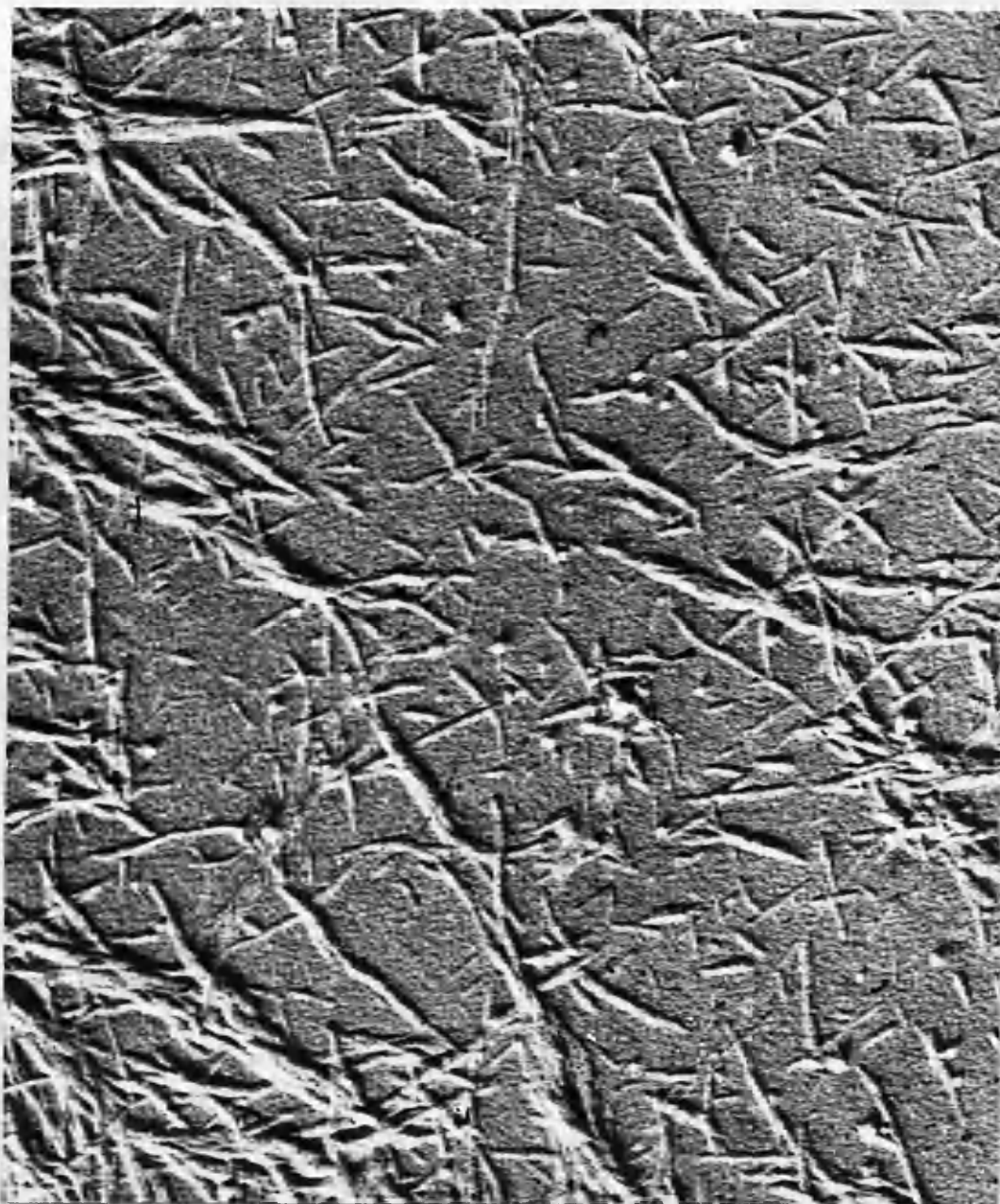


Plate 1.

Sisal disintegrated with sulphuric acid. Shadowed with Pd - Ni at 15° . Magnification x 41,800.



