2 cys

UCRL-19906

X-RAY STRUCTURE INVESTIGATION OF SOME SUBSTITUTED INDOLES, AND THE X-RAY CRYSTAL OF 1, 1'-BISHOMOCUBANE

William G. Quarles
(Ph. D. Thesis)

May 1970

AEC Contract No. W-7405-eng-48

LAWRENCE RADIATION LABORATORY
UNIVERSITY of CALIFORNIA BERKELEY

UCRL-19906

DISCLAIMER

Dedicated to my parents.

THIS PAGE

WAS INTENTIONALLY

LEFT BLANK

X-ray Structure Investigation of Some Substituted Indoles, and the X-ray Crystal Structure of 1,1'-Bishomocubane.

Contents

Abstract v

Introduction 1

I. The X-ray Crystal Structure of 1,1'-Bishomocubane 6

A. Introduction 7

B. Experimental 8

C. Solution of the Structure 17

D. Discussion of the Structure 23

E. Conclusion 36

II. A Computer Program for the Solution of Crystal Structures Containing a Planar Moiety with Known Geometry 38

A. Introduction 39

B. Use of the Program 41

C. Usefulness of the Program in General . . . 50

D. Conclusion 57

III. The Crystal and Molecular Structure of 5-Methoxytryptamine 59

A. Introduction 60

B. Experimental 63

C. Solution of the Structure 73

D. Discussion of the Structure 93

E. Conclusion 123

IV. The Crystal and Molecular Structure of Melatonin 128

A. Introduction 129

B. Experimental 130

C. Solution of the Structure 133

D. Discussion of the Structure 138

E. Conclusion 159

V. Acknowledgments 161a

VI. Appendix A

1. Derivation of the matrix a_{ij} 162

2. Fortran Listings 166

X-RAY STRUCTURE INVESTIGATION OF SOME SUBSTITUTED INDOLES, AND THE X-RAY STRUCTURE OF 1,1'-BISHOMOCUBANE

William G. Quarles

Lawrence Radiation Laboratory and
Chemistry Department
University of California
Berkeley, California
May, 1970

Abstract

The crystal structures of 5-methoxytryptamine, melatonin, and the p-bromobenzoate of 1,1'-bishomocubane have been solved by x-ray diffraction methods. A computer program for the trial and error solution of crystal structures is also described here.

The molecular structure of 1,1'-bishomocubane has been solved to an R factor accuracy of 5.5% using 990 independent manual diffractometer data, of which 865 were non-zero. Standard deviation on bond lengths is $\pm$ .02 Å. The compound crystallizes in space group $P2_1/c$ with $\underline{a}$ = 6.379 $\pm$ .001 Å; $\underline{b}$ = 26.07 $\pm$ .005 Å; $\underline{c}$ = 8.443 $\pm$ .002 Å; β = 96.87 $\pm$ .01°; λ = 1.54051 Å; Z = 4. Density calculated on these cell

dimensions is 1.575 g/cc^3, and density measured by floatation in ethylene bromide and ethylene chloride is 1.560 g/cc^3. The compound is sensitive to x-rays, and decomposes anisotropically. Addition of two extra carbons into one of the cubane cyclobutane rings causes two of the remaining cyclobutane rings to pucker by 27° from a planar configuration while the other two cyclobutane rings stay planar within the standard deviations of the determination.

The crystal structure of 5-methoxytryptamine has been solved using a trial and error computer program which is also outlined in this thesis. The compound crystallizes in the non-centric monoclinic space group Pc, and $\underline{a}$ = 6.110 $\pm$ .002 Å; $\underline{b}$ =9.532 $\pm$ .003 Å; $\underline{c}$ = 8.831 $\pm$ .003 Å; β - 98.72 $\pm$ .01°; λ - 1.54051 Å; Z = 2. The density calculated on the basis of these cell dimensions is 1.242 g/cc^3. The density measured by floatation at room temperature in ethyl acetate and ethylene chloride is 1.245 g/cc^3. The structure was refined against the 759 independent θ-2θ scan automatic diffractometer data, of which 17 were zero to an R factor of 2.5%. Standard deviations on bond lengths are .003 Å. Within the standard deviations of the determination, the indole ring is not planar. Two carbons of the benzene portion are warped above the plane of the ring at an angle of 1.6°. Short bonds correlate with high π electron density as calculated by molecular orbital theory. One of the shortest N-H-N hydrogen bonds yet reported, 2.916 Å, is formed between

the primary amine nitrogen of the aliphatic chain and the nitrogen of the indole ring, which donates its hydrogen for the formation of this bond.

The crystal and molecular structure of melatonin has been solved using statistical methods and automatic diffractometer θ-2θ scan data. This compound crystallizes into space group $P2_1/c$, with $\underline{a}$ = 7.707 $\pm$ .002 Å; $\underline{b}$ = 9.252 $\pm$ .002Å; $\underline{c}$ = 17.007 $\pm$ .004 Å; β = 96.78 $\pm$ .03°; λ = .709261 Å; Z = 4. Density calculated on the basis of these cell constants is 1.276 g/cc^3, and the density measured by floatation in ethylene chloride, ethylene bromide, and ethyl acetate is 1.272 g/cc^3. The structure was refined to a 3.5% R value against the 1140 independent data, of which 808 were non-zero weight. The indole ring is not planar within the standard deviation of the structure determinations. The carbons C(3) and C(10) of the pyrrole ring are warped above the ring by 1.8°.

The findings of the x-ray work are entirely consistent with the dual conformation theory of serotonin.

Introduction

One of the most exact and thorough characterizations of a molecular structure possible results from careful analysis of the way a single crystal diffracts x-rays. Standard deviations of bond lengths are ± .003Å for one of the compounds studied in this thesis. These highly accurate structures are the result of improved techniques. Diffraction data are collected with the help of a computer-controlled diffractometer. Many of the calculations are done by high-speed computers such as the CDC-6600.

This thesis is concerned with the application of x-ray crystallography to a precise molecular geometry determination of some substituted indoles, and 1,1'-bishomocubane. The thesis is divided into four independent sections, each with its own bibliography.

The crystal and molecular structure of 1,1'-bishomocubane is presented in Section I. This compound is important in the study of strained hydrocarbon ring systems.

The logic and instructions for use of a computer program designed to solve crystal structures by trial and error is related in Section II. This program will work best for molecules possessing a planar moiety with a known geometry. This computer program was used to solve the crystal structure of 5-methoxytryptamine. The solution of this crystal

structure is related in Section III. The crystal of a similar compound, melatonin, or N-acetyl-5-methoxytryptamine is given in Section IV. A Fortran listing of the trial and error computer program can be found in Appendix A, and the derivation of the orientation matrix for the program may also be found in Appendix A.

The major part of this thesis is concerned with the molecular structure of substituted indole compounds, and thus these compounds will be described more completely here in terms of their role in biochemical metabolism.

Since the isolation of serotonin from clotted blood[1] there has been a flurry of scientific activity concerning the metabolism and fate of the indole alkyl amines. Approximately 500 papers were published last year which treated the internal metabolism of these interesting indoles. It would be impossible, obviously, to do all these papers justice in a PhD. thesis. I do wish to outline here first the broad outlines of serotonin metabolism in the human brain, and then consider the possible implications of the molecular structures of 5-methoxytryptamine and melatonin.

The essential amino acid tryptophane is hydroxylated in the human brain to 5-hydroxytryptophane.[2] This hydroxylation is followed by decarboxylation initiated by 5-hydroxytryptophane decarboxylase, and serotonin is thus synthesized in the human brain.[3] Serotonin is stored after synthesis

in "granules" morphologically similar to pinched nerve endings.[4] These "granules" are surrounded by mitochondria containing monoamine oxidase.[5] This enzyme will destroy unprotected serotonin very quickly. The half-life of serotonin in the brain is estimated with the use of monoamine oxidase inhibitors to be 10 to 30 minutes.[6] Monoamine oxidase will oxidize serotonin to 5-hydroxyindoleacetic acid. This compound is excreted in the urine. Schizophrenics excrete an excess of this compound.[7] Serotonin is concentrated in the hypothalmus, mesencephalon, and pineal gland of the human brain. Concentration of serotonin in the brain has been correlated with many central and important biological effects. Briefly, serotonin concentration of turnover rate is correlated with regulation of body temperature, sleep, sexual activity, and hallucinogenic activity of indole alkyl amines such as N,N-dimethyltryptamine, and psilocin, or 4-hydroxy-N, N-dimethyltryptamine. The structural similarity of these compounds with 5-hydroxytryptamine, serotonin, is obvious.[8] It is impossible to enumerate here all the pathways of biological metabolism in which serotonin has been implicated even if we narrow our focus to the human brain. Several reviews are listed in the Bibliography to Section III.

One of the most interesting structural problems concerned with the metabolism of serotonin as follows: exactly how is serotonin stored in the brain unharmed by monoamine

oxidase, and structurally how does serotonin interact with the serotonin receptor sites to cause physiological and physical changes. There are currently no specific ideas about how serotonin is bound or protected from monoamine oxidase.

Serotonin has been shown to interact specifically with both smooth muscle and nerve tissue. Drugs which inhibit the nerve interactions do not inhibit the smooth muscle interactions, and vice versa. This fact caused Gaddum to propose the dual conformation theory of serotonin interaction with receptor sites.[9] Kier analysed calculated possible conformations of serotonin as a function of energy, using Hueckel orbital theory. His conclusion was that serotonin should have only one conformation in solution.[10]

The crystal structure of 5-methoxytryptamine is presented in Section III, and the crystal structure of melatonin or N-acetyl-5-methoxytryptamine is presented in Section IV of this thesis. The x-ray work is completely consistent with the dual conformation theory of Gaddum. Melatonin crystallizes into the minimum energy conformation calculated by Kier. On the other hand, 5-methoxytryptamine crystallizes into a much higher energy conformation. Further discussion of serotonin metabolism related to molecular structure can be found in the conclusions of Sections III and IV.

Bibliography

1. M. M. Rapport, A. A. Green, and I. H. Page, Science, 108, 319 (1948).
2. E. M. Gal, M. Poczik, and F. D. Marshall Jr., Biochem., Biophys Research Communs., 12, 39 (1963).
3. W. Sacks, J Appl. Physiol, 16, 1050 (1961).
4. E. A. Carlini and J. P. Green, Brit. J. Pharmacol., 20, 264 (1963).
5. D. F. Bogdanski, A. Weissbach, and S. Udenfriend, J. Neurochem., 1, 272 (1957).
6. S. Udenfriend, H. Weissbach, Proc. Soc. Exper. Biol Med., 97, 748 (1958).
7. G. G. Brune, and H. E. Himwich, Science, 133, 190 (1961).
8. I. H. Page, "Serotonin," Year Book Medical Publishers, Chicago, 1968.
9. J. H. Gaddum and K. A. Hameed, Brit. J. Pharmacol., 9, 240 (1954).
10. Lemont B. Kier, J. Pharmacol Sci., 57, (7), 1188 (1968).

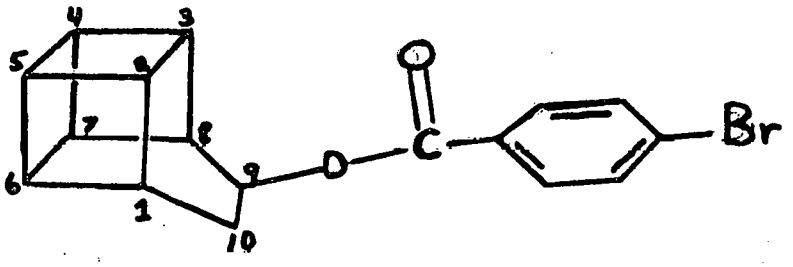
Section I

The X-ray Crystal Structure of 1,1'-Bishomocubane

A. Introduction

Since the original synthesis of the carbon cage compound cubane,[1] and the publication of its x-ray crystal structure,[2] there has been much synthetic and analytical work on this series of cage-type hydrocarbons. Cubane is an octane which assumes the shape of a cube. Cubane and its analogues are interesting model compounds in the study of strained saturated hydrocarbon rings.[3]

Although cubane itself is crystalline at room temperature, other members of this series are extremely volatile liquids, or near liquids at room temperature; and derivatives which form better crystals are studied in x-ray work. The names for these compounds are varied. {One possible name for the compound studied here is the p-bromobenzoate of pentacyclo ($4.4.0.0^{2,5}.0^{3,8}.0^{4,7}$) decan-9-ol.[4]} A schematic of the compound with this numbering system is shown below. It is more convenient for writing to abbreviate this complete chemical name with the common name for the hydrocarbon suggested by Dauben; 1,1' -bishomocubane.[5] This is the name for the compound which will be used throughout this section.



B. Experimental

Crystals of the p-bromobenzoate derivative of 1,1'-bishomocubane were supplied by Professor W. G. Dauben of the chemistry department here. The crystals were colorless rectangular plates that were largely transparent to white visible light. They were soft, and they cleaved easily and cleanly with a razor blade. These crystals are stable in air and show no pronounced tendency to absorb water.

Preliminary oscillation and Weissenberg X-ray photographs of a rather large (.43 x .38 x .43 mm) crystal showed Laue symmetry and extinctions consistent with space group $P2_1/c$. This crystal was mounted with the fiber direction nearly parallel to the b axis. All crystals of 1,1'-bishomocubane upon which measurements were made were mounted on thin glass fibers with General Electric Number 1202 Clear Industrial Glyptal Varnish. The a and c axes were chosen consistent with the extinction rules 0k0, k = 2n; h0ℓ, ℓ = 2n. A truncated set of Weissenbergs, k = 0-3, was taken on this crystal with unfiltered iron radiation. A complete set of Weissenbergs was taken later on another crystal. The crystal received a total of 33 hours of iron radiation at 10 ma and 30 kv, and was transferred to a General Electric XRD-5 manual diffractometer equipped with a NaI scintillation counter, a pulse height discriminator, a molybdenum

x-ray tube, and a quarter circle Eulerian cradle goniostat. The molybdenum x-rays were filtered through a .0004 inch zirconium filter mounted on the receiving slit of the scintillation counter, and Bragg reflection angles for the resolved $K\alpha_1$ and $K\alpha_2$ doublet were measured through a narrow slit along the h00, 0k0, and 00ℓ diffraction directions at a 2 degree tube take-off angle. During the measurements, the crystal received an additional 20 hours of 20 ma, 40 kv molybdenum x-rays. It was found out later that these crystals were sensitive to x-ray damage, and that the cell dimensions as well as intensities were a function of this damage.

As a consequence, the cell dimensions obtained in the above way, though precise, were not very accurate. They are listed here for later comparison. All calculations were based on the high angle $K\alpha_1$ (.709261), and $\underline{a}$ = 6.378 ± .003; $\underline{b}$ = 26.16 ± .01; $\underline{c}$ = 8.480 ± .004; β = 97.17; Z = 4. A check of a few diffraction peaks showed that there was considerable overlap of diffraction intensities in the counter window at a 4° take-off angle with no slit.

In order to eliminate overlapping of diffraction peaks, longer wavelength copper radiation, which gives larger diffraction angles for the same lattice spacings was chosen as more suitable for intensity measurement than molybdenum.

With the same diffractometer set-up as before, except for a copper x-ray tube and a .0005 inch nickel filter, cell

constants were determined on another, smaller crystal (.26 x .27 x .15 mm) which had less than one hour of 1 ma, 40 kv preliminary copper x-ray exposure. These cell constants were: $\underline{a}$ = 6.379 ± .001; $\underline{b}$ = 26.070 ± .005; $\underline{c}$ = 8.443 ± .002; β = 96.87 ± .01, λ = 1.54051. As a function of x-ray exposure, the $\underline{b}$ and $\underline{c}$ axes increased in length and the β angle expanded. Since the crystal was mounted parallel to the $\underline{b}$ axis, it was possible to obtain accurate measurements of the β angle as a function of x-ray exposure. The damage also resulted in an anisotropic reduction of a few diffraction intensities which were also measured in the experiment. This semi-quantitative estimate of damage is summarized in Table I-1.

Table I-1

Estimate of 40kv Copper X-ray Radiation Damage in 1,1'-Bishomocubane.

β (degrees)[a]	Time (hours)		Anisotropic Damage[b]	
	6ma	20ma	$(I_{002}/I_{00\bar{2}})$	$(I_{200}/I_{\bar{2}00})$
96.87	4	--	.88	.85
96.87	5	--	.88	.85
96.89	9	--	.78	.79
96.99	12	5	.80	.80
97.06	--	13	.72	.75

[a] Standard deviation of β measurement ± .01 degree

[b] I_{hkl} means total counts of reflection hkl measured in ten seconds minus a ten-second background count measured one degree lower in 2θ.

Density calculated on the basis of the copper cell constants listed above was 1.575 g/cc^3. Density measured by flotation in ethylene chloride and ethylene bromide was 1.560 g/cc^3.

In order to reduce damage, the equipment was modified to minimize x-ray exposure. A .0005 inch nickel filter was carefully taped to the x-ray collimator such that the narrow orifice near the x-ray window on the beam side was covered with nickel foil, but no tape was in the path of the beam. A fresh crystal was mounted which looked suitable for intensity measurements. The fiber direction was closest to the c* axis, and alignment on the diffractometer was adjusted until the c* axis of the crystal was approximately parallel to the phi rotation axis of the instrument.

The chi = 90° absorption test showed a 12% variation of intensity as a function of 360 degree rotation in phi for the 002 reflection under data-taking conditions of 4° tube take off angle, and 20 ma, 40 kv x-rays. The linear absorption coefficient, μ, for copper x-rays and 1,1'-bishomocubane was 44.1 cm^{-1}, and μt in the longest direction was .596. The dimensions of the crystal were (.10 x .10 x .036 mm), and no absorption correction was made.

In the preliminary line-ups, the crystal received less than one hour of 1 ma, and less than 30 minutes of 20 ma 40 kv copper radiation. In order to reduce exposure time,

all of the backgrounds were estimated from a curve of 2θ versus intensity, which was prepared for a number of different values of chi and phi. Diffraction intensities were measured at a 4° take-off angle with a stationary crystal, stationary counter technique. The crystal received radiation only during the 10 seconds that each peak was counted. Net intensities were obtained by direct subtraction of estimated 10 second backgrounds from the measured 10 second point counts for each hkl. A complete set of three dimensional intensity data was taken out to 85 degrees in 2θ ($\sin\theta/\lambda = .43855$) over a period of 36 straight hours. The Laue reflections hkl and $\bar{h}$kl only were measured; thus, there was one measurement for each independent reflection. A total of 990 independent reflections were measured, of which 38 were zero and 87 others had intensities smaller than one standard deviation of intensity.

Four diffraction standards measured at intervals of about 2 hours showed anisotropic decomposition of between 8 and 16 percent despite the precautions taken to minimize damage. Over the period of time involved in the intensity measurements, the b axis increased in length by approximately 1 percent.

All calculations made on 1,1'-bishomocubane were carried out on the CDC-6600 computer. The standard Fourier, Least Squares, Distan, and data processing programs were all

written by Dr. Allan Zalkin of this laboratory. The data processing programs sort, blend, correct for Lorentz and polarization effects, and estimate standard deviations for all input intensities such that a corrected, Fo, observed structure factor is produced. The Least Squares program minimizes the function

$$\Sigma \underline{w}(|kFo| - |Fc|)^2/\Sigma \underline{w}|kFo|^2$$

where Fc is the calculated structure factor, Fo is the observed structure factor from the corrected intensities, k is a linear scale factor, and $\underline{w}$ is a weighting factor. In the early stages of refinement $\underline{w} = 1$. In the later stages of refinement $\underline{w} = [\sigma(Fo)]^{-2}$. In the equations which follow, Fo will be abbreviated with F. The quantity $\sigma^2(F)$ is calculated from $\sigma(F)$. Thus, if $I \leq \sigma(I)$,

$$\sigma(F) = [\sigma(F^2)]^{1/2}$$

and if $I > \sigma(I)$

$$\sigma(F) = F - [F^2 - \sigma(F^2)]^{1/2}$$

where

$$\sigma(I) = [I + 2I_b + (\Delta I_b)^2 + \{(S)(I)\}^2]^{1/2}$$

and

$$\sigma(F^2) = (LP)^{-1}\sigma(I)$$

In the equations above $\sigma(F)$, $\sigma(F^2)$, $\sigma(I)$ are the standard deviations of the quantities involved. The symbol F represents the observed structure factor; I represents observed intensity; (LP) is the Lorentz-polarization correction; I_b is the intensity of the background; ΔI_b is the uncertainty in the background; S is an estimate of the fraction of observed intensity suffering from systematic error.

For the final refinements of 1,1'-bishomocubane, S was set equal to .06, and w was set equal to zero for reflections where $I = 0$ or $I \leq \sigma(I)$. The uncertainty in the background, ΔI_b, was ten counts, except in the last refinements, where it was twenty counts.

After each cycle, the Least Squares program produced the following criteria of fit between observed and calculated structure factors:

$$R_1 = \Sigma(|kFo| - |Fc|)/\Sigma|kFo|$$

where R_1 is the conventional R factor. The quantities involved were defined earlier. The R_1 reported throughout this thesis does not include zero weight data unless otherwise specified. The weighted R factor, R_2, is also produced by the program.

$$R_2 = \{\Sigma w(|kFo| - |Fc|)^2/\Sigma w|kFo|^2\}^{1/2}$$

A third criterion of fit is the standard-deviation-of-

observation unit weight, SD.

$$SD = [(\Sigma w(|Fo| - |Fc|)^2/(\underline{n}-\underline{p})]^{1/2}$$

where $\underline{n}$ is the number of data and $\underline{p}$ is the number of parameters.

The Least Squares program also produced estimated standard deviations of bond distances and angles. These standard deviations will be reported throughout this thesis, and are the larger of the two quantities

$$(\Sigma\Delta_c^2)^{1/2}/(n-1) \quad \text{or} \quad (\Sigma\sigma_i^2)^{1/2}/n$$

where Δ_i is the difference between the ith measurement and the average of n measurements, and σ_i is the error of the value estimated from the accuracy of the atomic coordinates.

The temperature factors used are of the form

$$\exp(-B \cdot (\sin\theta/\lambda)^2)$$

if isotropic, and

$$\exp(-h^2\beta_{11} - k^2\beta_{22} - \ell^2\beta_{33} - 2hk\beta_{12} - 2k\ell\beta_{23} - 2h\ell\beta_{13})$$

if anisotropic. The thermal parameters B, where B is a constant, and B_{ij} where $B_{ij} = 4\beta_{ij}/a_i^* a_j^*$ and a_i^* is the length of the ith reciprocal cell dimension, will be reported throughout the text.

Atomic scattering factors used were those of Cromer

and Mann[6] for the non-hydrogen atoms and those of Stewart, Davidson, and Simpson[7] for the hydrogen atoms. The anomalous dispersion corrections for bromine, $\Delta f' = -.96$ and $\Delta f'' = 1.46$ were those of Cromer.[8] The $\Delta f'$ and $\Delta f''$ used for the light atoms were zero.

C. Solution of the Structure

Since there was a heavy atom, bromine, in this derivative of 1,1'-bishomocubane, solution of the phase problem for this case was rather straightforward. The symmetry-equivalent positions for the space group $P2_1/c$; (x,y,z), $(\bar{x},\bar{y},\bar{z})$, $(\bar{x}, 1/2 + y, 1/2 - z)$, $(x, 1/2 - y, 1/2 + z)$ give both a Harker plane, and a Harker line upon vector interaction in a Patterson map. The glide plane interactions between symmetry-related atoms form a line at u = 0, v, w = 1/2 in the Patterson function. The screw axis interatomic vectors fall in the plane u, 1/2, w. Since the bromine atoms are so large compared with the rest of the scattering material, bromine-bromine vectors were easy to locate in the Harker plane and Harker line. The bromine interactions occurred in the Harker plane at u = 0, v = .50, w = .25. The relationship between the Patterson coordinates and the crystal coordinates for the Harker plane at v = +1/2 is u = -2x, w = 1/2 - 2z. The x coordinate of the bromine was, therefore, zero, and the z coordinate was 1/8. In the Harker line v = .32 for the bromine interaction peaks, and since -2y + 1/2 = v, the y coordinate of the bromine was .09. The bromine was at x = 0, y = .09, z = .125. The general positions (u,v,w) in the Patterson function for $P2_1/c$ result from centrosymetrically related atoms such that u = 2x,

v = 2y, w = 2z. There were also large peaks in the Patterson function at (0, .18, 125), and thus the bromine positions deduced from the Harker line and the Harker plane agreed with the centrosymmetrically related bromine peaks.

A Fourier was calculated with the observed structure factors using the phases calculated from the position of the bromine, and the rest of the non-hydrogen atoms appeared immediately. When these 19 atoms were refined along with the bromine atom for 4 cycles of full-matrix least squares with isotropic temperature factors, and unit weights, except for zero intensities which were given zero weight, the refinement converged to a conventional R of 18.2%. The temperature factors were all positive and B ranged from 2.765 to 4.858 A^2 for the atoms involved. The difficulty was due to erroneously indexed data. Because of confusion in the choice of left and right handed coordinate systems, all the hkℓ data was really $\bar{h}k\ell$ data, and all the $\bar{h}k\ell$ data was really hkℓ data.

When the error of sign was corrected, the same structure refined in the same way to a conventional R factor of 9.6%. Since the bromine had an x coordinate of zero, the Patterson function was not sensitive to the indexing error, and since the Fourier had been properly phased, the correct structure appeared. As a check, a Fourier of Fobs was again calculated phased on the bromine positions. This Fourier gave the same

structure as the Fourier based on incorrectly indexed data.

A refinement of this structure in the same way except for the use of a weighting scheme of the form outlined in I-B with S = .05 gave an R factor of 9.2% and a standard-of-observation-unit-weight 3.534. With the same weighting scheme, and the bromine plus the two oxygens given anisotropic temperature factors, the R factor was 7.4% and the standard deviation was 2.935.

A least squares refinement as above, except all 20 atoms given anisotropic temperature factors, gave R = 7.1% and a standard deviation of 2.861. A difference Fourier was calculated using the phase relationships of the refined atomic coordinates of the 20 non-hydrogen atoms from the 7.1% refinement. From this difference map, peaks that were reasonable distances from the non-hydrogen atoms were identified as hydrogens. Hydrogen positions on the benzene ring were calculated from known geometry.