Plate 2.

Section of intact fibre. Unshadowed. Magnification x 7,300.



Plate 3.

Longitudinal section of purified material. Shadowed with Pd - Ni at 15° . Magnification x 9,200.

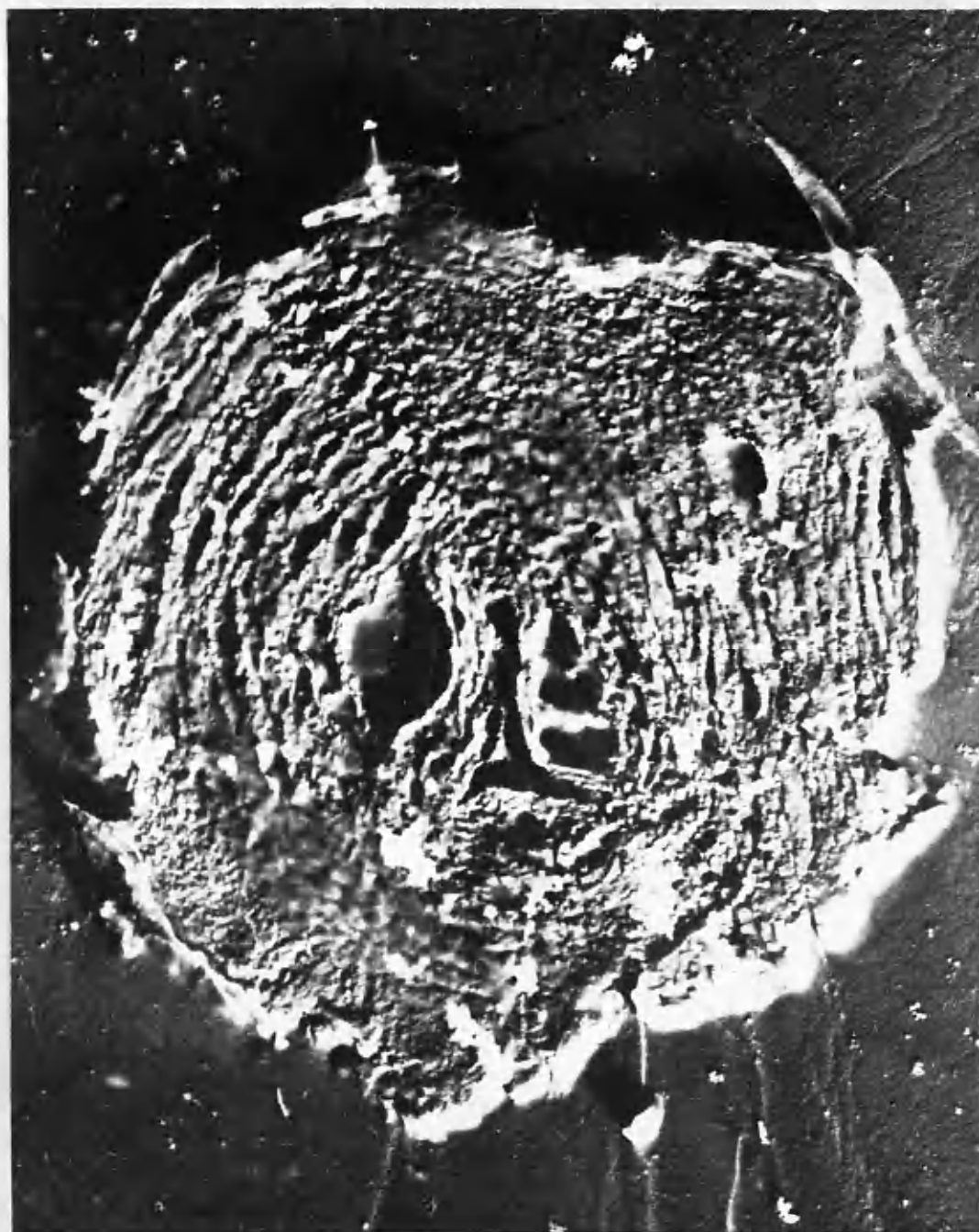


Plate 4.