A full-matrix least squares refinement with the weighting scheme mentioned above gave after four cycles of refinement with the bromine and two oxygens anisotropic, an R factor of 6.3% and a standard deviation of 2.438. Unfortunately, the hydrogens refined to unreasonable positions, and the temperature factors on them ranged from -54.9 to 14.6.

It was reasoned that the data was not good enough to allow refinement of hydrogens. At this point the data were carefully scrutinized for mismeasurements and mis-punched

cards. The mis-punched cards were corrected, and the crystal was put back on the diffractometer for remeasurement of intensities that showed bad agreement of observed and calculated structure factors. It was found that the crystal had decomposed further after it had been taken off the x-ray diffractometer. Intensities were down by about 50% of their original values. Nevertheless, the badly agreeing data were remeasured and scaled up to the rest of the data. Most of the bad agreement was in the very weak, or the very intense data. The Fo for the weak data was systematically high, due probably to an underestimation of background. The Fo for the intense data was systematically low due perhaps to counter saturation. Remeasurement of the decomposing crystal allowed correction of the most serious blunders due to mis-set angles, but could do little to improve the bad agreement of the weak and intense data.

When the data were corrected for these mistakes, a new data tape was prepared. The weighting scheme was changed to weight down the discrepancies of the weak and very intense data. The constant S, which accounts for systematic errors in intense data, was raised to .06, and ΔI_b, which affects the weighting on the weak reflections, was increased to 20 counts.

A least-squares refinement with this weighting scheme without hydrogens and with all the non-hydrogen atoms given

anisotropic temperature factors gave, after 4 cycles, an R factor of 6.5% and a standard deviation of 2.718. The weighted R, R2, was 9.7%.

The rest of the refinements on 1,1'-bishomocubane were done with this data tape. A refinement with 3 anisotropic atoms, bromine and the 2 oxygens, the 17 non-hydrogen atoms given variable isotropic temperature factors, and the hydrogens given a constant isotropic temperature factor of 3 A^2.

gave after 4 cycles an R factor of 5.3% and a standard deviation of 1.971. The hydrogen positions still did not refine well. Some of the distances dropped to .44 A and others increased to 1.41 A.

A refinement giving all non-hydrogen atoms anisotropic temperature factors, and all hydrogen atoms isotropic temperature factors which were allowed to vary, gave an R factor after 2 cycles of 4.9%, a standard deviation of 1.932, and a weighted R, R2, of 6.6%. The temperature factors on the hydrogens were again somewhat amazing, ranging from -11 to +10. The hydrogens would not refine, and in the final refinement on 1,1'-bishomocubane the hydrogen positions or temperature factors were not allowed to vary, but structure factors were calculated based on their probable positions.

The final refinement of 1,1'-bishomocubane with all non-hydrogen atoms given anisotropic temperature factors, and the hydrogen atoms used only for the calculation of

structure factors based on their probable positions with a fixed isotropic temperature factor of $3A^2$ gave an R factor of 5.5% for 865 non-zero weight data, an R factor of 6.4% for all the 990 independent reflections, a weighted R of 7.9%, and a standard deviation of observation unit weight 2.233. On this final refinement no parameter shifted by more than 10% of its estimated standard deviation. No peak on the final difference Fourier was larger than .44 electrons.

Examination of individual agreement between observed and calculated structure factors from this refinement showed that the three most intense Fo were observed systematically too low, due most likely to counter saturation. Deletion of the 002,111,112 reflections from the least squares calculations followed by refinement exactly as above gave a conventional R factor of 5.0%, an R factor of 5.9% against all the 887 data, a weighted R of 7.4%, and a standard deviation of observation unit weight 2.07. Standard deviations of individual bond distances dropped by .001.

D. Discussion of the Structure

The final positional coordinates of the non-hydrogen atoms are given in Table I-2, and their thermal parameters are given in Table I-3. The hydrogen positions on the benzene ring were calculated from known geometry. All other hydrogen positions were determined from less accurate diffuse peaks of the difference Fourier. The hydrogen positions would not refine in least squares, but their coordinates are given in Table I-4 and may be of use. All intramolecular distances are shown in Figure I-1 along with the atomic numbering system used for the crystallography of the compound. Least squares estimated standard deviations on the bond distances ranged from .011 to .021, and since these estimates represent a minimum, the standard deviation of .02 will be taken as a reasonable estimate of the true standard deviation for the purpose of this discussion.

The average bond length of the 15 bonds in the 1,1'-bishomocubane cage is 1.55 $\pm$ .02Å. On the average a carbon-bon single bond length in this peculiar cage compound is about the same as the bond length in a free hydrocarbon cahin.[9] This result is consistent with the published values of C-C bond distances in cubane,[2] 1.555 $\pm$ .003Å, in homocubane carboxylic acid,[10] 1.56 $\pm$ .03Å, and 6,6-ethylenedioxyheptachloropentacyclo(5.2.0.$0^{2,5}$,$0^{3,9}$.$0^{4,8}$)nonane-3-carboxylic acid,[11] 1.55Å. The average of all the bonds in the benzene ring is 1.34Å, a value which is not significantly different

Table I-2 (a)

Atomic Coordinates and their Standard Deviations (a) for all Non-hydrogen Atoms in 1,1'-Bishomocubane.

ATOM	X	Y	Z
BR	.9957(2)	.09402(5)	.1308(2)
O(1)	.476(1)	.3121(3)	.2858(9)
O(2)	.277(2)	.2561(3)	.399(1)
C(1)	.338(2)	.3529(5)	.328(1)
C(2)	.464(2)	.4014(4)	.328(1)
C(3)	.343(2)	.4522(4)	.351(1)
C(4)	.139(2)	.4564(4)	.231(1)
C(5)	.115(2)	.4076(4)	.122(1)
C(6)	.132(2)	.3553(4)	.205(1)
C(7)	.515(2)	.4240(4)	.171(1)
C(8)	.469(2)	.4781(4)	.229(1)
C(9)	.273(2)	.4835(4)	.109(1)
C(10)	.310(2)	.4282(4)	.047(1)
C(11)	.439(2)	.2624(7)	.322(2)
C(12)	.578(2)	.2237(5)	.277(1)
C(13)	.741(2)	.2349(5)	.200(2)
C(14)	.862(2)	.1978(7)	.156(1)
C(15)	.821(2)	.1480(5)	.192(1)
C(16)	.665(2)	.1352(4)	.271(2)
C(17)	.537(2)	.1723(6)	.317(2)

(a) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

Table I-3

Table of Anisotropic Temperature Parameters (a) and their Standard Deviations (b) in 1,1'-Bishomocubane.

ATOM	B11	B22	B33	B12	B13	B23
BR	5.34(9)	5.4(1)	5.17(9)	1.22(6)	.81(6)	.19(6)
O(1)	4.0(4)	4.0(4)	4.6(4)	.4(4)	1.2(3)	.6(4)
O(2)	8.1(7)	2.3(4)	6.0(5)	.7(4)	.5(5)	.7(3)
C(1)	4.5(7)	4.7(7)	4.2(7)	-.1(7)	1.2(6)	-.2(5)
C(2)	3.5(6)	4.4(8)	3.8(7)	-1.3(6)	-.8(5)	.4(5)
C(3)	8.1(9)	3.5(7)	4.2(7)	-.8(7)	2.1(7)	-1.4(5)
C(4)	3.8(7)	3.9(7)	5.8(7)	1.0(5)	1.2(6)	.2(6)
C(5)	2.9(6)	3.7(7)	5.3(7)	.9(5)	.6(5)	-1.4(6)
C(6)	4.5(7)	2.7(6)	5.0(7)	-.5(5)	.5(6)	-.9(5)
C(7)	2.8(6)	3.8(7)	4.7(7)	-.4(5)	.6(5)	-.2(5)
C(8)	5.6(8)	3.5(7)	4.9(7)	-2.2(6)	1.1(6)	-.3(6)
C(9)	4.3(7)	4.2(7)	4.4(7)	-.1(5)	.4(6)	.8(5)
C(10)	4.5(7)	3.3(6)	4.9(7)	.4(5)	1.8(6)	-.5(5)
C(11)	4.5(10)	9.3(13)	3.2(7)	-4.8(9)	-.2(6)	2.7(7)
C(12)	3.6(8)	4.7(9)	4.9(7)	3.7(8)	-.5(6)	-2.2(7)
C(13)	1.7(6)	5.8(8)	7.0(9)	-.3(6)	2.5(6)	.3(6)
C(14)	2.1(7)	6.5(9)	5.5(7)	-.3(7)	2.0(5)	.4(7)
C(15)	2.2(6)	4.8(9)	3.8(7)	.3(5)	.6(5)	1.2(5)
C(16)	6.2(8)	1.9(7)	6.1(8)	1.2(6)	2.3(7)	.9(6)
C(17)	6.6(9)	3.1(7)	6.1(8)	-1.2(7)	2.3(6)	2.0(6)

(a) Anisotropic thermal parameters, $\underline{B}$, in units of A^2, are given by $\underline{B} = 4\beta_{\underline{i}\underline{i}}/\underline{a}^*_{\underline{i}}\underline{a}^*_{\underline{i}}$, where $\underline{a}^*_{\underline{i}}$ is the $\underline{i}$th reciprocal cell length.

(b) Estimated standard deviations are given in parentheses following the parameter.

Table I-4

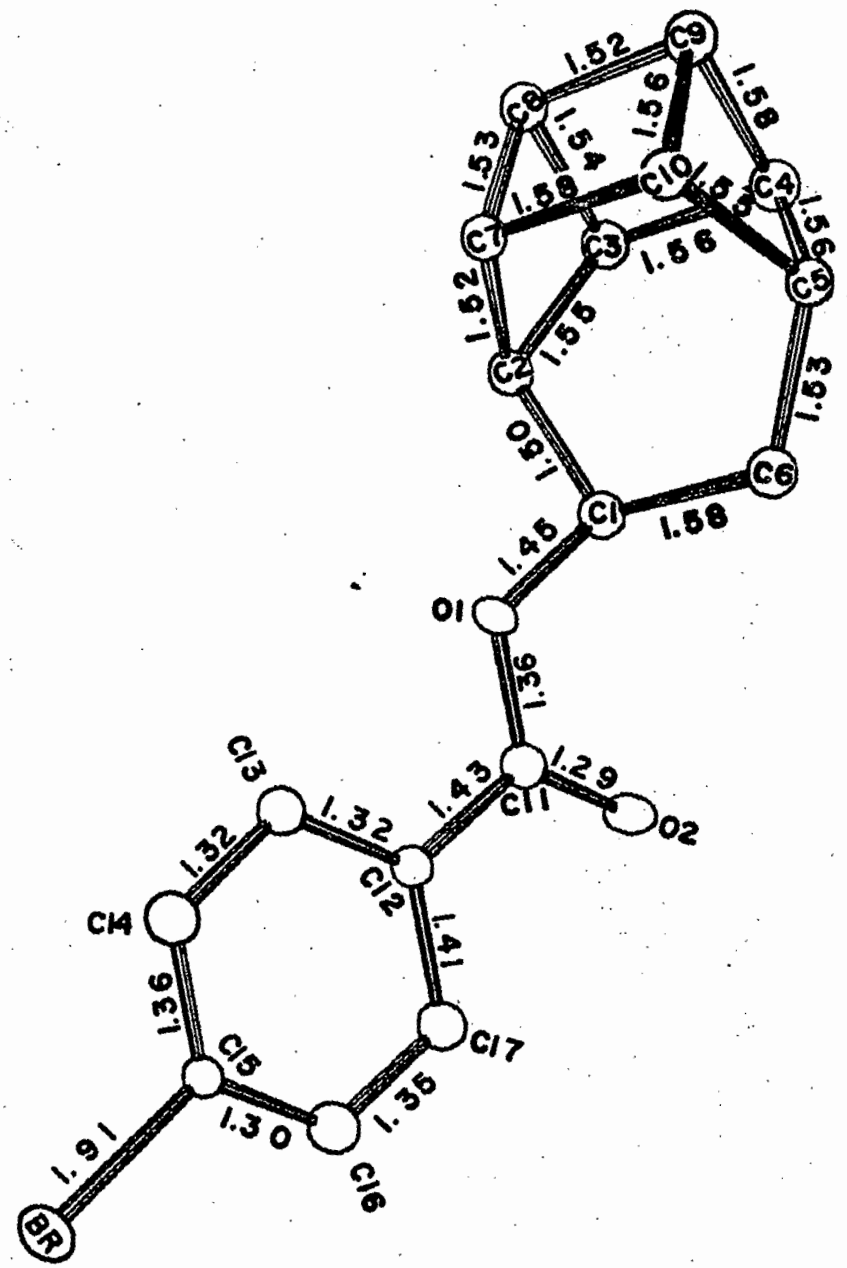
Atomic Coordinates of the Hydrogen Atoms in 1,1'-Bishomocubane.*

Atoms	X	Y	Z
H(1)	.2670	.3520	.4330
H(2)	.6100	.4050	.4000
H(3)	.3100	.4700	.4400
H(4)	0	.4700	.2500
H(5)	.9630	.4160	.0960
H(6)	.0210	.3320	.1440
H(7)	.0740	.3720	.3060
H(8)	.6400	.4100	.1200
H(9)	.5540	.5100	.2370
H(10)	.2300	.5100	.0200
H(11)	.4100	.0800	.4600
H(12)	.7868	.2741	.1720
H(13)	.9949	.2056	.0940
H(14)	.6242	.0976	.2962
H(15)	.3927	.1658	.3794

*Hydrogen coordinates are estimated from the difference Fourier; except benzene hydrogen coordinates are calculated from known geometry.

Figure I-1

Atomic Numbering System and Bond Distances in 1,1'-Bishomocubane.


C9
C8
C10
C4
C7
C3
C5
C2
C6
C1
O1
C11
O2
C13
C12
C14
C17
C15
C16
BR
1.52
1.56
1.58
1.54
1.53
1.58
1.55
1.56
1.52
1.56
1.58
1.53
1.50
1.58
1.45
1.36
1.29
1.43
1.32
1.32
1.41
1.36
1.35
1.30
1.91

XBL 704-805

Table I-5

Intramolecular Angles (in degrees) and their Standard Deviations (in parentheses) of 1,1'-Bishomocubane.

Cyclohexane

[C(1)-C(2)-C(3)-C(4)-C(5)-C(6)]

Atoms	Angles
C(1)-C(2)-C(3)	116.5(1.0)
C(2)-C(3)-C(4)	111.7(.9)
C(3)-C(4)-C(5)	109.7(1.0)
C(4)-C(5)-C(6)	117.6(1.0)
C(5)-C(6)-C(1)	109.9(.9)
C(6)-C(1)-C(2)	112.0(.9)

Cyclobutane

[C(2)-C(7)-C(8)-C(3)]

Atoms	Angles
C(2)-C(7)-C(8)	90.4(.9)
C(7)-C(8)-C(3)	86.8(.8)
C(8)-C(3)-C(2)	89.3(.9)
C(3)-C(2)-C(7)	86.4(.8)

Benzene

[C(12)-C(13)-C(14)-C(15)-C(16)-C(17)

Atoms	Angles
C(12)-C(13)-C(14)	119.7(1.6)
C(13)-C(14)-C(15)	120.0(1.6)
C(14)-C(15)-C(16)	122.4(1.6)
C(15)-C(16)-C(17)	119.2(1.5)
C(16)-C(17)-C(12)	118.2(1.6)
C(17)-C(12)-C(13)	120.5(1.6)

Cyclobutane

[C(4)-C(5)-C(10)-C(9)]

Atoms	Angles
C(4)-C(5)-C(10)	86.3(.8)
C(5)-C(10)-C(9)	91.3(.8)
C(10)-C(9)-C(4)	85.4(.8)
C(9)-C(4)-C(5)	90.2(.8)

Cyclobutane

[C(3)-C(8)-C(9)-C(4)]

Atoms	Angles
C(3)-C(8)-C(9)	91.8(1.0)
C(8)-C(9)-C(4)	89.7(.9)
C(9)-C(4)-C(3)	88.6(.9)
C(4)-C(3)-C(8)	89.9(.9)

Cyclobutane

[C(7)-C(8)-C(9)-C(10)]

Atoms	Angles
C(7)-C(8)-C(9)	92.6(.9)
C(8)-C(9)-C(10)	89.7(.9)
C(9)-C(10)-C(7)	89.2(.8)
C(10)-C(7)-C(8)	88.5(.9)

Bromine

Atoms	Angles
Br(1)-C(15)-C(14)	120.4(.6)
Br(1)-C(15)-C(16)	117.2(.6)

Carbonyl

Atoms	Angles
C(12)-C(11)-O(2)	127.3(1.2)
C(12)-C(11)-O(1)	118.8(1.2)
C(11)-O(1)-C(1)	121.2(1.2)
O(1)-C(1)-C(2)	106.2(.8)
O(1)-C(1)-C(6)	110.5(.7)

within the standard deviations from the established value for the C-C double bond distance. All the other bond distances agree within the standard deviations with the well-established values published in the International Tables.[9]

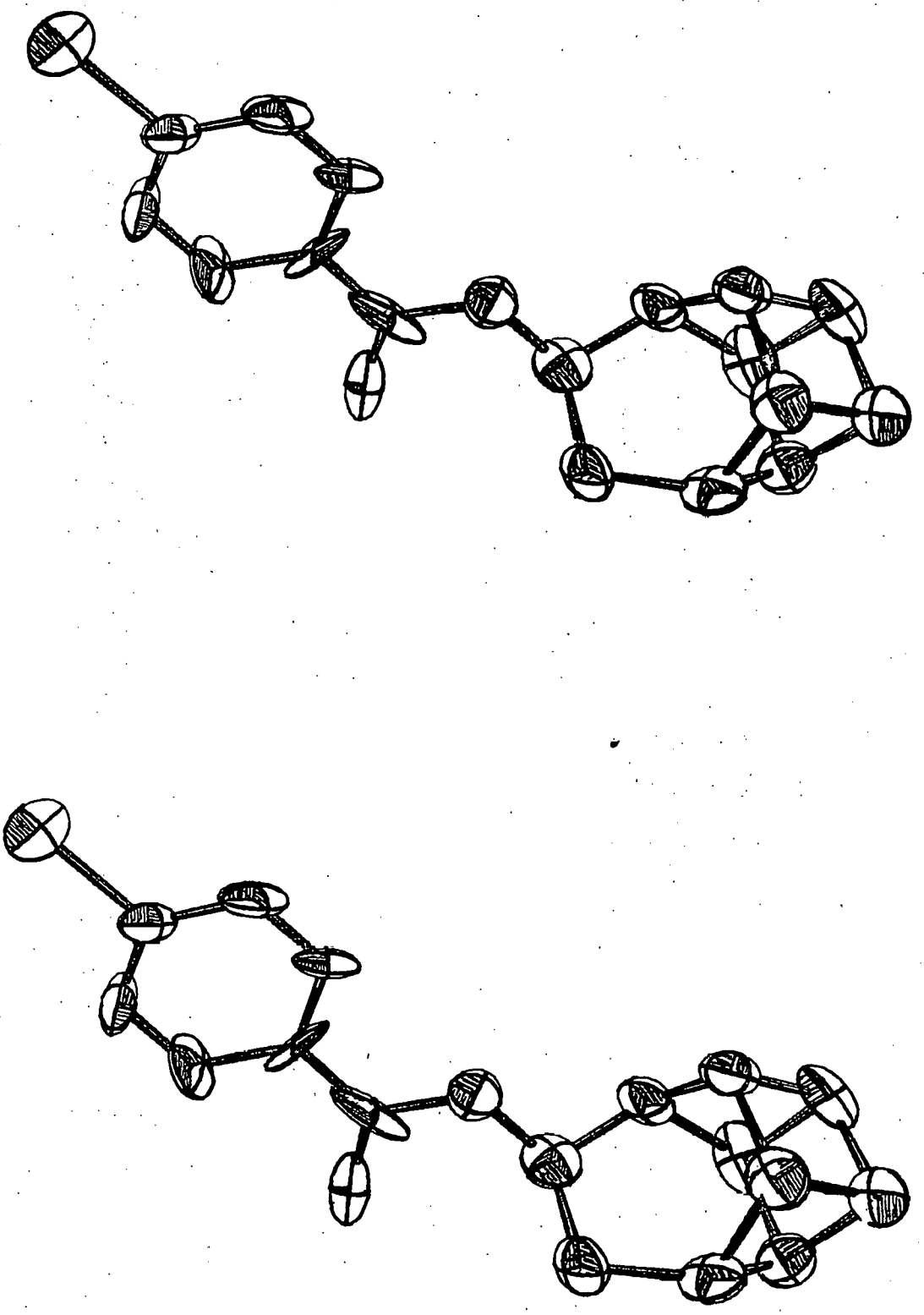
All the thermal parameters of 1,1'-bishomocubane are given in Table I-3. An ORTEP[12] thermal ellipsoid stereo representation of thermal motion is shown in Figure I-2.

The most interesting part of this structure is the angular configurations of the carbons in the rings of the cage. Intramolecular angles for the hydrocarbon rings, and all the atoms of the structure are given along with their standard deviations in Table I-5. Reference to the drawing of the compound in Figure I-1 will be helpful in interpreting Table I-5.

The average of the four angles in the cyclobutane rings is: (3-8-9-4) 90.0°, (7-8-9-10) 90.0°, (2-7-8-3) 88.2°, and (4-5-10-9) 88.3°. The average angles of the cyclobutane rings in cubane are all 90° and the cyclobutane rings are all planar.[2] Least squares planes through cyclobutane rings (3-8-9-4) and (7-8-9-10) show these two rings to be planar within the standard deviations of the structure determination. Deviations of the atoms in the various rings of 1,1'-bishomocubane from least squares planes drawn through the rings are given in Table I-6. The angles 10-9-4 and 4-5-10 of cyclobutane ring (4-5-10-9) are significantly smaller than 90°.

Figure I-2

Anisotropic Thermal Motion in 1,1'-Bishomocubane.



XBL 704-830

Table I-6

Deviation of Atoms in Angstroms from Least Squares Planes.

Atoms	C(7)	C(8)	C(9)	C(10)
Distance	.005	-.006	.006	-.005
Atoms	C(3)	C(4)	C(8)	C(9)
Distance	.004	-.004	-.004	.004
Atoms	C(2)	C(3)	C(4)	C(5)
Distance	-.000	.000	.000	.000
Atoms	C(2)	C(1)	C(6)	C(5)
Distance	.007	-.004	.004	-.007
Atoms	C(2)	C(3)	C(8)	C(7)
Distance	.134	-.134	.135	-.135
Atoms	C(4)	C(5)	C(9)	C(10)
Distance	.133	-.136	-.133	.136

This constriction in the cyclobutane ring results in a deviation from planarity and a puckered ring. The same is true for angles 7-8-3 and 2-3-7 in cyclobutane ring (2-7-8-3). Cyclobutane ring (2-7-8-3) is puckered such that the dihedral angle between planes through 7-8-3 and 7-2-3 is 27.9°. Similarly, the cyclobutane ring (5-10-9-4) is puckered, the carbon C(5) is pushed out of the plane, and the angle between planes 4-9-10 and 4-5-10 is 27.2°. Within the standard deviations of the structure determination, these two rings are puckered the same amount.

The cyclohexane ring in 1,1'-bishomocubane is flattened compared to its configuration in an unconstrained cyclohexane ring. The average for the six angles of the cyclohexane ring is 112.9°. This average is significantly larger than the tetrahedral angle of the unconstrained ring. Most of the cyclohexane ring distortion is about carbons C(5) and C(2). The angle 4-5-6 is 117.6° and the angle 1-2-3 is 116.5°. Carbons 2-3-4-5 of the cyclohexane ring are coplanar as we see in Table I-6. Carbons 5-6-2-1 of the cyclohexane ring are also coplanar within the standard deviations, and these two planar portions of the cyclohexane are folded toward each other at an angle of 131.7°. The sections of the ring are within 48.3° of a coplanar configuration.

The planar cyclobutane ring (7-8-9-10) is almost parallel with the planar portion of the cyclohexane ring 2-3-4-5.

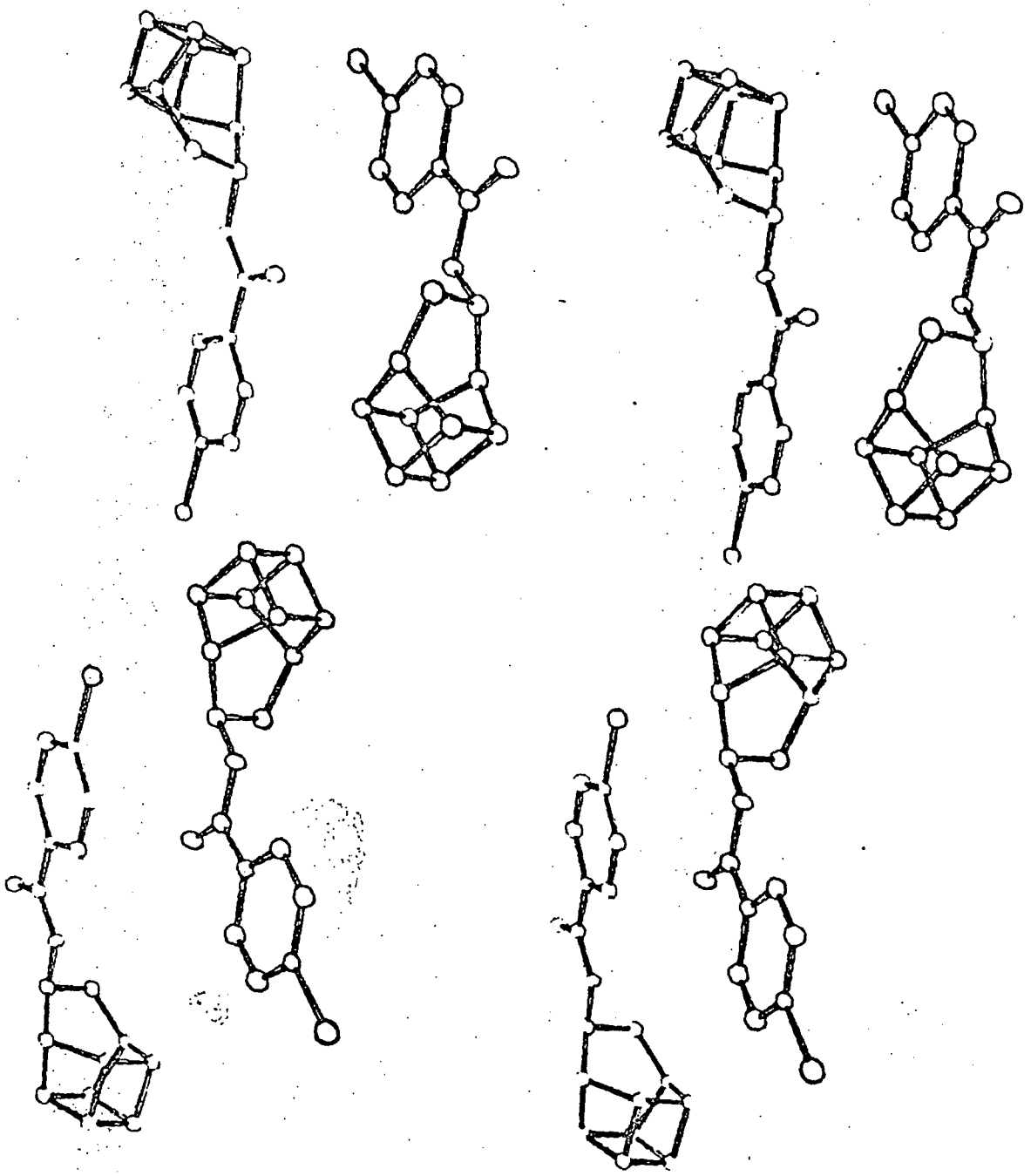
The dihedral angle between these two planes is 4.6°. This cyclobutane ring (7-8-9-10) is almost perpendicular to the planar cyclobutane ring (3-4-9-8), and the dihedral angle between the two planes is 86.1°. The planar portion of the cyclohexane ring (2-3-4-5) is also nearly perpendicular to cyclobutane ring (3-4-9-8) and the angle is 89.5°.