Cross-section of purified material. Shadowed with Pd - Ni
at 15° . Magnification x 5,800.

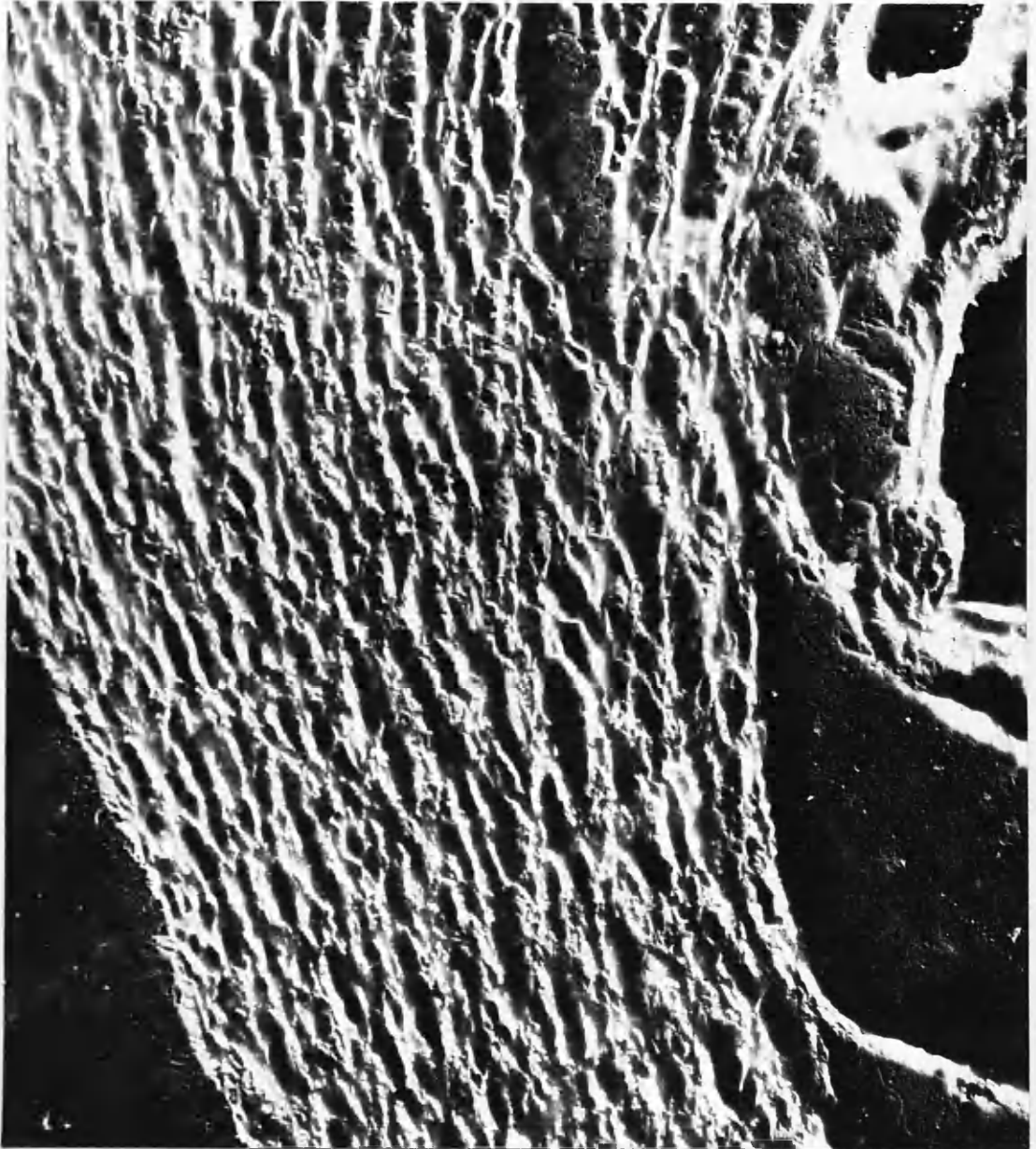


Plate 5.