A 3-dimensional representation of the intermolecular packing in 1,1'-bishomocubane is shown in Figure I-3. The closest intermolecular approach not involving hydrogens was 3.53Å between C(11) and C(13). Other close symmetry-related approaches between the molecules of the unit cell were C(12)-O(2) 3.56Å, C(9)-C(8) 3.60Å, C(11)-O(2) 3.63Å, and C(3)-Br(1) 3.63Å.

A list of observed and calculated structure factors for this compound is given in Table I-7.

Figure I-3

Intermolecular Packing in 1,1'-Bishomocubane



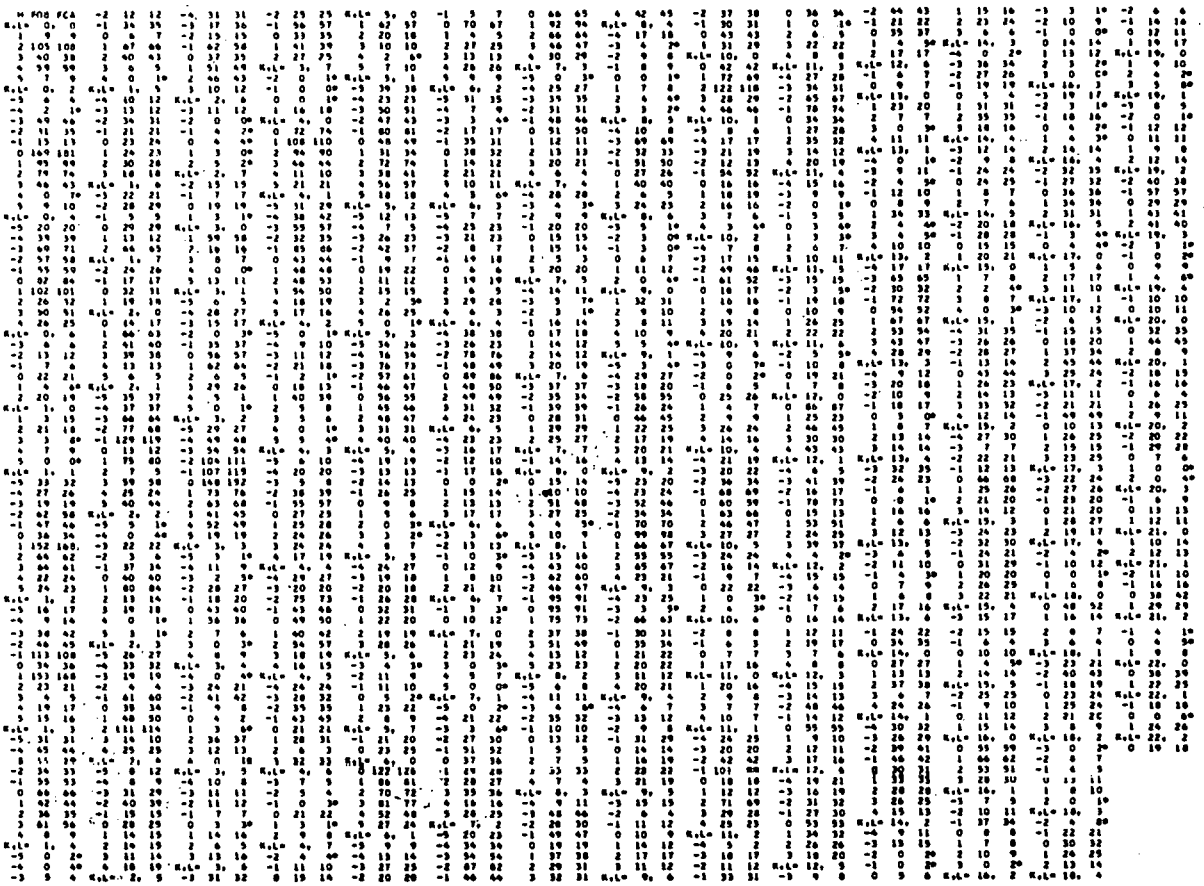
XBL 704 - 835

Table I-7

Observed and Calculated Structure Factors for 1,1'-Bishomocubane.

TABLE OF OBSERVED AND CALCULATED STRUCTURE FACTORS FOR 1,3' - DIBENZOTHIAZOLE

FCA(0, 0, 0) = 668



XBL 704-758

E. Conclusion

The carbon-carbon distances in all the rings of the hydrocarbon cage of 1,1'-bishomocubane are, on the average, no different from the carbon-carbon single bond distance expected for the unconstrained rings. A similar result was found for cubane.

All the cyclobutane rings in cubane were planar and mutually orthogonal. Addition of two extra atoms into one of the cyclobutane rings causes two of the remaining four cyclobutane rings to pucker from a planar configuration by about 27°. The other two cyclobutane rings remain planar and mutually orthogonal. The consecutive atoms 2-3-4-5 of the cyclohexane ring formed by this addition of two extra atoms into a cubane ring is planar and perpendicular to the cyclobutane ring (3-4-8-9).

Constraint within the cage causes the cyclohexane ring to flatten. The average of all the angles in the cyclohexane ring is 112.9°. The two planar portions of the cyclohexane ring are folded toward each other and the angle of separation is 131.7°.

Bibiography

1. P. E. Eaton, and T. W. Cole, J. Am. Chem. Soc., 86, 3157 (1964).
2. Everly B. Fleischer, J. Am. Chem. Soc., 86, 3889 (1964).
3. Dale L. Whalen, University of California, Berkeley, Ph.D. Thesis, (1966).
4. J. Meinwald, and J, K. Crandall, J. Am. Chem. Soc., 88, 1292 (1966).
5. Dale L. Whalen, and W. B. Dauben, Tetrahedron Letters, 31, 3743 (1966).
6. D. T. Cromer and J. B. Mann, Acta Cryst., A24, 321 (1968).
7. R. F. Stewart, E. R. Davidson, and W. T. Simpson, J. Chem. Phys., 42, 3175 (1965).
8. D, T. Cromer, Acta Cryst., 18, 17 (1965).
9. International Tables for X-ray Crystallography, 3, 276, Kynoch Press, Birmingham, England, 1962.
10. Roger Petterson, University of California Berkeley PhD. Thesis, UCRL-16975, 1966.
11. Y. Okaya, Acta Cryst., B25, 1882 (1969).
12. C. K. Johnson, "Ortep," A Fortran Thermal Ellipsoidal Plot Program for Crystal-Structure Illustrations, ORNEL-3794, Revised, Oak Ridge, Tennessee, 1965.

Section II.

A COMPUTER PROGRAM FOR THE SOLUTION OF CRYSTAL STRUCTURES CONTAINING A PLANAR MOIETY WITH KNOWN GEOMETRY.

A. Introduction

In this section there is outlined the use and logic of a computer program which can solve the crystal structure of a predominantly planar molecule by trial and error. The development of this program follows a path of logic which has its basis in practical crystallography. A useful starting point for the structure determination of a planar molecule by trial and error can be the juxtaposition of the molecular plane with a given Bragg reflection plane. This juxtaposition is a useful starting point because a planar molecule gives a Patterson function for which all the interatomic vectors within the molecule lie on a disk.[1] The orientation of this disk in vector space is the same as the Bragg orientation of the planar moiety of the molecule under consideration.[2] The orientation of the molecular plane can be extracted from the Patterson function. For example, if the disk emanates from the origin in vector space at (0,0,0), and remains parallel with the a, c plane, the probable orientation of the planar molecule is coincident with the 010 diffraction plane.

The program is extremely useful if the orientation of the molecular plane can be deduced, and if the number of possible positions of the molecules within the unit cell can

be limited. In competition with other methods, a trial and error program of this type is most successful for the solution of structures in non-centric space groups. Non-heavy atom problems in non-centric space groups can be solved by direct methods only through excessive use of computer time.

A trial and error program of the type outlined in this section works best in exactly these cases because the number of parameters is limited. Use of this program is outlined in part B of this section, and analysis of cases from the literature for which one could expect it to work well is given in part C. Use of the program to solve the crystal structure of 5-methoxytryptamine is given in Section III, and a Fortran listing along with the derivation of the orientation matrix is given in the Appendix. In its present form the program works for space groups in monoclinic or higher symmetry.

B. Use of the Program

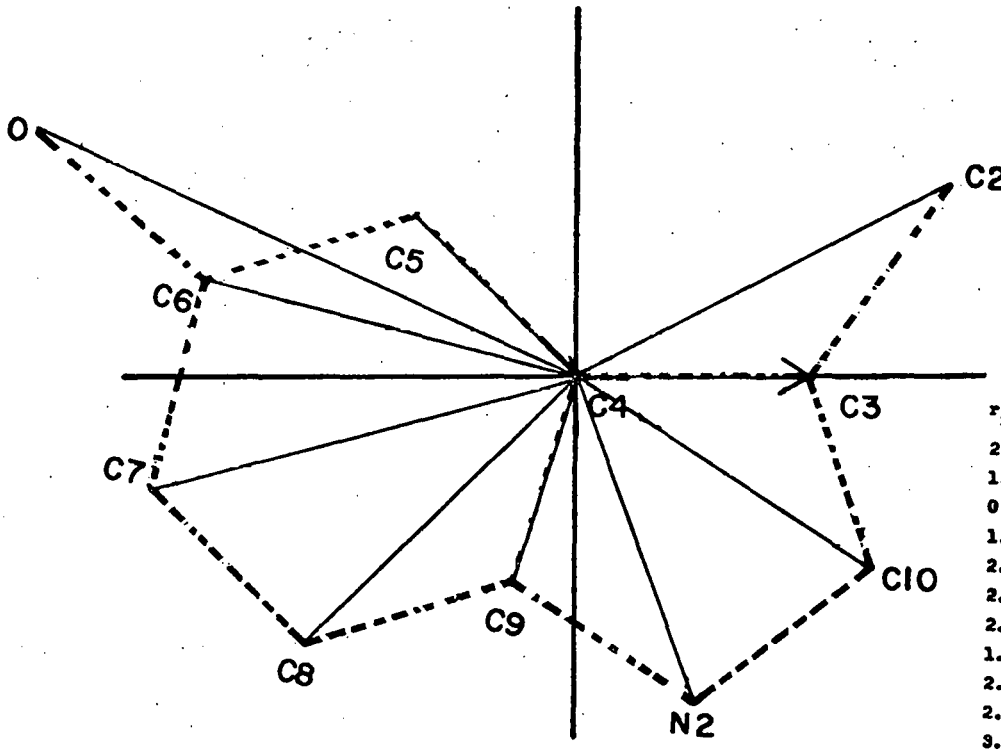
The use of this program is quite simple. All the necessary work involved in setting up the program for a specific case, and a brief outline of the logic is set forth here to aid those who wish to use it.

The first step is to look at your structure and choose one of your atoms as an origin. The atom should be rather centrally located. With this atom as origin, the part of the structure having known geometry is graphed onto polar coordinate paper. Then, choose another atom of the structure such that a vector drawn from the origin to this second atom defines simultaneously the positive abscissa of an orthogonal coordinate system and the zero angle of the polar coordinate system. The position of the Nth atom of the structure can now be described by its distance in Angstroms from the origin, r_n, and the angle θ_n between the positive abscissa and the vector $\underline{R}_n$ from the origin to this Nth atom. These two variables, r_n and θ_n, for each atom of your structure are input into the program on data card 6. An example of this graph for 5-methoxytryptamine is shown in Figure II-1.[3]

The program transforms each vector $\underline{R}_n$ first into an orthonormal basis set $\underline{e}_j$ such that $\underline{R}_n = \Sigma v_j \underline{e}_j$ $(j = 1,2)$ where $v_1 = r_n \cos\theta_n$ and $v_2 = r_n \sin\theta_n$.

Figure II-1

Graphical determination of r_n and θ_n for 5-methoxytryptamine. Approximate geometry of the indole ring is from the crystal structure of indole acetic acid.[3]



r_n(Å)	θ_n(deg.)	n
2.64	28	2
1.47	0	3
0	0	4
1.43	134	5
2.41	165	6
2.77	195	7
2.46	225	8
1.41	254	9
2.25	326	10
2.26	290	11≡12
3.75	155	12≡0(1)

XBL 704-852

The vector $\underline{R}_n$ is then represented as a linear combination of unitary monoclinic base vectors such that $\underline{R}_n = \Sigma u_i \underline{\nu}_i$ (i=1,3) and $u_i = \Sigma \underline{a}_{ij}\ v_j$ (i = 1,3; j = 1,2).

The components of $\underline{R}_n$, u_i, in this system are further transformed into crystallographic coordinates, x_i, with division by the cell dimensions, and $x_i = u_i/|\underline{b}_i|$, where $\underline{b}_i$ is the ith direct space cell dimension.

The second step in the use of this program is to decide with which Bragg diffraction plane your molecule is coincident. In general, Bragg planes can be described by the numbers hkl, which are related to the direction cosines of the angles which the Bragg plane normal makes with the crystal co-ordinate axes. The orientation matrix, $\underline{a}_{ij}$, is a function of these direction cosines, the angle β of the monoclinic axis system, and the cell dimensions of the crystal.

The orientation, hkl, of the molecular plane may be deduced from the Patterson function. Also, the orientation of the molecular plane may correspond to a high Ehkl coefficient.[4]

The derivation of the orientation matrix, $\underline{a}_{ij}$ is given in the Appendix to aid those who might want to extend the application of this program to the triclinic case or alter

this matrix in any way. The matrix $\underline{a}_{ij}$ is produced on a punched card in the proper format by program Norma. The input necessary for the production of this matrix is the hkl of the Bragg plane into which you propose to place your molecule and the cell dimensions of your crystal. Trial and error can be done in up to 50 Bragg plane orientations simultaneously. A Fortran listing of program Norma as well as the trial and error program Omoo is given in the Appendix. The orientation matrices $\underline{a}_{ij}$ for every case you wish to try are loaded as cards 4 in the trial and error program.

The observed structure factors, Fo, for your crystal are input as data cards 5. The format is in the program listing. These cards may be produced most easily by having your data tape punched directly into cards by the computer. Selection of data is the most critical factor in the trial and error approach. Any selection of data should have a large number of low angle Fobs as part of the data set. Care must be taken that all data is free from errors, both random and systematic. Extinction was a problem with the first data choice in the solution of the structure of 5-methoxytryptamine, as related in Section III. Low angle data should be present, as the low angle structure factors are less sensitive to the exact atomic positions of the atoms involved, and thus

give more lee-way for error in the approximate positions in the R factor test.[4]

Another possible means of selecting a sample of the data for the trial and error program is to choose Fo corresponding to high Ehkl values. One would expect more structural information pertaining to the position of a molecule plane in these numbers as high Ehkl values correlate with a number of atoms scattering in phase.[4]

Yet another way to supplement the low angle data, which are the backbone of this trial and error procedure, is to pick data cards at random out of the deck of observed structure factors. This tends to reduce the effects of possible systematic errors introduced by the other two methods. At present, up to 200 data can be loaded on cards 5.

Besides r_n and θ_n, which are estimated graphically; the orientation matrix, $\underline{a}_{ij}$, which is produced from crystal cell dimensions and a choice of the orientation numbers hkl by program Norma; and a choice of up to 200 observed structure factors, F_o; the program requires the average scattering factor for your compound as a function of $\sin\theta/\lambda$. The average scattering factor is used in the calculation of calculated structure factors, $|F_c|$. The form of the structure factor used by the program is

$$|F_c| = \hat{f}\cdot\hat{t}[(\sum_r A_r)^2 + (\sum_r B_r)^2]^{1/2}$$

where F_c is the magnitude of the calculated structure factor, $\hat{f}$ is the average scattering factor for the r atoms of the unit cell, $\hat{t}$ is an overall average isotropic temperature factor for the r atoms, and A_r and B_r are periodic functions of the position of the rth atom in the unit cell. The forms of A and B depends upon the space group and choice of origin, and the specific form of A and B for the problem at hand must be written into Fortran by the user. These coefficients are listed for all 230 space groups in the International Tables for X-ray Crystallography, Volume I.[5] For $P2_1/c$, B = 0 and

$$A_r = 4\cos 2\pi(h\, x_r + l z_r + (\tfrac{k+l}{4}))\, \cos 2\pi(k y_r - (\tfrac{k+l}{4})).$$

Specifically $\hat{t} = \exp(-B \cdot \sin^2\theta/\lambda^2)$ where B is a constant. Unless otherwise indicated, the program will use a value of $B = 3A^2$ for the overall thermal parameter. The average scattering factor is defined as[6]

$$\hat{f} = \sum_n f_n / N$$

where f_n is the scattering factor of the nth atom in the unit cell and N is the total number of atoms. The calculated structure factor $|F_c|$ is produced for each of the observed structure factors, F_o, on data cards 5, and a linear scale factor k is calculated such that

$$k = \Sigma|F_c| \,/\Sigma F_o$$

All possible orientations of a planar molecule in a crystallographic unit cell can be produced by various combinations of Bragg orientations set by the matrix a_{ij}, iterative choices of the positional parameters (x,y,z) of one atoms in the unit cell, and molecular rotations about the Bragg plane normal which passes through the each chosen molecular origin (x,y,z). The program has a nest of 5 Do loops which controls this iteration. For each trial orientation of the molecular plane, the program calculates atomic coordinates for the input structure in the proper crystallographic cell, calculates structure factors, scales the F_o to the $|F_c|$, and prints out the conventional R factor for the orientation involved where

$$R = \Sigma||kF_o| - |F_c|| \,/\Sigma|kF_o|$$

For 11 atoms and 100 data in space group P_c, the program will produce crystallographic coordinates and print an R factor in .20 seconds of CDC-6600 time. This makes it possible to try about 300 different structural possibilities in about one minute of computer time.

The speed with which the program runs is its main advantage. Other advantages are: it uses less than a 50,000 word memory, thus allowing rapid turnover in computer systems designed to give priority to jobs that use very little

memory; it loads no tapes and can be used without operator intervention; and the speed with which it runs is almost independent of the number of atoms. The number of atoms occurs only once in a DO loop involving two algebraic terms that are CDC-6600 library subprograms.

The principles necessary for the use of this program have now been outlined. Exact input information is available in the Fortran listing in the Appendix.

C. Usefulness of the program in general

This program will work for space groups of monoclinic or higher symmetry, and for molecules that have a moiety with a fixed, known geometry. Many organic compounds meet this criterion. For example, fused aromatic ring systems such as the indole system, benzene rings, ligand groups attached to a central metal atom, steroids, terpenes, porphyrins, and even DNA base pairs.[3]

In general the program reduces a known geometrical system of 3N positional variables, where N is the number of atoms, to 5 trial and error variables. These 5 variables are: (x,y,z), three crystallographic coordinates needed to specify the position of one atom of the molecule in the unit cell, the orientation of the normal to the Bragg plane with which the molecule is parallel, represented by the matrix $\underline{a}_{ij}$, and the rotation angle, omega, about an axis perpendicular to the Bragg plane which goes through the origin (x,y,z).

The success of any trial and error method depends on how successful one is in limiting the number of independent parameters. The number of trials can be somewhat limited with the use of low angle data for the R factor test. This has the effect of decreasing the atomic resolution and allowing a rather large grid size on the trial aand error variables. A resolution of .50A in (x,y,z), 10 degrees in the Bragg

orientation, and 15 degrees in the angle omega still permits a solution if the data are chosen carefully.

When all 5 independent trial and error variables must be blindly varied, the program uses an unreasonable amount of computer time. There are many cases, however, where some of the parameters can be eliminated or the range of search over them reduced.

The crystal structure of N-methyl-4-phenylisoxazolin-5-one is a good example of how crystallographic and chemical restrictions can limit the position of a molecule in a unit cell to the point that solution is possible by trial and error.[7] The compound crystallizes in the monoclinic space group $P2_1/c$. There are 2 molecules in the asymmetric unit, and Z = 8, a = 13.716, b = 10.925, β = 91.51. The presence of a strong 400 diffraction spot led the authors to deduce that the molecule was planar, that the planes of the molecules were in the 400 planes, and thus were spaced 1/4 a apart. Furthermore, since the space group was centric, x coordinates were deduced immediately at 1/8 and -1/8 for the molecular planes. The trial and error program can be applied to a molecule whenever the position of one atom in the unit cell can be limited. For $P2_1/c$ in general $(1/4)25^3$ trials would be necessary to locate the position of one atom within the unit cell if the other orientation parameters were known. For this case only $(1/4)25^2$ trials would be

needed although it is not a polar space group. If the center of gravity of the 5-membered ring is taken as an origin in the orthonormal coordinate system of the program, the omega angle need be rotated only through 180 degrees to find the rough ring orientation. Because the ring is substituted, this search would be unable to distinguish whether the true orientation were omega or 180 degrees + omega. This difficulty could be cleared up immediately once the position of the center of gravity of the ring were known in the unit cell. This structure could be solved with this program in $(1/4)(25^2)(11)$ computer trials, or 6 minutes of CDC-6600 computer time.

One rather common case occurs where the position of one atom in the unit cell can be deduced. If the structure to be solved has an atom large enough that its position in the unit cell can be located in the Patterson function, three of the trial parameters, (x,y,z) are eliminated. Furthermore, if the molecule is mostly planar such that the rough orientation of the molecular plane is known, only 4.4 seconds of computer time would be needed to solve the structure. The computer program tests structures at the rate of one per .20 seconds, or 300 per minute.

There are many space groups where the position of one of the atoms in the unit cell is limited by space group symmetry. There are sixty-seven polar space groups with

monoclinic or higher symmetry.[5] If a compound crystallizes in one of these space groups, the position of one atom of the structure can be specified by two or even one parameter rather than the usual 3, (x,y,z).

In competition with other methods, this program will in general be most useful for the non-centric, polar space groups. Programs which apply statistical methods to these cases tend to use quite a bit of computer time.

For a singly polar space group the origin can be specified with 25^2 trials if the whole cell must be scanned. This is rarely the case. For example, the rather common polar monoclinic space group P2 has equivalent positions (x,y,z) and (-x,y,-z). If the cell is searched from x = 0 to x = 1/2, z = 0 to z = 1/2, and y = 0 to y = 1.0, all the independent parts of the cell would be scanned. Only (1/4) x (25^2) trials would be needed to find the origin in the unit cell of one of the atoms being tested. Another rather common case, the orthorhombic space group $Pna2_1$, has equivalent positions (x,y,z), (-x,-y,1/2+z), (1/2-x,1/2+y,1/2+z), (1/2+x,1/2-y,z). The cell must be searched from x = 0 to x = 1; from y = 0 to y = 1/2; from z = 0 to z = 1/2. Only 1/4 of the unit cell must be search to find the origin of one of the atoms within the molecule. This means again (1/4) x (25^2) trials. If a predominantly planar molecule crystallizes in one of these space groups, in general only

1/4 to 1/2 of a unit cell contains all the independent parts of the structure in these polar space groups. Sometimes more and sometimes less of the unit cell must be searched, but (1/4) to (1/2) is rather representative. For a representative polar space group in which a Bragg plane orientation can be deduced for the molecule involved, between $(1/4)(25^2)x$ (22) and $(1/2)(25^2)$ x (22) trials must be calculated. Between 11.5 and 23 minutes of computer time would be used.

For a singly polar space group the method is unfeasible if a Bragg normal direction cannot be approximately deduced. These calculations again represent general upper limits for any specific case where packing considerations impose additional restrictions on where a molecule can be in a unit cell. For example, the crystal and molecular structure of 1:2,5:6-dibenzanthraquinone is an example of a structure in a polar space group that could have been solved quickly with this program.[8] The molecule crystallizes in the singly polar orthorhombic space group $Pca2_1$ $\underline{a}$ = 28.54, $\underline{b}$ = 3.85, $\underline{c}$ = 12.90, z = 4. There is no heavy atom. One of the planar napthalene residues could have been graphed into an orthonormal coordinate system. Since the $\underline{b}$ axis is so short, we know that the molecule must be nearly in the x,y plane. This orientation can be described by the normal to the 010 Bragg planes. In this polar space group the z coordinate of one of the atoms within the unit cell is arbitrary.

Only 1/4 of the cell need be searched. If we choose an origin at the center of the napthalene ring, we have to rotate through 180° in the angle omega. The orientation of this substituted napthalene could have been found with sufficient accuracy to phase a Fourier with $(25^2)(1/4)(11)$ trials. This number of trials amounts to about 6 minutes of CDC-6600 computer time.

Another polar space group case for which we would expect the program to be successful is (+)-(R)-N-Methyl-1-((naphthyl)ethyl)-(R)-O-methylmandelamide.[9] This structure does not have a heavy atom, and crystallizes in the non-centric polar space group $P2_1$. It was solved originally by successive application of the tangent formula.

Use of this trial and error program is especially attractive for cases such as these. Statistical methods applied to non-centric cases tend to take a lot of computer time. For non-centric cases where the number of parameters can be limited, solution of the problem by trial and error can be faster. In this case, the napthalene ring is planar and can be easily graphed into a Cartesian coordinate system since its geometry is known. With good data and a sharp Patterson function, it would be no problem to find the approximate Bragg orientation of the napthalene ring. The napthalene ring is graphed into an orthonormal set such that its center of gravity is at the origin of the orthonormal

system, only 180 degrees of rotation in the angle omega about the Bragg normal will be sufficient to include all orientations of the napthalene ring for any fixed position of this origin in the unit cell. For the space group $P2_1$ only 1/2 of the unit cell needs to be searched to include all parts of the asymmetric unit. Since the space group is polar, the z coordinate of the napthalene ring center of gravity within the unit cell is arbitrary. Only $(1/2)(25^2) \times (11)$ computer trials, or about 12 minutes of computer time would be necessary to find the position of the napthalene ring. Fouriers could then be phased using the location of the napthalene ring to finish the structure.

In a doubly polar space group this program is useful indeed. For these space groups only one crystallographic coordinate is needed to specify the position of one atom within the unit cell. In these space groups the trial and error approach with this program is possible even if the molecule involved is not planar.

The 5-methoxytryptamine molecule crystallizes in the doubly polar space group P_c. A description of the solution of the structure using this program is given in Section III of this thesis.

D. Conclusion

The trial and error program discussed in this section can be a very powerful tool for the solution of crystal structures. Predominantly planar molecules of all kinds can be solved with this program, and there are numerous examples of this kind of problem in the crystallography of organic molecules.

This program will be most useful in general for structures that crystallize in non-centric space groups. Direct methods or Patterson superposition function methods consume a lot of computing time in these space groups. Trial and error works best for these cases because the number of parameters can be limited by space group symmetry.

Types of problems for which the program works best were discussed in II-C, and a description of its role in the solution of the crystal structure of 5-methoxytryptamine is given in Section III.

Bibliography Section II

1. Martin Buerger, "Vector Space," John Wiley, New York, 1959.
2. P. Tollin, W. Cochran, Acta Cryst., 17, 1322 (1964).
3. Isabella L. Karle, Kathleen Britts, and P. Gum, Acta Cryst., 17, 496 (1964).
4. George H. Stout, Lyle H. Jensen, "X-ray Structure Determination," Collier-Macmillan Limited, London, 1968.
5. "International Tables for X-ray Crystallography," I, Kynoch Press, Birmingham, England, 1952.
6. D. Harker, and J. S. Kasper, Acta Cryst., 1, 70 (1948).
7. C. Sabelli and P. F. Zanazzi, Acta Cryst., B25, 182 (1969).
8. R. F. Entwistle, J. Iball, W. D. S. Motherwell, and B. P. Thompson, Acta Cryst., B25, 770 (1969).
9. Michael G. B. Drew, Acta Cryst., B25, 1320 (1969).

Section III

The Crystal and Molecular Structure of 5-Methoxytryptamine.

A. Introduction

The scientific interest in serotonin and its structural analogues since the original isolation of serotonin from clotted blood,[1] has been almost as widespread as the natural occurrence of these compounds. Serotonin may be found in such diverse places as tomatoes, bananas, toad skins, lizard eyes, and human gut, blood, spleen, and brain. Serotonin is concentrated in the mesencephalon, the hypothalmus, and the pineal gland of the human brain. Both serotonin (5-hydroxytryptamine), and the subject of the present study, 5-methoxytryptamine, are enzymatically synthesized from tryptophane. Serotonin and 5-methoxytryptamine have a similar chemical structure and undergo many of the same reactions, such as destruction of the unprotected amines by monoamine oxidase. They are both smooth muscle contractors and vasoconstrictors. They have been shown to inhibit nerve transmission in the adrenergic nervous system. Serotonin is also implicated as a chemical transmitter of nerve pulse through the serotonergic nerves of the mesencaphalon. It has been implied that serotonin is important in the regulation of body temperature.[2,3]

The biological activity of serotonin and 5-methoxytryptamine does not end with these physiological effects. Serotonin and 5-methoxytryptamine are well-established radiation protection agents in mammals. The nature of this protection

against lethal doses of x-ray and gamma radiations is not understood.[4] The way in which these two substances absorb and transmit high energy photons in solution may be different. Serotonin does not quench the scintillation of dioxane solutions of 2,5-diphenyloxazole excited by Cs^{137}, but 5-methoxytryptamine does.[5]

Serotonin is implicated in basic behavioral phenomena associated with the brain. The turn-over rate of serotonin is correlated with the hallucinogenic activity of LSD.[6] Suppression of serotonin synthesis with p-chlorophenylalanine causes heightened sex drive, increased copulation, and bizarre social activity in rats and rabbits.[7] Serotonin may play a central role in the regulation of sleep.[8]

The scope of the real and alleged biological functions of serotonin and 5-methoxytryptamine is beyond the limitations of this paper. Reviews of various aspects of the biological functions of serotonin and 5-methoxytryptamine are listed in the bibliography.[9,10,11,12]

The great biological importance of these molecules and the prevalence of indole ring compounds in so many pathways of biochemical metabolism motivated this study of the exact x-ray molecular structure of 5-methoxytryptamine. This exact structure should aid quantum mechanical calculations of the charge distribution on the substituted indole ring, and give clues to the structural nature of the enzyme-substrate inter-

actions of the indole alkyl amines.

B. Experimental

Crystals of 5-methoxytryptamine are commercially available from the Regis Chemical Company of Chicago. A sample of this commercial product given to us by Dr. W. B. Quay of the physiology department here was used for all crystallographic work published herein.

The 5-methoxytryptamine crystals are well-formed, thin, transparent, colorless plates which are rather stable, but slowly turn brown in air over a period of a month. This color-change is independent of x-ray exposure and may be due to oxidation of the primary amine group. The color-change correlates with crystalline decomposition as measured by loss of intensity of diffracted copper x-rays. Exposure to x-rays hastens the color-change and decomposition, but the diffraction standards showed variations of less than one standard deviation throughout a three-day period of data taking. After three days the standards dropped steadily. Crystals stored in a refrigerated dessicator with drierite remained unaffected.

Preliminary oscillation and Weissenberg photographs of a large unmeasured colorless elongated plate of 5-methoxytryptamine showed space group extinctions and symmetry consistent with either monoclinic space group P2/c, Z=4; or Pc, Z=2. Density calculated from the rough film cell dimensions was a reasonable value of about 1.20g/cc^3 for Z=2, and

a space group assignment of Pc was made.

A small crystal (.22 x .18 x .09mm) was glued onto the tip of a thin glass fiber with General Electric 1202 Clear Industrial Glyptal Varnish. The fiber direction was perpendicular to the plate of the crystal, and parallel to the crystalline b axis. After one oscillation photograph to check optical alignment, the crystal was transferred to a General Electric XRD-5 diffractometer equipped with a copper x-ray tube, a NaI scintillation counter, a pulse height discriminator, and a quarter-circle Eulerian cradle goniostat. The 40kv, 20ma copper radiation was filtered through a .001" nickel foil mounted on the receiving slit of the scintillation tube. The crystalline b axis and the glass fiber direction was parallel to the phi axis of the instrument. The x-ray tube was set at a 2° take-off angle, and Bragg reflection angles for the resolved $CuK\alpha_1$, $CuK\alpha_2$ doublet were measured through a narrow slit along the h00, 0k0, and 00ℓ diffraction directions. From these measurements precise cell dimensions were calculated:

$a = 6.110 \pm .002$; $b = 9.532 \pm .003$; $c = 8.831 \pm .003$;

$\beta = 98.72 \pm .01$

With Z=2 these parameters gave a calculated density of 1.242g/cc^3. The density measured at room temperature by flotation in ethylene chloride and ethyl acetate was 1.245g/cc^3, which agrees closely with the calculated value of 1.242g/cc^3.

With the crystal and instrumentation described above, a stationary crystal, stationary counter technique and ten second point counts were used to obtain raw intensity data. Backgrounds were estimated from curves of counts/10sec versus two theta angle for various combinations of fixed values of the instrument angles chi and phi. Net intensities were obtained by direct subtraction of these background values from the measured peak-height values for each reflection. All independent intensities within the copper sphere of reflection were measured in shells out to a maximum two theta angle of 120 degrees ($\sin\theta/\lambda$ = .5617). There were 759 independent intensities of which 17 were too weak to be measured. When it became evident that these data were suffering from some sort of systematic error, either due to counter saturation or extinction, another set of diffractometer data was taken on another crystal.

The second crystal was longer and more narrow than the first (.34 x .10 x .032mm), and thus was almost needle-shaped. The crystal is shown in Fig. III-1. The crystal was mounted on a thin glass fiber with epoxy resins (Carter's general purpose epoxy; batch 230; Carter's ink company, Cambridge, Mass.). The needle-like crystal was aligned with the needle axis, the c axis, within 30 degrees of the fiber axis. With no preliminary photographs the crystal was placed on a Picker four-circle automatic diffractometer. The cell dimensions obtained earlier were substantiated and were adequate along

with measured chi, phi, and omega settings of the 0,0,10 and the 020 reflections to set the orientation matrix of the PDP-8I computer which controlled the diffractometer.

The diffractometer was equipped with a copper tube, a full-circle goniostat, a NaI scintillation counter, a pulse height analyzer, and a graphite monochromator. The Picker diffractometer was controlled through an FASC-1 interface by a PDP-8I computer. All computer programs for the PDP-8I were furnished by the Picker Corporation. Copper x-rays generated at 30kv and 16ma were monochromatized to Kα radiation by a graphite monochromator set at 26.33 degrees, corresponding to the 000^2 reflection of graphite. No filters were used, but attenuators were automatically engaged for all reflections over 10,000 counts.

The computer was programmed for θ-2θ scans of the diffraction peaks. The peaks were sharp at both 2 and 3 degree tube take off angles, and from observation of some broad low angle peaks, the maximum scan width, s, was determined to be 1.80 degrees. The θ-2θ scans were done in the following way: The 2θ position of each peak was calculated for $CuK\alpha_1$ radiation. The diffractometer was then set at a value 2θ-s/2, and scanning was begun at the rate of 1 degree per minute through the peak until a value of Δ2θ+s was reached, where Δ2θ is the calculated separation of $K\alpha_1$ and $K\alpha_2$ for each reflection. Dispersion corrections on the x-rays were made before the calculation of the 2θ peak settings. For copper radiation,

$d\lambda/\lambda$ = .002487. Background was measured with offset in 2θ = .70 degrees above and below the terminal points of the 2θ peak scan, with counting times of 10 seconds in each position. The equivalent reflections $hk\ell, \bar{h}k\bar{\ell}$; $\bar{h}k\ell, hk\bar{\ell}$ were measured throughout the k hemisphere of reciprocal space in planes of constant h from $\bar{h}$ to h out to 124.5 degrees in 2θ($\sin\theta/\lambda$=.574). The 106 and $\bar{1}04$ reflections were chosen as standards, and were measured at intervals of 50 reflections. The time for each reflection averaged over the complete data set was 2.67 minutes.

The intensities and their standard deviations were calculated with the following formulas from the total counts measured on the Picker machine.

$$I = c - \frac{t_c(B_1 + B_2)}{2t_B}; \quad \sigma^2(I) = c + \frac{t_c^2(B_1 + B_2)}{4t_B^2}$$ where c

is the total count during the scan for a time t_c, B_1 and B_2 are the two background counts, each taken for time t_B. Observed structure factors, F's, were calculated from these intensities after they were corrected for Lorentz and polarization effects. A total of 1604 data were measured, of which 1594 corresponded to the equivalent Friedel pairs ($F\bar{h}k\ell = Fhk\bar{\ell}$). The OkO reflections were measured only once. These 1604 data, of which 19 were zero, gave 807 independent reflections after the F's of all equivalent reflections had been averaged. The standard deviation of each equivalent measurement from the

average Fobs was estimated as:

$$\sigma(F) = [\sigma^2(F)_{Av} + p^2F^2/4]$$

where $\sigma(F)_{Av}$ was taken as the greater of the two values

$$\frac{1}{n}\left(\Sigma \frac{k_i^2\sigma^2(I_i)}{4F_i^2}\right)^{1/2} \quad \text{and} \quad \frac{1}{n-1}(\Sigma\Delta_i^2)^{1/2}$$

where n is the number of equivalent reflections averaged, I_i is the intensity of the i^{th} reflection which was multiplied by correcting factor (including absorption and Lorentz and polarization factors) k_i, to be reduced into F_i^2; and Δ_i is the difference between F_i and the average F of the set. The term $p^2F^2/4$ was included to reduce the weight given to large intensity reflections, which are more liable to suffer from systematic errors. This has an effect similar to that of including a term $(pI)^2$, where p is a small fraction (.05 in this case), in the value of $\sigma^2(I)$. The linear absorption coefficient, μ, of 5-methoxytryptamine for copper radiation was $6.58cm^{-1}$, and μt for the largest dimension of the crystal was .22. No correction for absorption was made.

Refinements were made using a full-matrix least squares program, which minimizes $\Sigma\, w(|F_o| - |F_c|)^2/w|F_o|^2$, where $|F_o|$ and $|F_c|$ are the magnitudes of the observed and calculated structure factors, respectively, and $w = [\sigma(F)]^{-2}$.

All calculations were made on the CDC-6600 computer.

The Fortran computer programs used were:

1. Paper - for reduction of the total counts on paper tape to intensity data on cards.
2. Edit - for finding possible bad intensity measurements before averaging equivalent reflections.
3. Incor - for Lorentz and polarization corrections.
4. Sort - for blending and sorting equivalent reflections.
5. Fordap - for Fourier and Patterson summations.
6. Distan - for distance and angle calculations.
7. Least Squares - for adjusting final structure parameters, and calculations of structure factors.

All of these programs were written by Dr. Allan Zalkin of this laboratory, and are unpublished.

Additional programs used were:

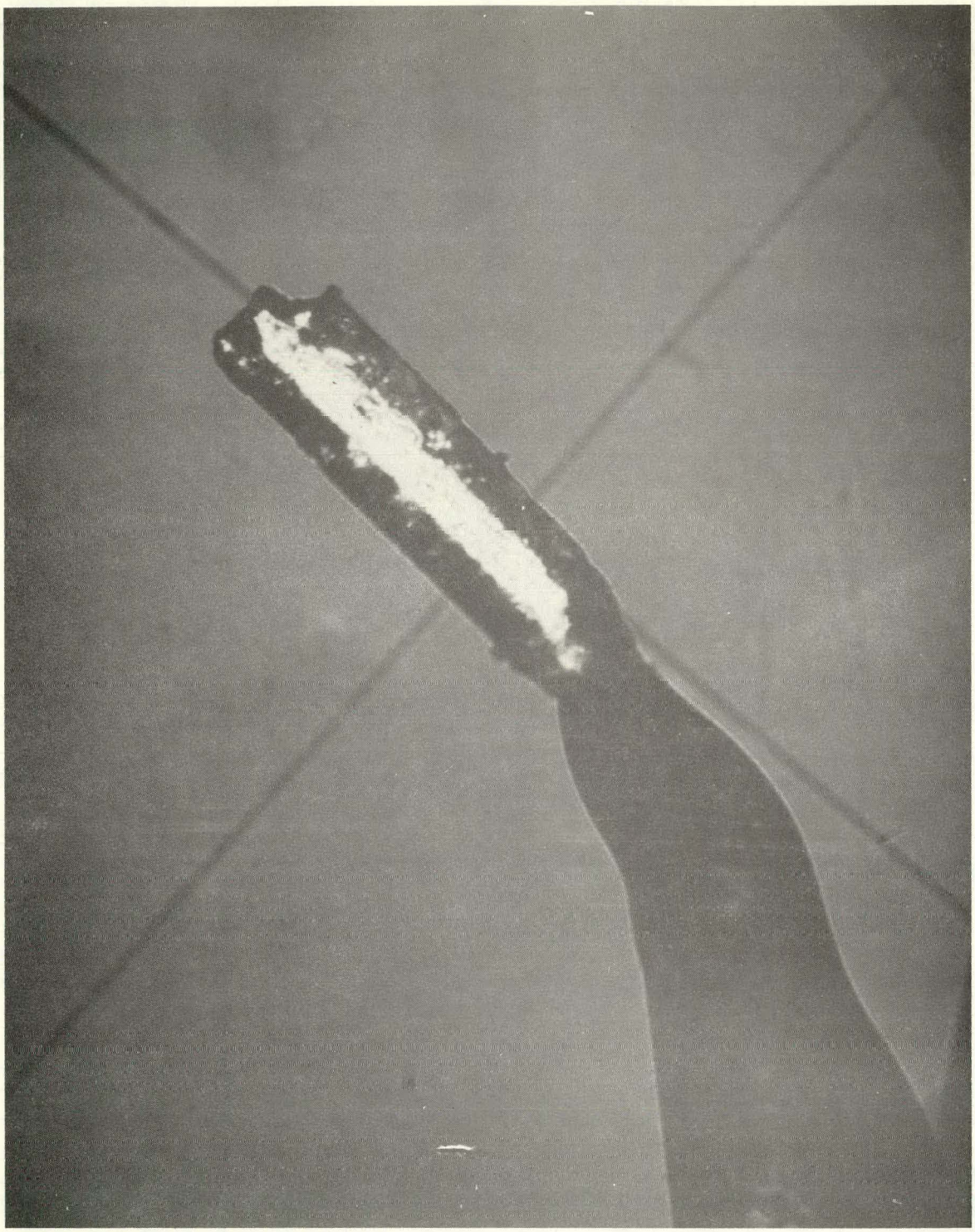
8. LSPLAN - from the University of Pittsburgh. This program calculates best fit planes through chosen input points.
9. Wilson - written by H. S. and M. L. Maddox for calculation of E's and an average temperature factor.
10. Ortep - a drawing program written by Carrol Johnson.[13]
11. Norma - a program which positions an orthonormal coordinate system in a crystal cell.
12. Omoo - a trial and error program for solution of crystal structures which have a moiety with a known rigid geometry.
13. Base - a program which calculates coefficients for a sharpened Patterson function.

The last three programs were written by the author of this thesis.

The source of the atomic scattering factors used was given in Section I-B. The anomalous dispersion corrections $\Delta f'$ and $\Delta f''$ for this light atom structure were zero.

Figure III-1

Photograph of the 5-methoxytryptamine crystal on which final intensity data were measured. The b* axis of the crystal is nearly perpendicular to the page in this projection. The dimensions of the crystal are (.34 x .10 x .032 mm).



XBB 706-2804

Figure III-1

C. Solution of the Structure

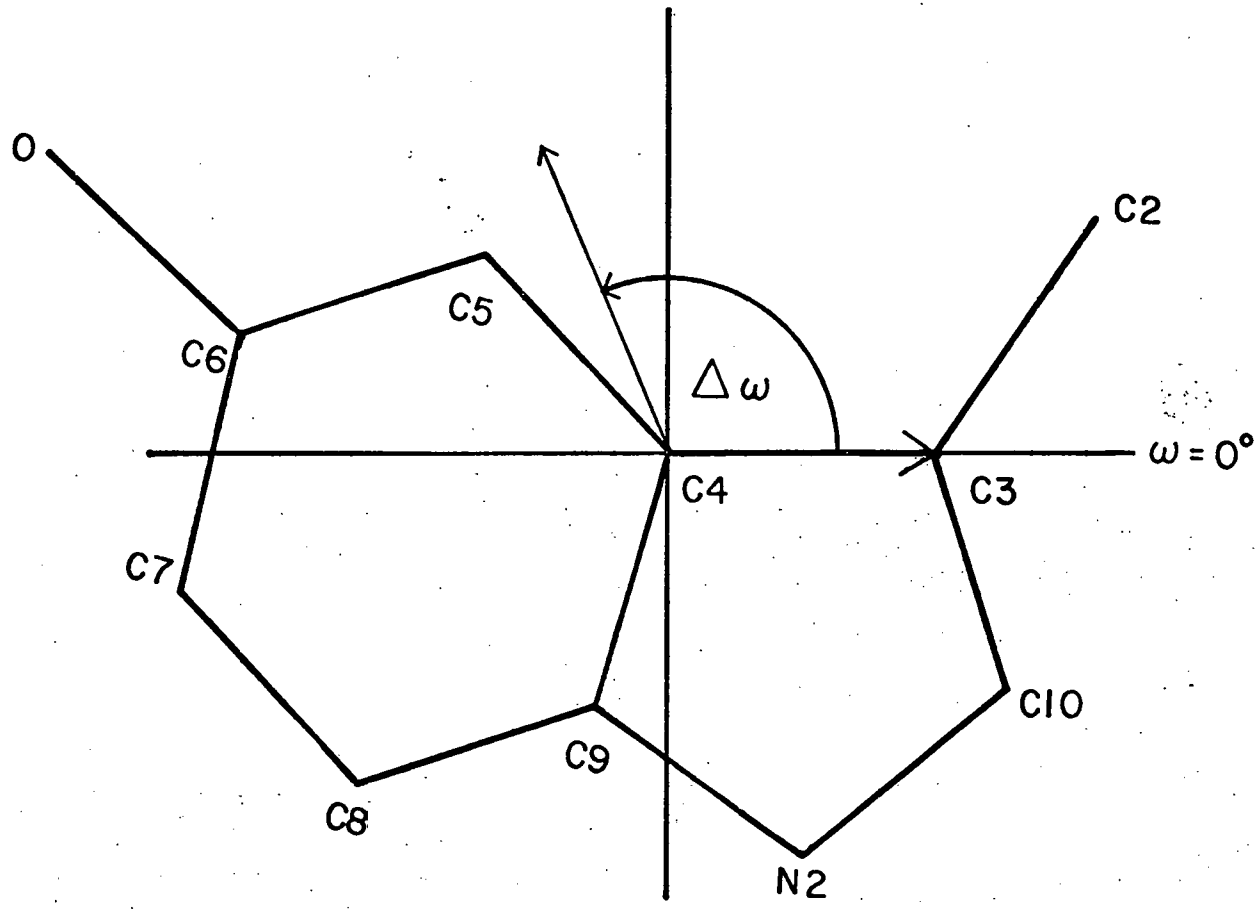
The crystal structure of 5-methoxytryptamine was solved by computer trial and error. The logic of the program is set forth in Section II, and will not be dealt with here.

In general, six parameters must be specified to place a known, rigid planar molecule in a unique position within a crystallographic cell. Three translational coordinates of any atom and in addition three angular coordinates are needed to specify position of the 5-methoxytryptamine molecule within the unit cell. Since the molecule is planar, the angular orientation can be fixed with the assignment of a Bragg plane, and a rotation angle of the molecular plane about the Bragg plane normal. As described in Section II, use of this program involves graphing known geometry into an orthonormal coordinate system. The approximate geometry of the indole ring was known from the structure determination of indole acetic acid. This graph, Figure III-2, indicates clearly the numbering system used for the indole ring, the choice of intramolecular origin, and the definition of the angle omega. The numbering system used for all the atoms will be given later.

The 33 positional parameters of the atoms within the indole ring and its immediate substituents can be reduced to three trial and error parameters: the y coordinate of C(4), the Bragg plane in which the molecule lies (this will be ab-

Figure III-2

Definition of the rotation angle omega used in the trial and error program. Counterclockwise motion of the C(4)-C(3) vector about the perpendicular to the plane through C(4) is defined as the direction of positive increasing angle omega.



XBL 704 - 804

breviated with hkl), and the rotation angle, omega, of the C(4)-C(3) vector about the Bragg plane normal.

The first attempts to solve this structure were made using the manual diffractometer data set collected as outlined in III-B. A Patterson function was calculated using all the Lorentz and polarization corrected intensities which were thus reduced to $(Fobs)^2$.

The Patterson function contained very few discrete peaks, and was quite diffuse except for a large peak at x = .50, y = .25, z = .09. This was the largest peak in the Patterson except for the origin and Harker peaks. The position of this peak and the general sense of the diffuse electron density distribution seemed to indicate the position of the ring to be nearly parallel with the -121 diffraction plane. This assignment was consistent with the fact that the Lorentz and polarization corrected structure factor for the -121 plane was the fourth largest present in the data set. A close examination of the films revealed pseudo-A-centering. Since there were glide planes at y = 0 and y = 1/2, this pseudo-A-centering meant that most of the x-ray scattering material was centered around y = 1/4. An attempt to calculate E values at this point was thwarted by a negative temperature factor in the Wilson plot. At the time the first experiments with this structure were done, it was not clear what this fact meant.

The first use of this trial and error program was undertaken with the knowledge that the plane of the molecule was approximately $\bar{1}21$, that the center of the x-ray scattering material was near y = .25, and that the y coordinate of C(4) should be larger than .10 in order to avoid collisions of the symmetry related molecules. Fifty observed structure factors were selected from the data set, including about 25 of the most intense. Structure factors calculated from the 33 atomic positional coordinates at each orientation of the molecular plane were compared with the observed structure factors. An R factor was computed at each orientation after a linear scale factor was applied to the Fobs. (A detailed description of how the R factor was calculated, and the method of calculating structure factors is given in Section II.)

With a plane orientation of $\bar{1}21$, the cell was swept in increments of .02 in y, and 10° in omega from y = 0 to y = .50 and from omega = 0° to omega = 360°. The x and z coordinates of C(4) were fixed at .60 and .30 respectively. More of the cell was swept than was necessary in this first use of the program. Two minutes of computing time was used. Several minima were found near y = 0. These were rejected because intermolecular collisions resulted at these values of y. There was, however, one unique minimum well-removed from the glide plane in the y direction that seemed to be

the best possibility. The trial parameters of this minimum are listed in Table III-1. The positional parameters for the atoms corresponding to this minimum at y = .30 and omega = 150° were used as input for four cycles of least squares refinement against all the 812 manual data, using isotropic temperature factors for the 11 atoms. All the parameters were allowed to vary except the x and z coordinates of C(4). The results were not spectacular. The R factor was 62% and all the temperature factors were negative.

Evidently the $\bar{1}21$ orientation was wrong, and a list of possibilities was in order: (1) The $\bar{1}21$ orientation gave a totally false R factor minimum and the true plane orientation was quite unrelated to y = .30, omega = 150°, $\underline{hkl} = \bar{1}21$. (2) The $\bar{1}21$ orientation was approximately correct, and could be adjusted slightly to obtain the right answer. (3) The diffractometer data were suffering from systematic errors of measurement. (4) The computer program had a systematic error in it.