Oblique section of purified material. Shadowed with pd - Ni
at 15°. Magnification x 7,700.



Plate 6.

Part of Plate 5 enlarged. Magnification x 15,300.



Plate 7.

Slightly oblique cross-section of purified material. Shadowed Pd - Ni at 15° . Magnification x 24,000.

REFERENCES

1. S. Nishikawa and S. Ono, Proc. Phys.-Math. Soc. Tokyo, 1913, 7, 131;
S. Nishikawa, Proc. Phys.-Math. Soc. Tokyo, 1914, 7, 296.
2. R.O. Herzog and W. Jancke, Ber., 1920, 53B, 2162.
3. K.H. Meyer and H. Mark, Ber., 1928, 61B, 593.
4. K.H. Meyer and H. Mark, Z. physik. Chem., 1929, B2, 115.
5. K.H. Meyer and L. Misch, Helv. Chim. Acta, 1937, 20, 232.
6. R.O. Herzog, J. Phys. Chem., 1926, 30, 457.
7. J. Hengstenberg and H. Mark, Z. Krist., 1928, 69, 271.
8. O. Kratky, Kolloid-Z., 1951, 120, 24.
9. A.N.J. Heyn, J. Amer. Chem. Soc., 1950, 72, 5768.
10. G.F. Davidson, J. Textile Inst., 1943, 34, T87.
11. R.F. Nickerson and J.A. Habrle, Ind. Eng. Chem., 1947, 39, 1507.
12. F.F. Morehead, Textile Research J., 1950, 20, 549.
13. B.G. Ranby and E. Ribl, Experientia, 1950, 6, 12;
B.G. Ranby, Acta Chem. Scand., 1949, 3, 649.
14. R.D. Preston, E. Nicolai, R. Reed and A. Millard, Nature, 1948, 162, 665.
15. A. Frey-Wyssling, K. Mühlethaler and R.W.G. Wyckoff, Experientia, 1948, 4, 475.
16. K. Mühlethaler, Biochim. et Biophys. Acta, 1949, 3, 15.

17. K. Mühlethaler, *Biochim. et Biophys. Acta*, 1949, 3, 527.
18. S.M. Mukherjee, J. Sikorski and H.J. Woods, *Nature*, 1951, 167, 821; S.M. Mukherjee and H.J. Woods, *Biochim. et Biophys. Acta*, 1953, 10, 499.
19. A.J. Hodge and A.B. Wardrop, *Nature*, 1950, 165, 272.
20. E. Ribi, *Nature*, 1951, 168, 1082.
21. R.D. Preston, *Discussions Faraday Soc.*, 1951, No.11, 165.
22. A. Frey-Wyssling, *Science*, 1954, 119, 80.
23. W.T. Astbury, *Trans. Faraday Soc.*, 1933, 29, 193.
24. A. Frey-Wyssling, *Protoplasma*, 1936, 25, 261.
25. O. Kratky and H. Mark, *Z. Physik. Chem.*, 1937, B36, 129.
26. J.A. Howsmon and W.A. Sisson, in *High Polymers*, Vol. V, 2nd ed., New York-London: Interscience 1954, p. 249.
27. A.J. Turner, *J. Textile Inst.*, 1949, 40, P972.
28. W.T. Astbury, T.C. Marwick and J.D. Bernal, *Proc. Roy. Soc.*, 1932, B109, 443; R.D. Preston and W.T. Astbury, *Proc. Roy. Soc.*, 1937, B122, 76.
29. R.D. Preston and G.W. Ripley, *Nature*, 1954, 174, 76.
30. W.A. Sisson, *Ind. Eng. Chem.*, 1935, 27, 51.
31. R.D. Preston, *The Molecular Architecture of Plant Cell Walls*, London: Chapman and Hall, 1952, p. 130.
32. R.D. Preston, Reference 31, pp. 26-30.