The computer program was thoroughly de-bugged, and the data set was compared with visually estimated intensities from a set of Weissenberg films. The diffractometer data generally agreed with the film data, and the problem was reduced for the moment to a systematic investigation of possibilities (1) and (2).

First, 70 plane orientation normals, $\underline{hkl}$'s, ranging

Table III-1

Values of the Trial and Error Parameters HKL, Y (fractions of translational direction along b), and Omega (degrees) for R (%) Minima less than Sixty Per Cent.

The search included 70 plane normal possibilities, hkl's. The approximate angles A, B, C (all in degrees) which each hkl plane normal makes with the coordinate axes a, b, c are also listed.

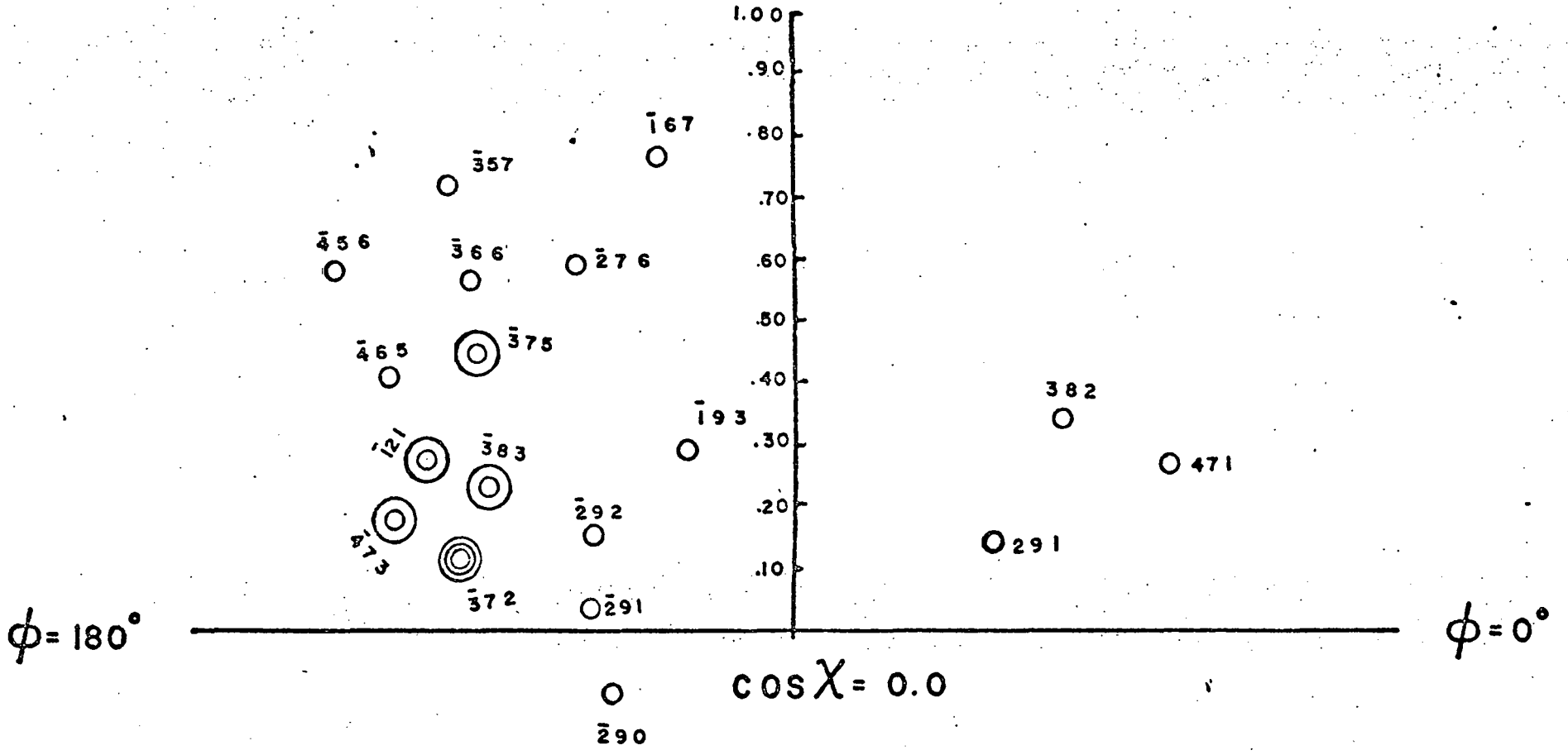
HKL	Y	Omega	R	A	B	C
-121	.30	150	47.8	54	42	66
-383	.12	140	48.6	61	34	70
-473	.32	120	49.8	50	44	70
-375	.28	140	48.0	61	44	56
-291	.32	120	57.2	71	20	83
471	.16	140	55.3	49	44	83
291	.20	140	58.1	71	22	83
-292	.24	80	54.1	72	22	77
382	.12	140	55.3	61	35	77
-193	.32	40	57.9	81	21	70
-465	.32	120	51.3	51	53	56
-366	.28	140	51.4	61	52	47
-456	.28	120	56.2	51	60	48
-276	.24	80	58.8	72	45	48
-167	.32	80	59.9	81	52	38
-357	.28	140	59.1	61	59	39
-290	.12	20	55.9	71	20	90

over a hemisphere of the reciprocal lattice were generated. These planes corresponded to all the Bragg reflection planes in a thin shell of the reciprocal lattice for $100° \leq 2\theta \leq 104°$ and $.49685 \leq (\sin\theta/\lambda) \leq .51110$. The cell was searched for all values of the parameters $0 \leq y \leq 1/2$ and $0° \leq$ omega $\leq 180°$ with $\Delta y = .04$ and $\Delta\omega = 20°$ for each of the 70 hkl's. Because of the glide symmetry, this search included plane normal possibilities ranging over a complete sphere, and within limitations of grid size, all possible values of y and omega. The values of trial parameters corresponding to the 16 cases with $R \leq 60\%$ are listed in Table III-1. For clarification, the angles which each plane normal makes with the coordinate axes are also listed in Table III-1, and an heuristic map of the search is shown in Figure III-3. This systematic search of the entire reciprocal lattice for the best-fit plane normal orientation vectors showed, besides the $\bar{1}21$, only three possibilities with $R \leq 50\%$. Two of these vectors were within 6 degrees of the $\bar{1}21$ vector, and the third was within 14 degrees of the $\bar{1}21$ possibility. From these results it was concluded that the plane orientation normal had to be near $\bar{1}21$. Also, it was noticed that most of the y values were in the range $.20 \leq y \leq .32$, and this fact seemed to substantiate the earlier reasoning that the y coordinate of C(4) must not be near either $y = 0$ or $y = 1/2$.

Since the approximate orientation seemed to be $\bar{1}21$,

Figure III-3

Map of the Trial and Error Search in 5-Methoxytryptamine. The circles represent termination points of plane-normal vectors projected onto the ac plane. Terminal points of plane normal orientation vectors which gave an R factor of less than 50% are circled twice. The correct orientation vector is circled three times.


1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
$\bar{1}67$
$\bar{3}57$
$\bar{4}56$
$\bar{3}66$
$\bar{2}76$
$\bar{4}65$
$\bar{3}75$
382
$\bar{1}93$
$\bar{1}21$
$\bar{3}83$
471
$\bar{2}92$
291
$\bar{3}72$
$\bar{2}91$
$\phi = 180°$
$\phi = 0°$
$\cos\chi = 0.0$
$\bar{2}90$

XBL 704-834

125 plane orientations near the $\bar{1}21$ were generated by systematic variation of the hkl $\bar{5}$ 10 5. The integers hkl are proportional to the direction cosines of the hkl plane normal, and an hkl of $\bar{5}$ 10 5 describes the same plane as hkl = $\bar{1}21$. At this time it seemed reasonable to limit the search to the y and omega values that had led to the lowest R with hkl = $\bar{1}21$. About 20 seconds of computing time was used to calculate R factors for the 125 hkl's at y = .30 and omega = 150°. Forty-eight of these hkl's correlated with R factors of less than 53%, and of these, eight showed a minimum of 48%. In order to distinguish between these eight possibilities, some of the intense low angle Fobs were removed from the data deck of the trial and error program and high angle Fobs were added. The total number of Fobs used in the R factor calculation still remained 50. The eight possibilities resulting from this search are listed in Table III-2. The R factor reported there refers to calculations made with some of the intense low angle Fobs removed from the deck.

The coordinates of the 11 atoms for each of the four most promising orientations were refined with full matrix least squares against 170 low angle data. Isotropic temperature factors were applied to all the atoms in 8 cycles of refinement. The input geometry was wildly distorted in every case. The $\bar{4}95$ orientation gave the lowest conventional R value in the least squares, R = 36%.

Table III-2

Trial Parameters Near the $\bar{1}21$ Orientation.

The y parameter is fixed at .30 (fraction of the b translational direction), and omega is fixed at 150°. A, B, and C are the angles each hkl vector makes with the coordinate axes a, b, c.

HKL	R(%)	A	B	C
$\bar{5}105$	52.2	54	42	66
$\bar{5}106$	52.2	56	43	61
$\bar{4}85$	52.5	56	43	60
$\bar{4}86$	51.9	58	46	56
$\bar{4}94$	49.5	57	38	68
$\bar{4}95$	47.4	59	40	63
$\bar{3}63$	53.2	55	44	66
$\bar{3}64$	52.3	57	44	59

Further refinements tested thoroughly the $\bar{4}95$ possibility. This orientation was refined again with constrained geometry against the same 170 data. Only the linear scale factor and the isotropic temperature factors of the 11 atoms were allowed to vary. This refinement gave, after six cycles, negative temperature factors for all the atoms. The $\bar{4}95$ orientation was refined again in the same way against about 500 high angle data, and all the temperature factors remained positive. Refinement of the same geometry for six cycles of least squares against 500 high angle data allowing all parameters to vary except the x and z coordinates of C(4) gave an R = 36%, and the geometry, though distorted by refinement, was much more reasonable. Refinement against all the data allowing all positional parameters except those fixed by symmetry and all the isotropic temperature factors to vary for six cycles gave R = 46%. These results, combined with the knowledge of the negative overall isotropic temperature factor calculated from the Wilson plot, led to a suspicion of the low angle data. It was felt that they were systematically low due either to counter saturation or extinction. It was realized that there was little chance of success with this method unless the data chosen for trial and error could be relied upon as representative. A complete set of new diffractometer data was taken as outlined in III-B.

Against the new data set the $\bar{4}95$ orientation gave an R

value in least squares against all the data of 43%. In this refinement all parameters were varied except those fixed by symmetry. As before, isotropic temperatures were used for six cycles of refinement. The 50 structure factors chosen for trial and error had led to a false structure. Somewhat surprisingly, the Wilson plot still gave a negative temperature factor B = -2.123 $\AA^2$, and again it was impossible to calculate reliable Ehkl values.

The problem was now clearly defined. A data set for trial and error had to be chosen that was free of systematic errors and that was representative of the total data set. The following experiment was performed in order to decide more quantitatively what effect the data sample had on the R factor in the trial and error program: First, the coefficients Phkl = Fhkl/[$\hat{f}$hkl x exp(-B x $\sin^2\theta/\lambda^2$)] were calculated for all structure factors Fhkl in the data set. The $\hat{f}$hkl is the average scattering factor for all the atoms of the structure, and the isotropic temperature factor B was fixed at 3$\AA^2$. The hkl's corresponding to the largest of these coefficients Phkl were used as plane orientation possibilities in this experiment to test the data sampling procedure. The results are summarized in Table III-3. It was evident that the structure selected by the R factor probe was a function of the data sample. The new automatic diffractometer data, (1a), gave the same minima as the old

Table III-3

Values of Y (expressed as fractions of the translational direction b), Omega (degrees), and R (%) as a Function of Data Choice for Plane Orientations (HKL) with Large Phkl Coefficients.

Phkl				9.17			----			9.23			11.3		
HKL				$\bar{2}51$			$\bar{4}95$			$\bar{3}72$			106		
Data Set	Range of y	Δy	Δω	y	ω	R	y	ω	R	y	ω	R	y	ω	R
(1)	$0 \leq y \leq .50$	.04	15	.12	120	48.8	.12	135	48.4	.12	135	47.6	.24	0	73.9
(1a)	$0 \leq y \leq .50$	.04	15	.12	120	47.7	.12	135	46.7	.12	135	46.9	.24	0	73.7
(2)	$0 \leq y \leq .50$	.04	15	.36	120	36.8	.12	105	39.2	.20	0	35.2	0	30	53.9
(3)	$.20 \leq y \leq .30$	.02	10	.40	0	40.6	.12	135	42.2	.26	10	40.6	---	--	----
(4)	$.20 \leq y \leq .30$	.025	15	.25	0	39.3	.27	30	42.2	.25	15	35.2	---	--	----

Data Sets (1) and (1a) refer to the same set of 50 low angle, intense Fhkl. Data Set (1) was measured manually, and Data Set (1a) was measured on the Picker Automatic Diffractometer. Data Sets (2), (3), and (4) were all measured on the Picker machine. Data Set (2) refers to 50 Fhkl chosen at random from the data deck of 807 cards. Data Set (3) refers to 50 low angle Fhkl of medium to medium-strong intensity, and Data Set (4) refers to 100 Fhkl produced by the addition to Data Set (3) of 50 Fhkl taken from the data deck at random. The $\bar{4}95$ orientation is included for comparison. Omega is in the range $0 \leq \omega \leq 180°$.

manual data (1) for the same set of 50 low angle intense data. Fifty structure factors chosen at random from the new data deck of 807 cards, Data Set (2), gave a different set of R factor minima for the four plane normal orientations tested. Fifty low angle data of medium to medium-strong intensities, Data Set (3), gave another set of R factor minima different from the first two. Fifty data were not enough to allow the unique orientation of the 5-methoxytryptamine ring to be chosen by trial and error. It was noticed that the $\bar{3}72$ plane orientation, however, consistently gave lower R values than the other possibilities. It was also noticed that the y coordinate of the minimum in Data Set (3) allowed for maximum spacing of the molecules away from the glide planes at y = 0 and y = 1/2. The atomic coordinates corresponding to the trial parameters hkl = $\bar{3}72$, y = .26, omega = 10° were refined in the full-matrix least squares program with isotropic temperature factors against all the data in eight cycles to an R factor of 33%.

After the 33% R factor had been calculated for the $\bar{3}72$ orientation at y = .26, omega = 10°, the course of the work was split into two parts. The grid size for the trial and error program was reduced to $\Delta y = .01$, $\Delta\omega = 1°$ and the orientations around hkl = $\bar{3}72$, y = .26, and omega = 10° were examined more closely. Data Set (4) was used in this probe. The minimum from this search was at y = .25, omega = 7°.

Positional parameters from this orientation gave an R factor of 28% in least squares. (The details of these final refinements of 5-methoxytryptamine are given in Table III-4 and will not be elaborated in the text. The R factor mentioned in the text is the conventional R factor, R_1.) At the same time, seventeen possibilities of hkl with plane normal direction cosines similar to those of $\bar{3}72$ were used to orient the molecular plane for trial and error. The search range $.20 \leq y \leq .30$, $0 \leq \omega \leq 180°$ was used with a grid size $\Delta y = .025$, $\Delta\omega = 15°$. Another minimum was found with hkl = $\bar{4}92$, y = .25, omega = 15°.

As can be seen in Table III-4, the $\bar{4}92$ orientation refined to the same structure as the $\bar{3}72$ orientation when the 11 atoms obtained by trial and error were input into least squares. The dihedral angle between least squares planes through the $\bar{3}72$ orientation plane and the final refined position of the ring was 4°. The angle between the $\bar{4}92$ and the final refined position was less than 1°. The fact that another plane orientation refined to the same structure as the $\bar{3}72$ was added confirmation of the structure. The phases from the 28% least squares refinement of the $\bar{3}72$ orientation were used to calculate a Fourier. The peaks in this Fourier ranged from 0 to 13.22. The 11 largest peaks ranged from 8.98 to 13.22 and corresponded to the atoms used to calculate the phase information. The three largest peaks besides

Table III-4

Schedule of Full Matrix Least Squares Refinements for 5-methoxytryptamine.

HKL	No. Atoms	Type of Temp. Factor	FSD	R1(%)	R2(%)	R3(%)	SD	k' x 10^{-6}	No. Cycles
372	11	I	.05	27.8	32.8	29.7	8.98	0	8
492	11	I	.05	27.8	32.8	29.7	8.98	0	8
372	14	I	.05	8.67	10.9	9.14	3.008	0	6
372	14	A	.05	6.37	8.0	6.86	2.327	0	4
372	28	I	.05	6.19	7.1	6.65	2.054	0	4
372	28	A	.05	3.32	3.5	3.64	1.081	0	4
372	28	A	.05	3.06	3.3	3.38	1.004	.08	4
372	28	A	.04	2.51	2.8	2.84	.988	.30	4
372	28	A	.03	2.51	2.6	2.84	1.115	.30	4

HKL refers to the orientation of the molecular plane chosen for refinement.
No. Atoms includes the number of hydrogens when applicable.
Type of temperature factor, either isotropic (I) or anisotropic (A), refers to non-hydrogen atoms since all hydrogens were refined isotropically. The form of these temperature factors is given in III-C.
FSD is the proportion of the intensity to be used as a weighting factor in least squares, and is explained in III-C.
R1 is the conventional R factor excluding 61 zero weight data, R2 is the weighted factor, and R3 is the conventional R factor including zero weight data. Data for these refinements were obtained with a Picker automatic diffractometer. There were a total of 807 independent data, of which 61 were given zero weight in the calculation of R1.
SD is the standard deviation of observation unit weight for each refinement.
k' is the extinction correction applied.

these, ranging from 3.28 to 3.94, corresponded to the reasonable distances expected for the remaining non-hydrogen atoms. (These numbers are merely indicative of peak size unless the Fourier is calculated on the basis of all the atoms. Then the numbers associated with peak height refer to the number of electrons. This total of 14 non-hydrogen atoms refined isotropically to 8.7%, and anisotropically to 6.3%.

A difference Fourier was calculated, phased on the anisotropically refined positions of these 14 atoms. As an aid in the search for hydrogens in 5-methoxytryptamine, the postulated geometry of the ring hydrogens was graphed into an orthonormal set, and the coordinate calculation subroutine of the trial and error program was used to calculate the expected positions of the ring hydrogens in the crystallographic unit cell. On the difference Fourier, peaks ranged from zero to .36 electrons. The calculated positions of the ring hydrogens were in every case near the position of rather large peaks in the difference map. Thirteen of the top 16 peaks corresponded to the reasonable distances from carbon and nitrogen atoms expected for hydrogen atoms. These peaks were all in the range .19 to .36 electrons. The hydrogen attached to the ring nitrogen had a peak height on the difference map of .16 electrons. These 14 hydrogens were added to the 14 other atoms and the structure was refined isotropically to an R factor of 6.19%. When the non-

hydrogen atoms were refined anisotropically and the hydrogens isotropically, an R factor of 3.32% resulted. A close examination of observed and calculated structure factors from this refinement revealed that Fobs was systematically observed too weak for the intense reflections, by 20% in the worst case.

An empirical, linear extinction correction was applied, using the relation $Fc = kFobs(1 + k'Iobs)$. The least squares program scales the data in exactly this way. In this equation Fc is the calculated structure factor for each hkl, and Fobs is the observed structure factor. Iobs is the observed intensity. The constant k is the linear scale factor for Fobs, and k' is the proportion of Iobs that must be added to Fobs to correct for extinction. A graph of (Fc/kFobs) versus Iobs was prepared, and the best visually estimated straight line was drawn through the resulting points. The slope of this line k' was an estimate of the magnitude of the extinction effect. From this graph $k' = 8.3 \times 10^{-7}$. This value was too large and resulted in overcorrection of the Fobs for extinction as Fobs now systematically exceeded Fc for the intense reflections. A refinement was then done with $k' = 8.3 \times 10^{-8}$. This value of k' resulted in an R factor of 3.06%, but it was not large enough to remove the systematic discrepancies. An extinction correction, k', of 3.0×10^{-7} randomized the systematic

error in the intense Fobs without overcorrecting. The least squares weighting factor P, described in III-B, partially accounts for systematic effects such as absorption and extinction. When a proper extinction correction was applied to the observed data, the fraction P=.05 proved to be too large. The intense data were over-weighted and the standard deviation of observation unit weight, SD, dropped below one. P was accordingly reduced to .03 to avoid over-weighting the data for systematic errors.

The final extinction correction with $k' = 3.0 \times 10^{-7}$ lowered the R factor from the value uncorrected for extinction of 3.32% to a low of 2.51%. The least-squares-estimated standard deviation of bond lenght on C-C distances dropped from .004 to .003. In this final refinement no parameter shifted by more than 3% of its estimated standard deviation. Final discrepancy factors are: R1 = 2.51% for 746 non-zero-weight data; R3 = 2.84% for all 807 independent data; and the weighted R2 = 2.6%. The standard deviation of observation unit weight is 1.115. There is no systematic trend in either $|Fobs/Fc|$ or $\underline{\omega}^{1/2}|\Delta F|$ as a function of intensity or Bragg scattering angle. No peaks in the final difference Fourier based on the final structure were higher than .15 electrons.

D. Discussion of the Structure

The atomic coordinates of all the non-hydrogen atoms are given in Table III-5, and the thermal parameters of these atoms are given in Table III-6. The atomic coordinates of the hydrogen atoms along with their thermal parameters are given in Table III-7. The atomic numbering system and bond distances are in Fig. III-4. The observed an calculated structure factors for 5-methoxytryptamine are given in Table III-8. First, the intramolecular aspects of the structure will be discussed. After this, the packing arrangements of the molecules will be described.

The most predominate feature of the 5-methoxytryptamine structure is the indole ring. Least squares planes were calculated through the indole, the benzene, and the pyrrole rings separately. Distances of atoms from these least squares planes are shown in Table III-9. Atoms given zero weight in the least squares are indicated with parenthesis. A least squares plane through the indole ring giving all nine atoms of the ring full weight in least squares show it to be planar to within .02Å. The first member of the side chain, C(2), is .035Å above this plane, and the oxygen, O(1), substituent at C(6) is -.02Å below this plane. The indole rings in glycyl-ℓ-tryptophan,[14] ℓ-tryptophan hydrobromide,[15]

Table III-5

Atomic Coordinates and their Standard Deviations (a) for all Non-hydrogen Atoms in 5-Methoxytryptamine.

ATOM	X	Y	Z
C(1)	.3467(0)	.3571(3)	.9726(0)
C(2)	.5789(7)	.2923(3)	.0040(5)
C(3)	.6839(6)	.3048(2)	.1689(4)
C(4)	.6088(6)	.2435(2)	.3008(4)
C(5)	.4265(6)	.1572(2)	.3189(4)
C(6)	.4050(6)	.1152(3)	.4647(4)
C(7)	.5609(6)	.1535(3)	.5922(4)
C(8)	.7384(7)	.2375(3)	.5765(5)
C(9)	.7606(6)	.2834(2)	.4291(4)
C(10)	.8715(6)	.3782(3)	.2241(4)
C(11)	.0721(7)	.0109(3)	.8763(5)
N(1)	.3332(6)	.4970(3)	.5195(5)
N(2)	.9196(6)	.3666(3)	.3801(4)
O(1)	.2350(6)	.0321(2)	.4990(4)

(a) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

Table III-6

Table of Anisotropic Temperature Parameters (a) and their Standard Deviations (b) in 5-Methoxytryptamine.

Atom	B11	B22	B33	B12	B13	B23
C(1)	3.5(1)	4.0(1)	2.8(1)	-.3(1)	-.06(9)	.6(1)
C(2)	3.9(1)	4.1(1)	2.9(1)	.1(1)	.79(9)	.1(1)
C(3)	2.9(1)	3.2(1)	2.9(1)	.43(9)	.60(9)	.15(9)
C(4)	2.4(1)	2.51(9)	2.7(1)	.31(8)	.21(8)	.02(8)
C(5)	2.7(1)	2.7(1)	2.5(1)	.18(8)	-.07(9)	-.17(8)
C(6)	3.3(1)	3.0(1)	3.0(1)	.00(9)	.40(9)	.16(9)
C(7)	4.2(1)	4.0(1)	2.5(1)	.1(1)	.1(1)	.35(9)
C(8)	3.7(1)	3.7(1)	2.7(1)	.0(1)	-.51(9)	-.2(1)
C(9)	2.5(1)	2.7(1)	3.2(1)	.25(9)	.17(9)	-.24(9)
C(10)	2.9(1)	3.8(1)	4.0(1)	.2(1)	.87(9)	.4(1)
C(11)	3.8(1)	4.0(1)	4.2(1)	.7(1)	.6(1)	-.2(1)
N(1)	2.9(1)	3.7(1)	6.0(2)	-.12(9)	-.2(1)	-1.0(1)
N(2)	2.67(9)	4.0(1)	4.0(1)	-.27(9)	.13(8)	-.11(9)
O(1)	4.14(9)	4.88(9)	3.09(8)	-1.19(8)	.31(6)	.89(7)

(a) Anisotropic thermal parameters, B, in units of A^2, are given by $B = 4\beta_{ij}/a_i^* a_j^*$, Where a_i^* is the i^{th} reciprocal cell length.

(b) Estimated standard deviations are given in parentheses following the parameter.

Table III-7

Final Positional Parameters and Isotropic Thermal Parameters (a) and their Standard Deviations (b) in 5-Methoxytryptamine for all the Hydrogen Atoms.

ATOM	X	Y	Z	B
H(1)	.394(7)	.501(4)	.630(6)	9.0(12)
H(2)	.411(7)	.449(4)	.455(5)	7.7(12)
H(3)	.248(4)	.300(2)	.027(3)	3.1(5)
H(4)	.292(5)	.354(3)	.862(4)	4.0(6)
H(5)	.563(4)	.189(3)	.977(3)	3.6(6)
H(6)	.673(6)	.339(3)	.938(4)	5.6(8)
H(7)	.318(4)	.126(3)	.229(3)	2.6(5)
H(8)	.532(5)	.119(3)	.692(3)	4.3(6)
H(9)	.845(5)	.262(4)	.663(4)	5.4(7)
H(10)	.046(6)	.415(4)	.434(4)	5.9(8)
H(11)	.972(6)	.433(3)	.160(4)	5.2(7)
H(12)	.969(5)	.072(3)	.929(3)	4.3(6)
H(13)	.006(5)	.067(4)	.320(4)	5.4(8)
H(14)	.131(5)	.068(3)	.792(4)	4.8(6)

(a) The isotropic temperature factor has the form $T=\exp(-B(\sin\theta/\lambda)^2)$.

(b) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

Table III-8

Observed and Calculated Structure Factors

for 5-methoxytryptamine.

TABLE OF OBSERVED AND CALCULATED STRUCTURE FACTORS FOR 5-METHOXYTRYPTAMINE

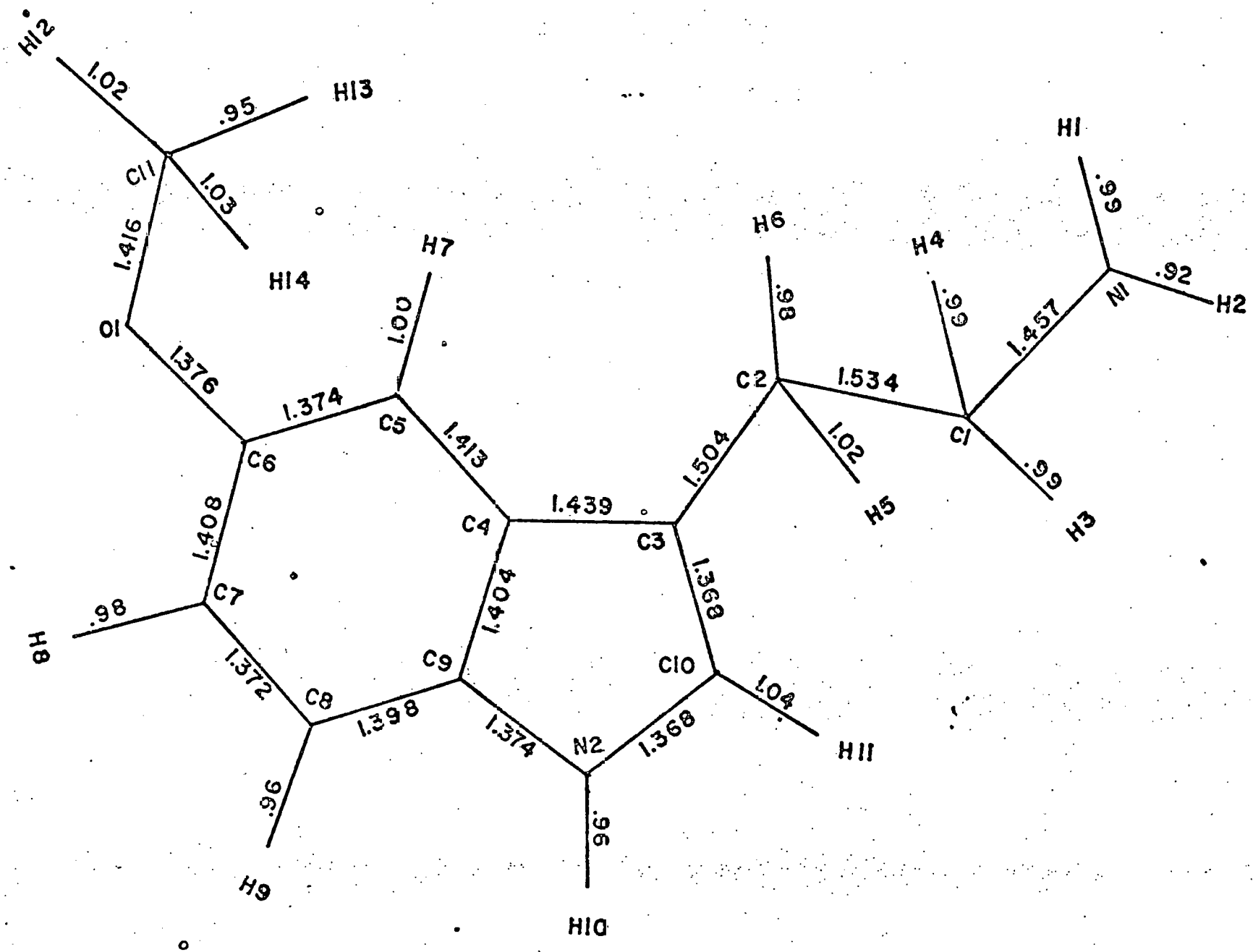
FCAL(0,0,0) = 2040

[illegible]

XBL 703-635

Figure III-4

Atomic Numbering System and Bond Distances


H12
1.02
.95
H13
C11
1.03
H14
1.416
O1
1.376
1.374
C6
C5
1.00
H7
1.413
1.408
C7
C4
1.439
.98
H8
1.372
C8
C9
1.404
1.398
1.374
N2
.96
H9
H10
C3
1.504
C2
1.534
C1
1.457
N1
H6
.98
H4
.99
H1
.99
.92
H2
1.02
H5
.99
H3
1.368
C10
1.04
1.368
H11

XBL 704-827

Table III-9

Distance (Å) of Atoms in the Indole Ring from Least Squares Planes.

Distances in parentheses refer to atoms given zero weight in the calculation of the plane.

Atoms	N(2)	C(3)	C(10)	C(9)	C(4)	
Indole Plane	-.012	.015	.003	-.010	-.005	
Benzene Plane	(-.009)	(.030)	(.016)	-.008	.004	
Pyrrole Plane	-.002	.003	-.001	.004	-.004	
Atoms	C(5)	C(6)	C(7)	C(8)	C(2)	O(1)
Indole Plane	-.004	-.010	.012	.010	(.035)	(-.020)
Benzene Plane	.005	-.010	.006	.003	(.060)	(-.021)
Pyrrole Plane	(-.004)	(.002)	(.040)	(.036)	(.007)	(-.006)

and indole acetic acid[16] were all reported as planar to within the accuracy of the structure determinations. In the most accurate of these studies, that of indole acetic acid, the R factor was 18.2%, and the least-squares-estimated standard deviation of bond lengths was .015 to .022 Å. The least-squares-estimated standard deviation of bond length in 5-methoxytryptamine is .003 Å for C-C distances and .004 Å for C-N distances. Within the accuracy of these standard deviations, the indole ring is not planar. According to Cruikshank, the criterion for significant difference between two measurements is that they be separated by at least three standard deviations.[17] As we can see in Table III-9, N(2), C(3), and C(6) are all more than three standard deviations away from the least squares plane through the indole ring. A least squares plane giving full weight only to benzene ring atoms shows that all the atoms of the benzene ring are within three standard deviations of this plane. The benzene portion of the indole ring is planar, and the indole ring as a whole is not.

Furthermore, a least squares plane giving weight only to the five members of the pyrrole ring N(2), C(10), C(3), C(4), and C(9) show the pyrrole portion of the indole ring to be planar within two standard deviations. The dihedral angle between the benzene ring and the pyrrole ring is about one degree. All the atoms of the indole ring and the two

substituents O(1), at carbon C(6); and C(2), at carbon C(3), lie within .008 Å, or two standard deviations, of the pyrrole ring plane. The exceptions to this are the carbons C(7) and C(8) of the benzene portion of the indole ring which depart from the pyrrole ring plane by about .040 Å, or ten standard deviations, and thus are significantly above the plane.

These results can best be summarized with the following: The indole ring and its immediate substituents lie in a common plane except for carbons C(7) and C(8) which lie significantly above the plane. A least squares plane through the two non-planar atoms and their immediate neighbors shows the indole plane to be puckered at C(7) and C(8) by 1.63 degrees.

Other authors have noticed this deviation from planarity of conjugated aromatic ring systems when the accuracy of the structure determination was high. The phthalocyanine structure was at first thought to be planar, and later more accurate work revealed that this was not the case.[18] The substituted naphthalene ring is not planar within the accuracy of the structure determination.[19]

No explanation for the deviation of C(7) and C(8) from the indole plane is obvious from considerations of packing or conformation. The closest approach of atoms outside the molecule to C(7) and C(8) is in every case greater than 3 Å. The nitrogen, N(1), in the side chain of the same molecule is

closest and is within 3.576 and 3.482 Å of these atoms respectively.

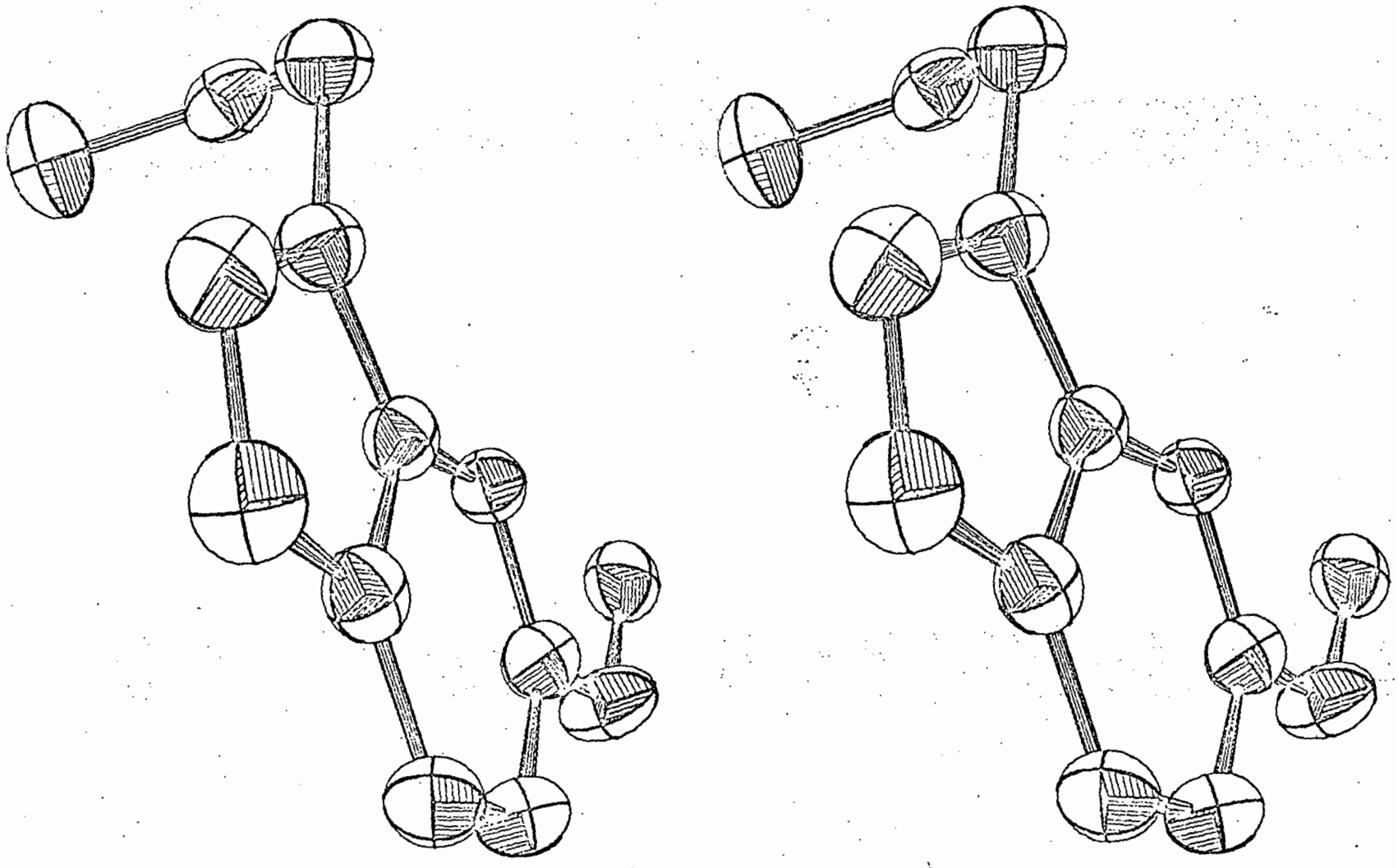
The anisotropic thermal parameters for 5-methoxytryptamine are reasonable, and are set forth in Table III-6. The anisotropic thermal motion of 5-methoxytryptamine is also shown in a stereoscopic illustration, Figure III-5. As we see in this diagram, the nitrogen N(2) of the ring has the largest amount of anisotropic thermal motion.

The carbons C(7) and C(8) certainly do not deviate enough from the plane to destroy the indole π electron resonance system. We would expect all the bonds in the indole ring to be significantly lengthened if this were true. The average of all bonds in the benzene portion of the indole ring in 5-methoxytryptamine is 1.395 ± .003 Å. All intramolecular distances in 5-methoxytryptamine are given in Table III-10. The most accurate bond distance for the carbon-carbon double bond in benzene is 1.397 ± .001 Å.[20] The average bond length for the benzene portion of 5-methoxytryptamine is equal to this most accurate value within the standard deviation.

The C(5)-C(6) and C(7)-C(8) bonds are six standard deviations shorter than the benzene average, and the C(4)-C(9)

Figure III-5

Anisotropic Thermal Motion in 5-Methoxytryptamine.



XBL 704-836

Table III-10

Intramolecular Distances (in Å) and their Standard Deviations (a) in 5-Methoxytryptamine.

Atoms	Distance	Atoms	Distance
N(1)-H(1)	.99(5)	C(7)-C(8)	1.372(4)
N(1)-H(2)	.92(5)	C(8)-H(9)	.96(3)
N(1)-C(1)	1.457(4)	C(8)-C(9)	1.398(3)
C(1)-H(3)	.99(3)	C(9)-C(4)	1.404(3)
C(1)-H(4)	.99(3)	C(9)-N(2)	1.374(3)
C(1)-C(2)	1.534(4)	N(2)-H(10)	.96(4)
C(2)-H(5)	1.02(3)	N(2)-C(10)	1.368(4)
C(2)-H(6)	.98(4)	C(10)-C(3)	1.360(4)
C(2)-C(3)	1.504(3)	C(10)-H(11)	1.04(3)
C(3)-C(4)	1.439(3)	O(1)-C(6)	1.376(3)
C(4)-C(5)	1.413(3)	O(1)-C(11)	1.416(3)
C(5)-H(7)	1.00(2)	C(11)-H(12)	1.02(3)
C(5)-C(6)	1.374(3)	C(11)-H(13)	.95(3)
C(6)-C(7)	1.408(4)	C(11)-H(14)	1.03(3)
C(7)-H(8)	.98(3)		

(a) Standard deviations have been estimated by the method of least squares and are indicated in parentheses.

bond is six standard deviations larger than the average. All the other bonds of the benzene portion of the indole ring are equivalent within six standard deviations. The bond average for the indole ring of 1.393 ± .003Å is not significantly different from that found for indole acetic acid,[16] ℓ-tryptophane hydrobromide,[15] and glycyl-ℓ-tryptophane.[14]

In the pyrrole portion of the indole ring in 5-methoxytryptamine, the C-N distances C(9)-N(2) = 1.374(3), C(10)-N(2) = 1.368(3) are equivalent. They are shorter than the average for the pyrrole ring by six standard deviations. The C(2)-C(10) bond is also shorter than the average by six standard deviations, while the C(3)-C(4) bond is larger by 12 standard deviations. The average for the pyrrole ring is 1.391 ± .003 Å, and the value agrees within the standard deviations with other x-ray work.[14,15,16]

There was disagreement among earlier workers concerning the length of the C-N bonds in the pyrrole ring. The C(9)-N(2) bond of 1.31 ± .02 Å in glycyl-ℓ-tryptophane was found to be significantly shorter than the corresponding bond in ℓ-tryptophane hydrobromide, and ℓ-tryptophane hydrochloride (1.43 ± .02 Å).[21] The length of this bond in 5-methoxytryptamine is 1.374 ± .003 Å. This is near the average (1.37 ± .02 Å) obtained by averaging the results of the disagreeing workers. Pauling believed the C-N distances in pyrrole were equivalent, and calculated a value of 1.42 ± .02 Å for them,

based on electron diffraction data.[22] The angular geometry of the indole ring in 5-methoxytryptamine is not significantly different from the results found in the structure determinations mentioned above. All intramolecular angles not involving hydrogen in 5-methoxytryptamine are shown in Table III-11.

The C(5)-C(6), C(7)-C(8), N(2)-C(9), N(2)-C(10), and C(3)-C(10) bonds of the indole ring in 5-methoxytryptamine are all six deviations or more shorter than the average for the ring. If this bond-shortening effect is due to a concentration of π electrons, we should expect molecular orbital calculations to give high electron densities in the regions near these bonds. The frontier electron density, which correlates with π electron density, has been calculated for 5-methoxytryptamine and tryptamine with the approximations of Hueckel molecular orbital theory.[23] The concentration of π electron density, listed in decreasing order, is C(10) > C(3) > C(5) > C(7) > N(2) > C(8). These regions of high electron density do not correlate with any long or average indole ring bond in 5-methoxytryptamine as determined by the x-ray work. Furthermore, the shortest bonds as found by the x-ray work are C(3)-C(10) and N(2)-C(10); each bond is 1.368 ± .004 Å. These regions correspond to the highest π electron concentration as calculated by the Hueckel method. The x-ray work and the molecular orbital calculations for 5-methoxytrypta-

Table III-11

Intramolecular Angles (in degrees) and their Standard Deviations (a) for all Non-hydrogen Atoms in 5-Methoxytryptamine.

Atoms	Angles	Atoms	Angles
N(1)-C(1)-C(2)	115.04(.25)	C(5)-C(6)-O(1)	123.84(.20)
C(1)-C(2)-C(3)	112.98(.23)	C(7)-C(6)-O(1)	114.51(.25)
C(2)-C(3)-C(4)	127.57(.25)	C(6)-C(7)-C(8)	121.31(.25)
C(2)-C(3)-C(10)	126.63(.27)	C(7)-C(8)-C(9)	117.79(.24)
C(4)-C(3)-C(10)	105.80(.21)	C(8)-C(9)-C(4)	121.40(.22)
C(3)-C(4)-C(5)	132.98(.21)	C(8)-C(9)-N(2)	130.29(.24)
C(3)-C(4)-C(9)	106.89(.20)	C(4)-C(9)-N(2)	108.30(.20)
C(9)-C(4)-C(5)	120.13(.21)	C(9)-N(2)-C(10)	108.04(.22)
C(4)-C(5)-C(6)	117.70(.21)	N(2)-C(10)-C(3)	110.97(.22)
C(5)-C(6)-C(7)	121.65(.23)	C(6)-O(1)-C(11)	117.70(.21)

(a) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

mine seem to agree in this respect. The average bond length for the 5 ring hydrogens is .99 Å. The difference between this average bond length and the most accurate benzene C-H distance, 1.084 ± .005 Å,[20] from electron diffraction work, is consistent with the expected magnitude of effects of thermal motion and concentration of electrons in the bond. These effects are well known to cause the x-ray determinations for bonds to hydrogen to be shorter than the average internuclear distance. All hydrogen distances are reported in Table III-10. The standard deviations on hydrogen bond lengths range from .02 Å on the best determined ring hydrogens, to .05 Å on the primary amine hydrogens of N(1). All thermal parameters on the hydrogens are reasonable, and have been set forth earlier in Table III-7. The ring hydrogens do not significantly depart from the plane of the indole ring. The hydrogens H(8) and H(9) of the puckered portion of the ring bend upward with the two carbons, but the effect is not significant within the hydrogen standard deviations on bond lengths. The angular geometry of the ring hydrogens is consistent with other accurate work and is shown in Table III-12. Interestingly enough, the average length of the other nine hydrogens was also .99 Å.

In the ether side chain at C(6), rotation of the methyl group about the carbon-oxygen single bond may be hindered by C(5). The C(11)-C(5) distance is 2.797 Å, and the H(7)-H(13)

distance is 2.252 Å. The ether side chain bends down below the indole ring and points in the general direction of C(5). The methyl carbon C(11) is 1.04 Å below the indole ring, and the ether side chain makes an angle of 19.8 degrees with the indole plane. The internal geometry of this side chain is about what one would expect, based on reported values. The C(6)-O(1) bond (1.376 ± .003 Å), and the C(11)-O(1) bond (1.416 ± .003 Å) are consistent with values found in the International Tables; 1.36 ± .01 Å for shortened oxygen-carbon distance due to the influence of an aromatic ring, and 1.43 ± .01 Å for aliphatic carbon-oxygen distance.[24]

The C(6)-O(1)-C(11) bond angle in 5-methoxytryptamine is 117.7 degrees. This compares well with the angles in 1,4-dimethoxy benzene (121°) within the standard deviations of the determinations.[25]

The side chain at C(3) in 5-methoxytryptamine may be thought of as a substituted ethane. Possible conformations for ethane, or better, n-butane, are shown in Figure III-6. The indole ring may be in the extreme case either cis- or trans- to N(1) of the side chain. The rotation angle about the ethane bond for the substituents is 0° for the cis- and 180° for the trans- position. The conformation of n-butane may be either eclipsed, with end-on projection of the C-C bond showing only three substituents, or staggered, with all six substituents visible.

Table III-12

Intramolecular Angles (in degrees) and their Standard Deviations (a) for all Bonds Involving Hydrogen in 5-Methoxytryptamine.

Atoms	Angles	Atoms	Angles
C(1)-N(1)-H(1)	102.6(2.3)	C(6)-C(7)-H(8)	116.0(1.6)
C(1)-N(1)-H(2)	104.1(2.3)	C(8)-C(7)-H(8)	122.7(1.6)
H(1)-N(1)-H(2)	118.6(5.0)	C(7)-C(8)-H(9)	121.3(1.9)
C(2)-C(1)-H(3)	107.9(1.4)	C(9)-C(8)-H(9)	120.9(1.9)
C(2)-C(1)-H(4)	109.4(1.7)	C(9)-N(2)-H(10)	132.4(1.8)
N(1)-C(1)-H(3)	108.8(1.3)	C(10)-N(2)-H(10)	119.6(1.8)
N(1)-C(1)-H(4)	106.8(1.5)	C(3)-C(10)-H(11)	126.5(1.8)
H(3)-C(1)-H(4)	108.8(3.0)	N(2)-C(10)-H(11)	122.5(1.8)
C(1)-C(2)-H(5)	107.1(1.5)	O(1)-C(11)-H(12)	103.4(1.5)
C(1)-C(2)-H(6)	108.8(1.8)	O(1)-C(11)-H(13)	111.8(1.7)
C(3)-C(2)-H(5)	108.5(1.4)	O(1)-C(11)-H(14)	115.0(1.8)
C(3)-C(2)-H(6)	109.6(1.8)	H(12)-C(11)-H(13)	116.2(3.3)
H(5)-C(2)-H(6)	109.9(2.9)	H(12)-C(11)-H(14)	109.2(3.4)
C(4)-C(5)-H(7)	121.5(1.4)	H(13)-C(11)-H(14)	101.7(3.2)
C(6)-C(5)-H(7)	120.8(1.4)		

(a) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

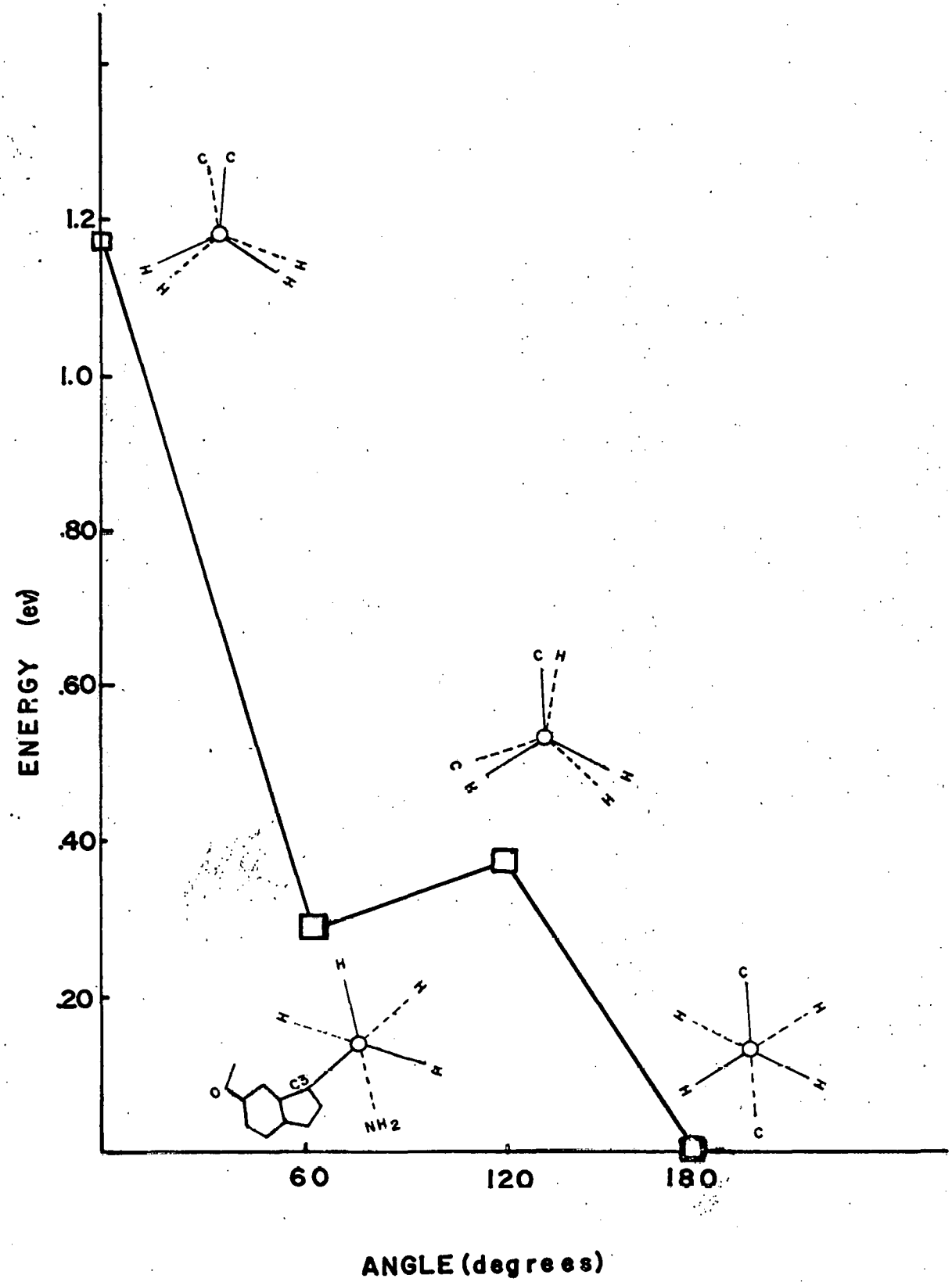
The conformation energies of n-butane were calculated by Hoffman using extended Hueckel molecular orbital theory.[26] This energy curve is also shown in Figure III-6. The minimum is for the trans- isomer. The maximum is for the cis- isomer. The energy separation is 1.2 ev or 27.7 kcal/mole. Notice that there is also a minimum with a gauche, staggered conformation at a rotation angle of about 60 degrees. The energy separation between this conformer and the trans- minimum is .26 ev or about 5.99 kcal/mole.

The rotation angle between the terminal nitrogen N(1) of the side chain and carbon C(3) of the indole ring in 5-methoxytryptamine is 68.3 degrees. The conformation is thus staggered with a gauche orientation. The conformation of 5-methoxytryptamine in the solid state falls into a higher energy minimum than the preferred trans- arrangement calculated for the isolated molecule.

Kier has made a similar extended Hueckel molecular orbital conformational analysis of serotonin.[27] He found the lowest energies associated with the trans- configuration of the terminal nitrogen N(1) and the carbon C(3). The configuration in 5-methoxytryptamine, as mentioned earlier, is a gauche, staggered arrangement with the rotation angle between N(1) and C(3) about the ethane bond of 68.3 degrees. This conformation corresponds to an energy of about .3 ev or 6.9 kcal/mole higher than the most stable, trans- conformation. This estimate of the energy separation

Figure III-6

Conformation Energies for the n-Butane, and Conformation Energy of the Primary Amine Side Chain in 5-Methoxytryptamine.


ENERGY (ev)
1.2
1.0
.80
.60
.40
.20
60
120
180
ANGLE (degrees)
C
C
H
H
H
C
H
C
H
H
H
H
H
H
O
C3
NH2
C
H
H
H
H
C

XBL 704-831

between the most stable conformation and that found in 5-methoxytryptamine is 1.0 kcal/mole higher than the analysis based on n-butane.

As mentioned above, the energy calculations were made using extended Hueckel molecular orbital theory. This theory works best for aromatic systems, but is only semi-quantitative for aliphatic systems. For aliphatic systems, energy barriers are always estimated too high. The estimates are best for compounds such as ethane, where very few terms, due to steric repulsion, are added to the energy calculation. The energy gap between the eclipsed and staggered positions of ethane is calculated to be 4.0 kcal/mole, and is actually observed to be 25 to 33 per cent less. This over-estimation of energy separations tends to increase with the complexity of the molecule until qualitative as well as quantitative errors occur with the branched pentanes.[26]

Kier also calculated the probable rotation angle of the ethane bond with the indole ring. For the purpose of this discussion we define the angle ϕ such that when $\phi = 0$ the ethane bond of the side chain is parallel to the plane of the indole ring and pointing generally toward the benzene ring. When $\phi = 0$, serotonin has the highest calculated conformational energy. When $\phi = 90°$, with the terminal nitrogen in the trans- staggered position, serotonin has the lowest conformational energy. The energy separation between $\phi = 0$

and $\phi = 90°$ is 1.5 ev or 34.6 kcal/mole when the terminal nitrogen N(1) is in the lowest energy, trans- conformation.

The angle ϕ in 5-methoxytryptamine is 35.8 degrees. The ethane bond points in the general direction of the benzene ring and is rotated 35.8 degrees below the indole plane. This bond conformation corresponds to an energy about .3 ev or 6.92 kcal/mole higher than the minimum with $\phi = 90°$ and the ethane bond perpendicular to the ring. Again, we must realize that these calculated energy barriers are calculated too high by 25 to 33 percent in the most favorable case, that of ethane.

If we assume that the conformation energies involved are independent and additive, 5-methoxytryptamine assumes a conformation in the solid state about 13.8 kcal/mole higher than the most stable conformation. If we use the more reasonable value for the conformation energy, that based on n-butane, the energy separation is 12.8 kcal/mole. We can compare this to the highest conformation energy that 5-methoxytryptamine could assume according to these calculations, 62.3 kcal/mole above the most stable configuration. The solid-state conformation energy of 5-methoxytryptamine is 20 percent of the highest conformation energy possible according to extended Hueckel molecular orbital theory. Since the energies for ethane were calculated 25 to 33 percent high, and the percentage error tends to increase in an upward

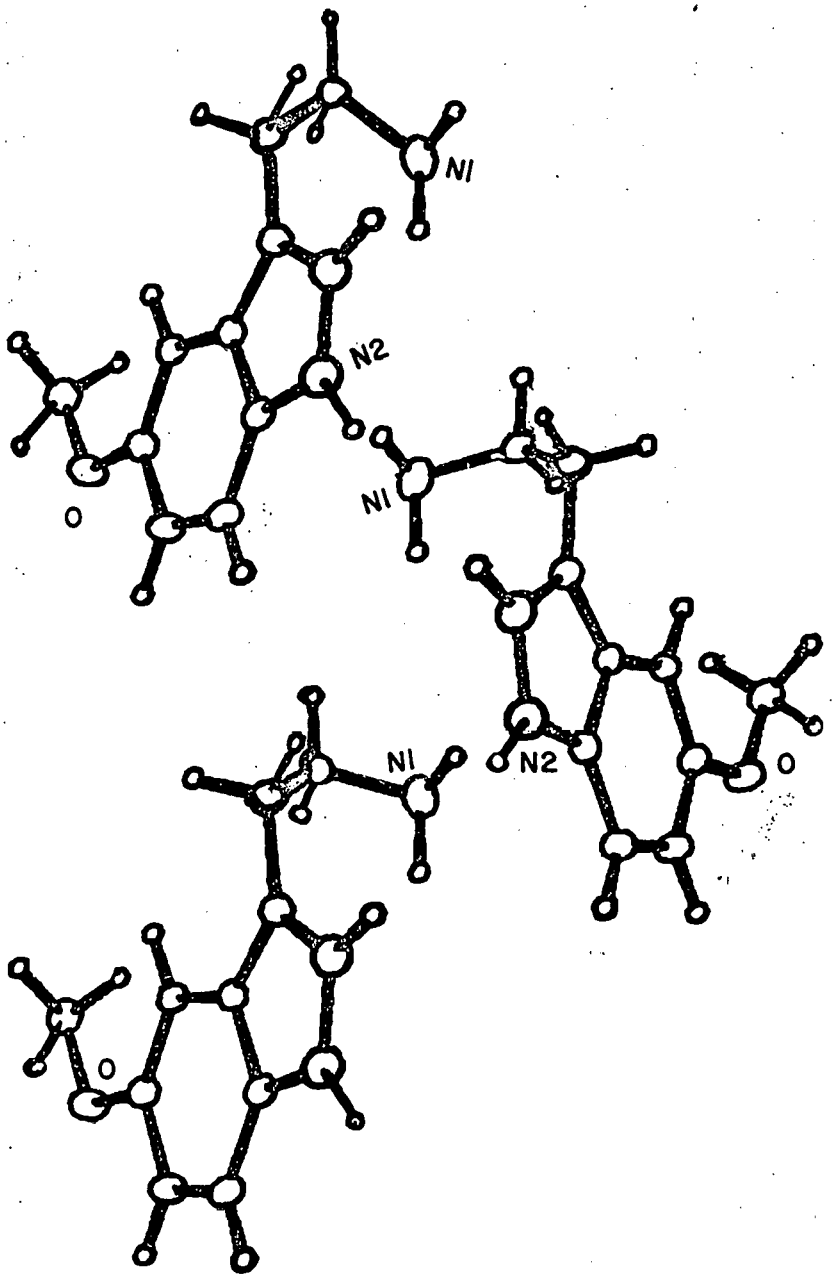
direction with the number of steric repulsions, we should probably reduce the conformational separation energy of 12.8 kcal/mole calculated for serotonin by more than 33 percent.

The 5-methoxytryptamine molecule moves to the higher energy conformer in the solid state partially in order to complete an intermolecular hydrogen bond. The indole ring makes an angle of 35.6 degrees with the ac plane, and the distance between translation equivalent indole rings is 3.50 Å. The nitrogen side chain bends below the plane of the ring, with the conformations mentioned earlier, toward the ac plane. Hydrogen bonds are formed with the nitrogen N(2) of the glide-related molecule translated one unit cell distance along a, (1 + x, -y, 1/2 + z). The hydrogen on the ring nitrogen N(2) is donated to the primary amine nitrogen N(1) on the side chain. The intermolecular distance between N(1) and N(2) is 2.916 ± .003 Å, and is the closest intermolecular approach in the crystal structure of 5-methoxytryptamine. The angular environment about the hydrogen-bonded atom N(1) is given in Table III-13, and Figure III-7.

In crystalline ammonia, the N-H-N distance is 3.38 Å. The energy of the N-H-N hydrogen bond in ammonia is estimated to be in the range 1.3 to 3.8 kcal/mole. The N-N approach distance in 5-methoxytryptamine is one of the shortest yet reported for this kind of bond. This distance of 2.916 Å is shorter than the N-H-N bond in Adenine-HCl, 2.99 Å; ammonium

Figure III-7

Hydrogen Bonding in 5-Methoxytryptamine.



XBL 704-803

Table III-13

Angles (in degrees) and Standard Deviations (a) about the Hydrogen-bonded Atom N(1).

The Hydrogen H(10) is donated by the Nitrogen N(2) in the symmetry-related molecule one unit cell away. Other atoms are in the same molecule as N(1).

Atoms	Angles
H(1)-N(1)-H(2)	118.6(5.0)
H(10)-N(1)-H(1)	122.2(4.0)
H(10)-N(1)-H(2)	93.9(4.6)
C(1)-N(1)-H(10)	111.6(1.1)
C(1)-N(1)-H(1)	102.6(2.3)
C(1)-N(1)-H(2)	104.1(2.3)
N(2)-H(10)-N(1)	171.0(2.3)

(a) Least squares estimations of the least significant digits are given in parentheses.

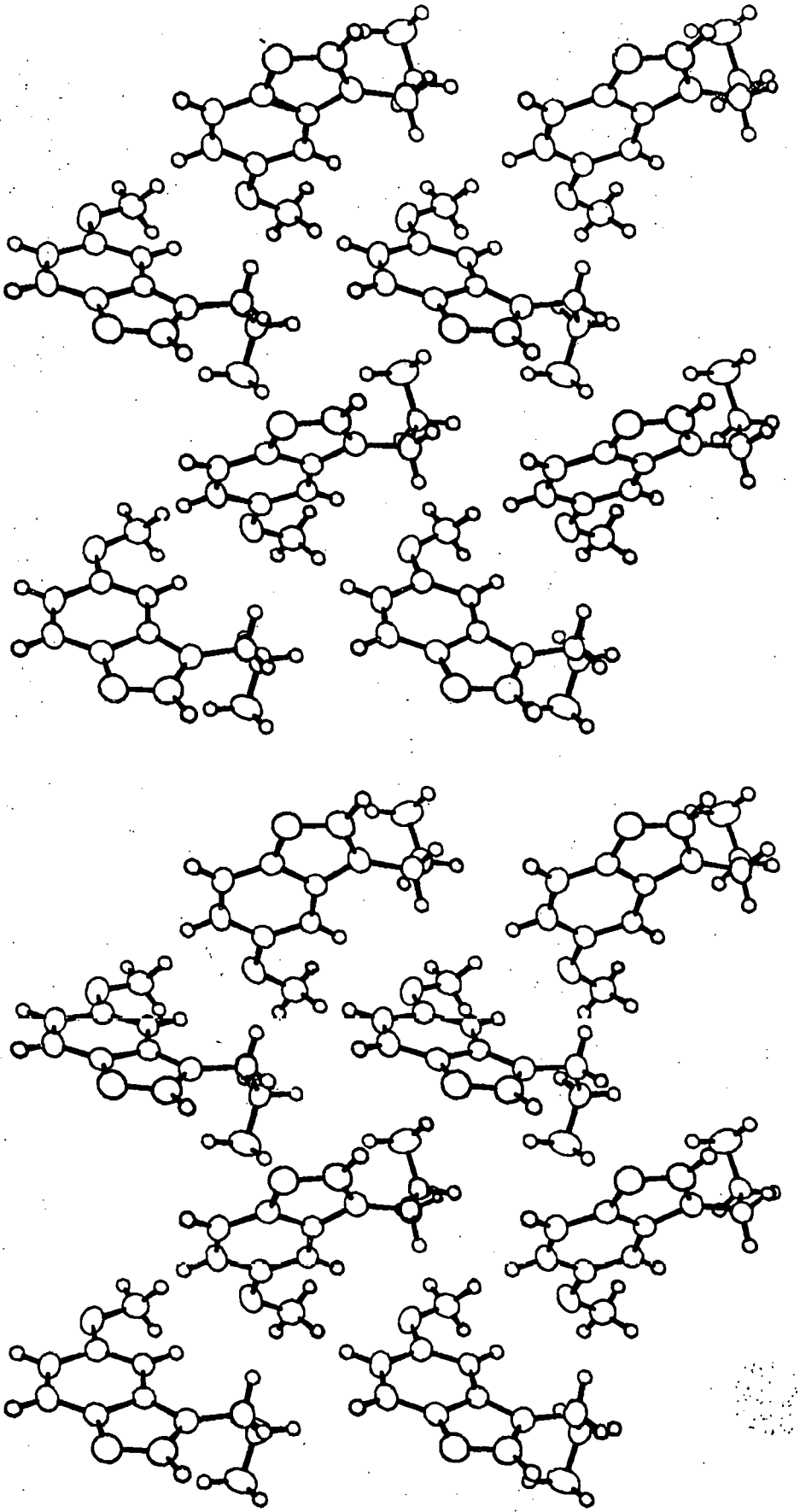
azide and dicyandiamide, 2.94 Å.[28] In view of the short N-H-N bond distance, an energy on the order of 3.0 kcal/mole is a reasonable estimate of the strength of the N-H-N bond in 5-methoxytryptamine. For comparison, the O-H-O hydrogen bond, in ice, is 5.0 kcal/mole, and the O-H-O bond in acetic acid is 7.6 kcal/mole.[29]

From these considerations, we can say that up to 30 per cent of the excess conformational energy in 5-methoxytryptamine can be accounted for by the formation of hydrogen bonds. Hueckel molecular orbital theory, then, overestimates the conformational energy separations in 5-methoxytryptamine by a value of between 33 and 70 percent.

The super-structure of 5-methoxytryptamine can be described as pleated sheets stacked along the a direction, held together by N-H-N hydrogen bonding between the sheets. Distance between equivalent planes is 3.50 Å. The nitrogen N(1) is 3.30 Å below the plane of the ring to which it is attached, and is within 2.916 Å of N(2) in the glide-related molecule, translated one unit cell length along a. The N-H-N hydrogen bond has a reasonable geometry as we can see in Table III-13. A least squares plane through this intermolecular N(2)-H-N(1) bond makes an angle of 60.5 degrees with the plane of the indole ring of N(2). The intermolecular packing is shown in Figure III-6. The view is down the a axis. The direction of the glide plane is along c.

Figure III-8

Intermolecular Packing in 5-Methoxytryptamine.



XBL 704-833

The closest intermolecular approach in 5-methoxytryptamine is between C(5)-O(1), 3.405 Å. The oxygen is in the glide-related molecule at (x, -y, 1/2 + z). Other close approaches are between O(1) of the methoxy side chain in the molecule (1 + x, y, z), and C(8) of the molecule (x,y,z). The distance involved is 3.759 Å. The C(7) in molecule (x,y,z) is within 3.734 Å of C(5) in the glide-related molecule at (x, -y, 1/2 + z).

E. Conclusion

The indole ring in 5-methoxytryptamine is not planar within the standard deviations of the structure determination. The carbons C(7) and C(8) are warped out of the plane of the indole ring by 1.6°. No explanation of this is obvious from a consideration of molecular packing or conformation. Short bonds in the indole ring correlate with regions of high π electron density as calculated by Hueckel molecular orbital theory.

The nitrogen of the aliphatic side chain forms a very strong N-H-N hydrogen bond with the nitrogen of the indole ring in the glide related molecule at (1+x,-y,1/2+z). Formation of this hydrogen bond partially compensates for the large energy separation between the minimum energy conformation and the conformation which actually occurs in 5-methoxytryptamine.

The approximations of Hueckel molecular orbital theory may substantially overestimate the energy separations involved.

Bibliography

1. M. M. Rapport, A. A. Green, and I. H. Page, Science, 108, 319 (1948).
2. Boris Sokoloff, Recent Results in Cancer Research, 15, 1-19 (1968).
3. Irvine H. Page, in "Serotonin," Year Book Medical Publishers, Chicago, 1968.
4. L. F. Semenov, Nucl. Sci. Abstr., 18 (14), Abstr. No. 23385 (1964).
5. A. N. Pisarevskii, A. F. Seleznev, G. M. Pashek, Radiobiologica, 5 (5), 768-70 (1965).
6. Robert C. Lin, S. H. Ngai, E. Costa, Science, 166, 238 (1969).
7. Alessandro Tagliamonte, Paola Tagliamonte, Gian L. Gessa, Bernard B. Brodie, Science, 166, 1434 (1969).
8. J. Mouret, Pierre Bobillier, Michael Jouvet, Eur. J. Pharmacol. 5 (1), 17-22 (1968).
9. S. Garattini, P. A. Shore, E. Costa, and M. Sandler (eds.), Advances in Pharmacology, Academic Press, New York, 1968, Vol. 6A.
10. S. Garattini, L. Valzelli, "Serotonin," Elsevier, Amsterdam, 1965.
11. D. W. Wooley, "The Biochemical Bases of Psychoses or the Serotonin Hypothesis about Mental Diseases," John Wiley, New York, 1962.

12. V. Erspamer, Recent Research in the Field of 5-hydroxytryptamine and Related Indole Alkylamines, in "Progress in Drug Research," E. Jucker (ed.), Interscience Publishers, New York, (1961).

13. C. K. Johnson, "Ortep," A Fortran Thermal-Ellipsoidal Plot Program for Crystal-Structure Illustrations, ORNEL-3794, Revised, Oak Ridge, Tenn., 1965.

14. R. A. Pasternak, Acta Cryst., 9, 341 (1956).

15. R. Ramachandra Ayyar, and R. Chandrasekharan, Curr. Sci., 6, 139 (1967).

16. Isabella L. Karle, Kathleen Britts, and P. Gum, Acta Cryst., 17, 496 (1964).

17. D. W. J. Cruikshank, Acta Cryst., 2, 65 (1949).

18. Mark S. Fischer, University of California Phd. Thesis, UCRL-19524, (1969).

19. Michael G. B. Drew, Acta Cryst., B25, 1320 (1969).

20. B. P. Stoicheff, Canad. J. Phys., 32, 339 (1954).

21. T. Takigawa, Y. Sasada, and M. Kakudo, Bull. Chem. Soc. Japan, 39, 2369 (1966).

22. V. Schomaker, and L. Pauling, J. Amer. Chem. Soc., 61, 1769 (1939).

23. Carl R. Merril, Solomon H. Snyder, and Dan F. Bradley, Biochem. Biophys. Acta, 118 (2), 316-24 (1966).

24. International Tables for X-ray Crystallography, 3, 276, Kynoch Press, Birmingham, England, 1962.

25. T. H. Goodwin, M. Przykylaska, and J. M. Robertson, Acta Cryst., 3, 279 (1950).

26. Roald Hoffmann, J. Chem. Phys., 39, 1397 (1963).

27. Lemont B. Kier, J. Pharm Sci., 57 (7), 1188-91 (1968).

28. J. Monteath Robertson, "Organic Crystals and Molecules," Cornell Univ. Press, Ithaca, N.Y., 1953.

29. Linus Pauling, "Nature of the Chemical Bond," Cornell Univ. Press, Ithaca, N. Y., 1960.

Section IV

The Crystal and Molecular Structure of Melatonin.

A. Introduction

Melatonin is formed in the pineal gland of the human brain by N-acetylation of serotonin.[1] This reaction is followed by methylation with hydroxyindole-Omethyltransferase.[2] The formation of melatonin is an added complication in the complicated biochemistry of serotonin. Serotonin which is acetylated is protected from monoamine oxidase destruction. The compound N-acetylserotonin is excreted in human urine.[3]

The crystal structure of melatonin was undertaken in order to provide a sound structural basis for molecular orbital calculations. It was hoped that some clue would be provided by the x-ray structure which would help establish the role molecular conformations have in the biochemical actions of serotonin and melatonin.

B. Experimental

The crystals of melatonin used in this structure determination were obtained as described in III-B. Preliminary oscillation and Weissenberg photographs of an unmeasured crystal had symmetry and extinctions consistent with monoclinic space group $P2_1/n$. (0k0, k = 2n; h0ℓ, h + ℓ = 2n). It was subsequently found that a better choice of cell could be made, and that the preliminary photographs were taken with the crystal aligned along the $\underline{a}^*$ axis of a $P2_1/c$ unit cell. The crystal structure of melatonin was solved in $P2_1/c$, and all further references to this problem will be based on this space group.

Rough cell dimensions were measured from the films of this crystal and another crystal (.35 x .27 x .10 mm) mounted on a thin glass fiber with General Electric Clear Industrial Varnish 1202 in a random orientation was transferred to the Picker Automatic Diffractometer for the accurate determination of cell dimensions and intensity measurements. The Picker Automatic Diffractometer was physically the same as outlined in Section III-B, except that a molybdenum x-ray tube was used, and the graphite monochromator was set at 12.048°.

Least squares cell dimensions were calculated on the PDP-8I computer from the 2θ values of 13 well-centered reflections using the standard program furnished by the Picker

Corporation. The molybdenum cell dimensions for $P2_1/c$ were $\underline{a}$ = 7.707 ± .002; $\underline{b}$ = 9.252 ± .002; $\underline{c}$ = 17.077 ± .004; β = 96.78 ± .03; λ = .709261; z = 4.

Calculated density based on these cell dimensions was 1.276 g/cc^3, and the experimental density as measured at room temperature in ethylene chloride, ethylene bromide, and ethyl acetate was 1.272 g/cc^3, which agrees well with the calculated value. The linear absorption coefficient, μ, for melatonin with molybdenum x-rays was .941 cm^{-1} and μt in the longest direction for this crystal (.35 x .27 x .10 mm) was .033. No absorption correction was made.

Intensity data was taken throughout the $+\underline{k}$ region of reciprocal space in planes of constant $\underline{h}$ from $-\underline{h}$ to $+\underline{h}$. All the molybdenum data in this hemisphere was measured from 2θ = 0 to 2θ = 40° ($\sin\theta/\lambda$ = .48222). The scan width, s, was 2.0 degrees, and data was taken at a 2.0 degree take-off angle. Definition of s and the diffractometer scan procedure was given in III-B.

The 006 and 206 reflections were chosen as intensity standards. The 206 reflection was measured 49 times for an average intensity of 21552 and an average deviation of 393. The 006 reflection was measured 48 times for an average intensity of 24313 and an average deviation of 315.

A total of 2275 reflections were measured. The $45\bar{4}$ reflection was measured only once due to a machine error, and

4 0k0 reflections were measured only once. There were two measurements of all other independent reflections as the Friedel-equivalent pairs $hk\ell$, $\bar{h}k\bar{\ell}$; and $\bar{h}k\ell$, $hk\bar{\ell}$ were measured. Thus, the 2275 measurements were reduced to 1140 independent observations after averaging.

The form of the temperature factors used and the weighting scheme for least squares refinement were given in III-B. The source of scattering factors used was given in I-B. The $\Delta f'$ and $\Delta f''$ anomalous dispersion corrections for the light atoms in melatonin were zero.

C. Solution of the Structure

The crystal structure of melatonin was solved with a direct method of phase determination. A Wilson plot of the data was prepared using a computer program written by Maddox and Maddox.[4] The overall temperature factor, 1.09 $Å^2$, calculated from the Wilson plot, was used by the program in the calculation of the normalized structure factors Ehkl. The 151 Ehkl $\geq$ 1.50 were tested for sign interactions with a Fortran computer program written by Michael Drew. Three reflections corresponding to a large number of interactions, high Ehkl values, and linear independence modulo two were chosen to determine the origin. These reflections were arbitrarily given positive signs. Once the origin was set, all the other signs were functions of the structure alone. A starting set of four reflections with a large number of sign interactions and large Ehkl values were given all possible combinations of plus and minus signs resulting in 16 possible solutions to the structure of melatonin. Signs for the other Ehkl values based on these 16 possible combinations of the starting set were calculated with iterative application of Sayre's equation,[5] which results also from the Σ_2 relationship of Hauptman and Karle.[6] This equation was applied to the 151 normalized structure factors Ehkl using a computer program written by R. E. Long.[7]

Sayre's equation may be expressed as

$$s(E_{\underline{A}}) = s(\sum_{\underline{A}=\underline{B}+\underline{C}} E_{\underline{B}} \cdot E_{\underline{C}})$$

where s( ) means "sign of", the dot implies multiplication, and $\underline{A}$, $\underline{B}$, and $\underline{C}$ are the vectors (hkl) for the reflections A, B, and C; the sum is over all combinations where $\underline{B} + \underline{C} = \underline{A}$.

The three origin determining reflections and the four structure invariant reflections which were chosen for sign permutation are given below, the origin determining reflections are listed first:

h	k	l	Ehkl	M
3	1	11	3.09	34
3	1	8	2.02	32
1	2	6	2.46	38
3	1	5	2.52	30
2	1	-2	5.93	49
2	1	-1	4.11	45
1	4	-3	2.02	32

The quantity M is tne number of sign interactions for the reflection involved, and M ranged from 8 to 51 for the 151 reflections considered.

Selection of the correct combination of signs from the 16 sets produced by Long's program is aided by a consistency index, C, defined as

$$C = \frac{\{|E_{\underline{A}} \sum_{\underline{A}=\underline{B}+\underline{C}} E_{\underline{B}} E_{\underline{C}}|\}}{\{|E_{\underline{A}}|\sum_{\underline{A}=\underline{B}+\underline{C}}|E_{\underline{B}}||E_{\underline{C}}|\}}$$

where the sums are over all pairs $\underline{B}$ and $\underline{C}$ for which $\underline{B}+\underline{C} = \underline{A}$, and where { } means the average over all values of $\underline{A}$.

A totally consistent solution has a consistency index of 1.0.

The program repeatedly applies Sayre's equation in cycles through the list of Ehkl such that signs predicted at the top of the list are used in the prediction of signs below until there are no changes or additions to the list. The solution with the correct combination of signs usually correlates with the highest consistency index, and the fewest number of cycles.

Although this correlation does not always hold, for melatonin it happened to be true. Of the 16 possible solutions, consistency indices ranged from .50 to .80, and the number of cycles needed to produce an unchanging set of signs ranged from 7 to the maximum of 17. Set number 11 had a consistency index of .80, and sign combinations for the

151 terms was calculated in 7 cycles. For this solution, the four signs of the starting set were positive except for the 315, and $31\bar{1}$ reflections.

A Fourier was calculated using as terms the 151 Ehkl with the signs predicted by Long's program. The top 17 peaks on this Fourier corresponded to all the non-hydrogen atoms of the melatonin structure. A full matrix least squares refinement of these 17 atoms with isotropic temperature factors gave a conventional R factor of 11.6%, a weighted R, R_2, of 13.4%, and a standard deviation of observation unit weight of 3.417. An examination of the signs of the calculated structure factors corresponding to this refined structure revealed that Sayre's equation and Long's computer program had correctly predicted every single one of the 151 signs.

This structure was refined with anisotropic temperature factors through four cycles of least squares refinement. The result of this refinement was an R factor of 9.1%, a weighted R of 10.4%, and a standard deviation of 2.881.

A difference Fourier phased on the refined atomic positions of the 17 non-hydrogen atoms from the anisotropic refinement revealed the positions of the hydrogen atoms. The top 19 peaks on this difference map ranged from .26 to .51 electrons, and the 16 hydrogen atoms of melatonin were among these 19 peaks.

Four cycles of least squares refinement giving the non-hydrogen atoms anisotropic, and the hydrogen atoms isotropic temperature factors resulted in an R factor of 3.45%, a weighted R of 3.4%, and a standard deviation of observation unit weight .995. All the refinements mentioned above were done with the weighting scheme outlined in IV-B Since the standard deviation at this point dropped below one, the intense data were overweighted for systematic errors, and the coefficient p was lowered to .04.

With this revised weighting scheme the above refinement was repeated. The R factor this time was 3.45% for 808 data; the weighted R was 3.3%, the standard deviation was 1.106, and the conventional R including zero weight data was 5.71% for the 1140 independent reflections. There was no systematic trend in either $|Fo/Fc|$ or $\underline{w}^{1/2}\ \Delta F$ as a function of intensity or Bragg scattering angle. The largest peak on the final difference Fourier was .19 electrons. In this final refinement no parameter shifted by more than 1% of its estimated standard deviation.

D . Discussion of the Structure

The atomic coordinate of all the non-hydrogen atoms in melatonin are given in Table IV-1, and the thermal parameters are listed in Table IV-2. The coordinates of the hydrogen atoms and their thermal parameters are given in Table IV-3. The atomic numbering system is entirely consistent with that of 5-methoxytryptamine in Section III-D. For this reason, numbering of the hydrogens in melatonin is not entirely sequential. Since there is only one hydrogen on the nitrogen, N(1), in the side chain of melatonin, and two in 5-methoxytryptamine, the label H(2) for the second hydrogen on this nitrogen has been omitted for melatonin. This numbering system, with the interatomic distances in melatonin, is presented in Figure IV-1. Interatomic distances are given along with their standard deviations in Table IV-4.

The average of all the bond distances of the benzene ring is 1.386Å. This average is not significantly different from the result found for 5-methoxytryptamine. As in 5-methoxytryptamine, bonds C(5)-C(6) and C(7)-C(8) are significantly shorter than the average. The bond C(4)-C(5) is significantly larger than the average. The definition of significance used and the probable cause of the bond-shortening have been adequately discussed in III-D.

The average of all the bonds in the pyrrole ring is

Table IV-1

Atomic Coordinates and their Standard Deviations (a) for all Non-hydrogen Atoms in Melatonin.

ATOM	X	Y	Z
C(1)	.4374(6)	.3060(4)	.1265(2)
C(2)	.6292(6)	.6757(4)	.4202(2)
C(3)	.3187(4)	.2059(3)	-.0055(2)
C(4)	.2388(4)	.3953(3)	.4384(2)
C(5)	.8116(4)	.0403(4)	.0554(2)
C(6)	.1182(4)	.6102(4)	.3777(2)
C(7)	.9049(5)	.0394(5)	.1950(2)
C(8)	.1418(5)	.3979(5)	.2971(2)
C(9)	.2138(4)	.3265(4)	.3650(2)
C(10)	.3377(5)	.1702(4)	.4541(2)
C(11)	.9128(7)	.3303(5)	.0534(3)
C(12)	.5977(4)	.3575(4)	.2551(2)
C(13)	.3268(7)	.7993(5)	.1667(2)
N(1)	.4906(4)	.7658(3)	.2944(2)
N(2)	.7280(4)	.6876(4)	.1235(2)
O(1)	.0660(3)	.7531(3)	.3751(2)
O(2)	.3862(3)	-.0144(2)	.2630(1)

(a) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

Table IV-2

Table of Anisotropic Temperature Parameters (a) and their Standard Deviations (b) in Melatonin.

ATOM	B11	B22	B33	B12	B13	B23
C(1)	5.9(2)	2.7(2)	3.1(2)	-.1(2)	.2(2)	.1(1)
C(2)	4.1(2)	3.0(2)	3.5(2)	.2(2)	.4(2)	.1(2)
C(3)	3.9(2)	2.6(2)	2.8(2)	-.2(1)	.5(1)	.3(1)
C(4)	3.3(2)	2.7(2)	3.0(2)	-.5(1)	.5(1)	-.0(2)
C(5)	4.1(2)	3.1(2)	3.3(2)	.4(1)	.6(2)	.3(2)
C(6)	4.1(2)	3.3(2)	4.1(2)	.2(2)	.3(2)	.3(2)
C(7)	4.6(2)	4.6(2)	3.7(2)	-.0(2)	-.4(2)	-1.1(2)
C(8)	4.6(2)	4.9(2)	3.0(2)	-.6(2)	.2(2)	-.2(2)
C(9)	3.9(2)	2.8(2)	3.5(2)	-.1(1)	.3(1)	-.3(2)
C(10)	4.7(2)	3.4(2)	3.3(2)	.1(2)	.4(2)	.3(2)
C(11)	6.0(3)	3.2(2)	7.1(3)	-.5(2)	.4(3)	.1(2)
C(12)	4.0(2)	2.9(2)	3.3(2)	-.4(1)	.9(1)	-.5(2)
C(13)	5.9(3)	4.1(2)	3.5(2)	.3(2)	-.0(2)	-.3(2)
N(1)	5.8(2)	1.9(1)	3.1(2)	.4(1)	.4(1)	-.4(1)
N(2)	5.6(2)	3.4(2)	3.6(2)	.0(1)	.8(1)	1.2(2)
O(1)	6.9(2)	3.6(1)	5.6(2)	1.4(1)	.1(1)	.7(1)
O(2)	7.3(2)	2.3(1)	4.0(1)	.6(1)	.9(1)	.07(9)

(a) Anisotropic thermal parameters, $\underline{B}$, in units of A^2, are given by $\underline{B} = 4\beta_{ij}/\underline{a}^*_i\underline{a}^*_j$, where $\underline{a}^*_i$ is the $\underline{i}$th reciprocal cell length.

(b) Estimated standard deviations are given in parentheses following the parameter.

Table IV-3

Final Positional Parameters and Isotropic Thermal Parameters (a) and their Standard Deviations (b) in Melatonin for all the Hydrogen Atoms.

ATOM	X	Y	Z	B
H(1)	.502(3)	.680(3)	.280(1)	1.9(7)
H(3)	.656(4)	.879(3)	.370(2)	4.7(8)
H(4)	.535(4)	.355(3)	.103(2)	4.9(9)
H(5)	.458(4)	.103(4)	.088(2)	4.6(9)
H(6)	.273(4)	.136(3)	.103(2)	3.8(8)
H(7)	.798(4)	.410(3)	.506(2)	3.5(7)
H(8)	.043(4)	.591(3)	.260(2)	4.3(8)
H(9)	.124(4)	.352(4)	.246(2)	5.3(9)
H(10)	.282(4)	.131(3)	.343(2)	2.2(8)
H(11)	.623(4)	.577(4)	.027(2)	4.6(9)
H(12)	.022(5)	.713(4)	-.011(2)	7.2(12)
H(13)	-.045(5)	.423(4)	.068(2)	6.1(10)
H(14)	.210(5)	.824(4)	.474(2)	6.5(11)
H(15)	.323(5)	.695(4)	.164(2)	6.5(10)
H(16)	.607(5)	.335(4)	.374(3)	7.8(12)
H(17)	.208(6)	.825(4)	.157(2)	7.2(12)

(a) The isotropic temperature factor has the form $T=\exp(-B(\sin\theta/\lambda)^2)$.

(b) Standard deviations of the least significant digits estimated by least squares are given in parentheses.

Table IV-4

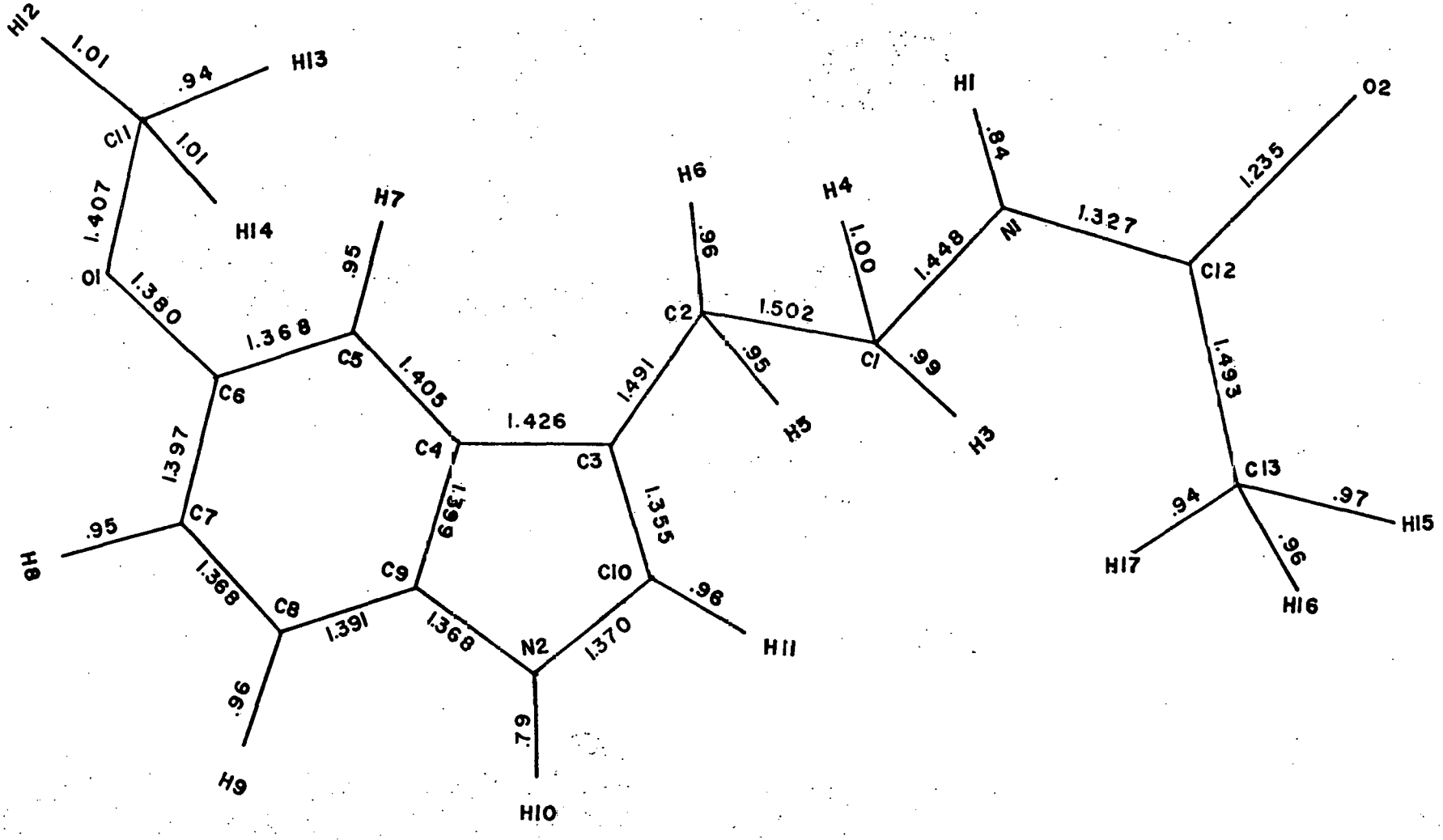
IntramolecularDistances (in Å) and their Standard Deviations (a) in Melatonin.

Atoms	Distance	Atoms	Distance
C(1)-C(2)	1.502(4)	N(1)-H(1)	.84(3)
C(2)-C(3)	1.491(4)	C(1)-H(3)	.99(3)
C(3)-C(4)	1.426(4)	C(1)-H(4)	1.00(3)
C(4)-C(5)	1.405(4)	C(2)-H(5)	.95(3)
C(4)-C(9)	1.399(4)	C(2)-H(6)	.96(3)
C(5)-C(6)	1.368(4)	C(5)-H(7)	.95(3)
C(6)-C(7)	1.397(5)	C(7)-H(8)	.95(3)
C(7)-C(8)	1.368(5)	C(8)-H(9)	.96(3)
C(8)-C(9)	1.391(5)	N(2)-H(10)	.79(3)
C(9)-N(2)	1.368(4)	C(10)-H(11)	.96(3)
N(2)-C(10)	1.370(4)	C(11)-H(12)	1.01(4)
C(10)-C(3)	1.355(4)	C(11)-H(13)	.94(4)
C(6)- O(1)	1.380(4)	C(11)-H(14)	1.01(4)
O(1)-C(11)	1.407(5)	C(13)-H(15)	.97(4)
C(1)-N(1)	1.448(4)	C(13)-H(16)	.96(4)
N(1)-C(12)	1.327(4)	C(13)-H(17)	.94(4)
C(12)-O(2)	1.235(3)		
C(12)-C(13)	1.493(5)		

(a) Standard deviations are estimated by the method of least squares and are indicated with parentheses.

Figure IV-1

Atomic Numbering System and Bond Distances in Melatonin.


H12
1.01
.94
H13
C11
1.01
H14
1.407
O1
1.380
1.368
H7
.95
C5
C6
1.405
1.426
C4
C3
1.491
H6
.96
C2
1.502
.95
H5
H4
1.00
1.448
N1
.84
H1
1.327
O2
1.235
C12
1.493
C1
.99
H3
C13
.94
H17
.97
H15
.96
H16
1.397
C7
.95
H8
1.399
1.355
1.368
C8
C9
C10
.96
H11
1.391
1.368
N2
1.370
.96
H9
.79
H10

XBL 705-949

1.384Å. As in 5-methoxytryptamine the bonds N(2)-C(9) and N(2)-C(10) are equivalent within one standard deviation and are three standard deviations shorter than the average for the ring. The bond C(3)-C(10) is seven standard deviations shorter than the average. All the results are consistent with the structure of 5-methoxytryptamine. As in 5-methoxytryptamine, short bonds in these aromatic rings correlate with large frontier electron density as calculated by Hueckel molecular orbital theory. The shortest bond correlates with the highest density.

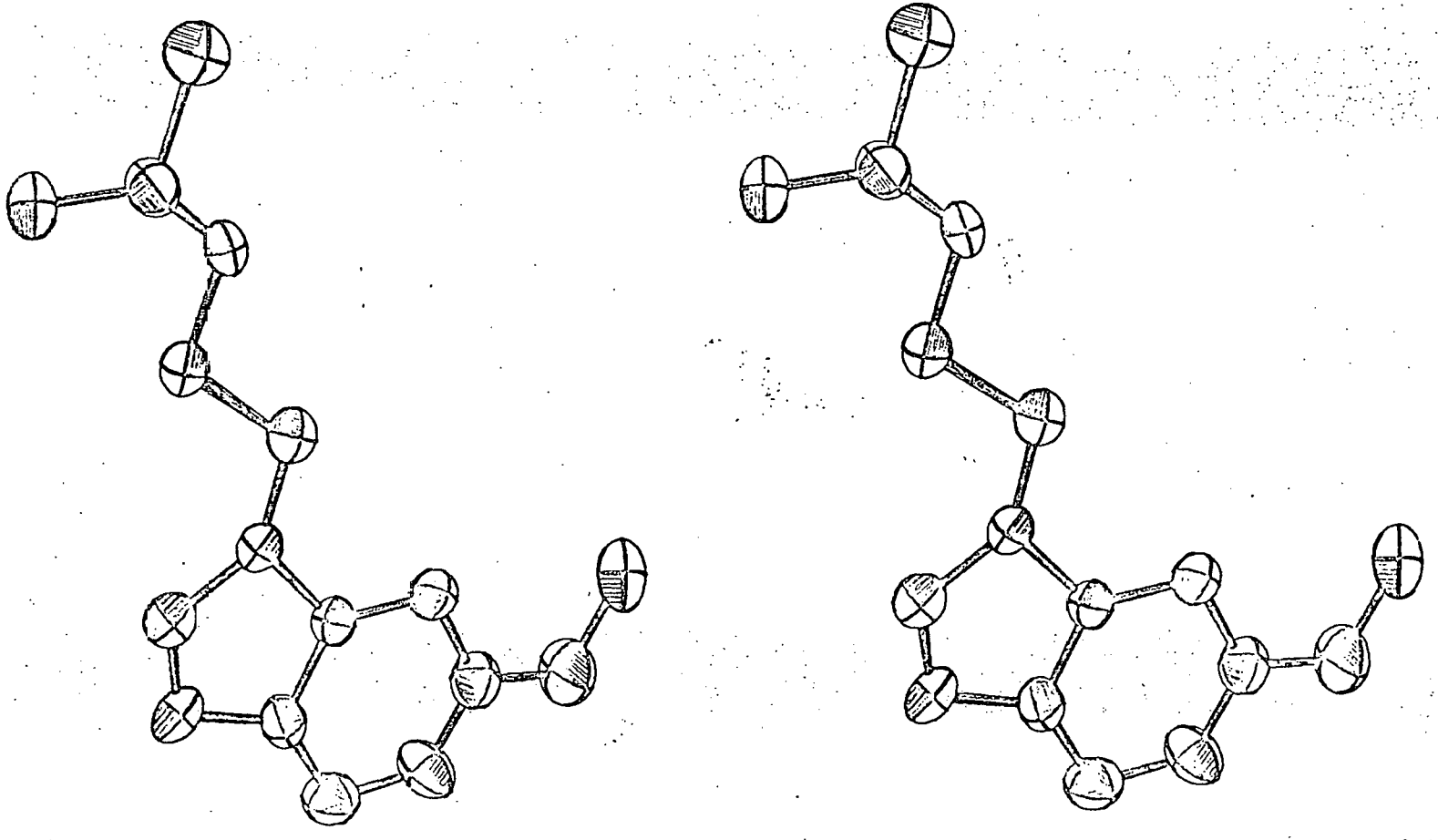
The average of all bonds for the indole ring is 1.385Å. This value is consistent with that found in 5-methoxytryptamine. All the intramolecular angles in the indole ring are consistent within the standard deviations with the results of 5-methoxytryptamine. Intramolecular angles not involving hydrogen are given in Table IV-5, and all intramolecular angles involving hydrogen are given in Table IV-6.

A stereoscopic view of thermal motion in melatonin is shown in Figure IV-2. Thermal parameters are reasonable, and they may be found in Tables IV-2 for non-hydrogen, and IV-3 for hydrogen atoms, as related earlier.

As in 5-methoxytryptamine, the indole ring is planar within .02Å, but is not planar within the standard deviations of the structure determination. Deviations of atoms from least squares planes in melatonin are given in Table IV-7.

Figure IV-2

Anisotropic Thermal Motion in Melatonin.



XBL 705-937

Table IV-5

Intramolecular Angles (in degrees) and their Standard Deviations (a) for all Non-Hydrogen Atoms in Melatonin.

Atoms		Atoms	
C(2)-C(3)-C(4)	125.3(3)	C(8)-C(9)-N(2)	131.2(4)
C(2)-C(3)-C(10)	128.6(3)	C(9)-N(2)-C(10)	109.1(3)
C(3)-C(4)-C(5)	132.7(3)	C(9)-C(4)-C(5)	119.4(3)
C(3)-C(4)-C(9)	107.9(3)	N(2)-C(10)-C(3)	110.1(3)
C(4)-C(5)-C(6)	118.3(3)	C(10)-C(3)-C(4)	106.1(3)
C(4)-C(9)-C(8)	122.0(3)	C(3)-C(2)-C(1)	113.8(3)
C(4)-C(9)-N(2)	106.8(3)	C(2)-C(1)-N(1)	111.0(3)
C(5)-C(6)-C(7)	121.1(3)	C(1)-N(1)-C(12)	122.7(3)
C(5)-C(6)-O(1)	124.4(3)	N(1)-C(12)-O(2)	120.9(2)
C(6)-O(1)-C(11)	117.0(4)	N(1)-C(12)-C(13)	117.1(4)
C(6)-C(7)-C(8)	121.9(3)	C(13)-C(12)-O(2)	122.0(3)
C(7)-C(8)-C(9)	117.1(4)		

(a) Standard deviations are estimated by the method of least squares and are enclosed in parentheses.

Table IV-6

Intramolecular Angles (in degrees) and their Standard Deviations (a) for all Bonds Involving Hydrogen in Melatonin.

Atoms	Angles	Atoms	Angles
C(12)-N(1)-H(1)	118.9(1.7)	H(12)-C(11)-H(13)	111.2(4.6)
C(1)-N(1)-H(1)	118.3(1.7)	H(13)-C(11)-H(14)	116.6(3.9)
N(1)-C(1)-H(3)	108.5(1.7)	H(14)-C(11)-H(12)	99.6(4.1)
N(1)-C(1)-H(4)	105.3(1.6)	C(6)-C(7)-H(8)	118.5(1.8)
C(1)-C(2)-H(5)	107.3(1.8)	C(8)-C(7)-H(8)	119.5(1.8)
C(1)-C(2)-H(6)	108.7(1.6)	C(7)-C(8)-H(9)	120.2(1.9)
C(3)-C(2)-H(5)	112.1(1.8)	C(9)-C(8)-H(9)	122.7(1.9)
C(3)-C(2)-H(6)	109.7(1.6)	C(9)-N(2)-H(10)	125.7(2.1)
C(2)-C(1)-H(3)	111.8(1.7)	C(10)-N(2)-H(10)	124.4(2.1)
C(2)-C(1)-H(4)	112.3(1.6)	N(2)-C(10)-H(11)	119.6(1.8)
H(3)-C(1)-H(4)	107.5(3.3)	C(3)-C(10)-H(11)	130.0(1.7)
H(5)-C(2)-H(6)	104.7(3.4)	C(12)-C(13)-H(15)	114.4(2.1)
C(4)-C(5)-H(7)	121.4(1.6)	C(12)-C(13)-H(16)	109.8(2.2)
C(6)-C(5)-H(7)	120.3(1.6)	C(12)-C(13)-H(17)	109.9(2.3)
O(1)-C(11)-H(12)	114.0(2.1)	H(15)-C(13)-H(16)	108.4(4.0)
O(1)-C(11)-H(13)	103.4(2.0)	H(16)-C(13)-H(17)	111.4(4.5)
O(1)-C(11)-H(14)	112.5(2.0)	H(15)-C(13)-H(17)	102.7(4.0)

(a) Standard deviations are estimated by the method of least squares and are enclosed in parentheses.

Table IV-7

Deviations (in Å) of Atoms from Least Squares Planes in Melatonin.

Atoms given zero weight in the least squares calculations are indicated with parentheses.

	C(3)	C(10)	N(2)	C(9)	C(4)	C(5)	C(6)	C(7)	C(8)	O(1)	C(2)
Indole	.012	.018	-.020	-.007	-.010	-.014	.004	.012	.006	(.026)	(.041)
Benzene	(.038)	(.042)	(-.008)	-.002	.004	-.004	.002	.001	-.001	(.018)	(.078)
Pyrrole	-.001	.008	-.012	.011	-.006	(-.004)	(.032)	(.053)	(.042)	(.062)	( )

	O(1)	N(1)	C(12)	C(13)
Acetyl	-.002	-.002	.005	-.001

Deviation from	C(1)	N(1)	C(12)	C(13)	O(2)	C(11)
Indole plane	(-.064)	(.166)	(.308)	(.618)	(.169)	(.038)

In 5-methoxytryptamine the indole ring was planar except C(7) and C(8) were above the plane by 10 standard deviations. The first members of the side chains C(2) and O(1) were in the plane of the ring. In melatonin the benzene portion of the indole ring is quite planar, and the nitrogen N(2) of the pyrrole ring lies in this plane. Carbons C(3) and C(10) of the pyrrole ring are above the plane of the indole ring by at least 9 standard deviations. The oxygen O(1) and carbon C(2) of the side chains lie significantly above the ring also. It is not clear why C(3) and C(10) should be out of the indole ring plane. This deviation from the plane does not seem to be a consequence of molecular packing.

All the atoms of melatonin lie approximately in the plane of the indole ring. Deviation of all the non-hydrogen atoms of the structure from a least squares plane through the indole ring is given in Table IV-7. As in 5-methoxytryptamine, the aliphatic and the ether side chains unfold on the same side of the ring. The side chain at C(3) drifts casually above the ring to a maximum distance above the ring of .62Å with the terminal atom, C(13).

The angle Φ of melatonin is 175.0°. It will be remembered from Section III that this is the angle between the ethane bond of the side chain and the indole ring such that if $\Phi = 0$ C(1) is pointed toward C(5) and is in the plane of the ring. The minimum energy conformation in the absence of outside

forces is with $\Phi = 90°$, but the energy separation between $\Phi = 90°$ and $\Phi = 180°$ is only 2.3kcal/mole . Within the approximations of Hueckel theory, these two conformations are equivalent.[8]

The ether side chain lies almost entirely in the plane of the indole ring. A least squares plane through C(11)-O(1)-C(6) is parallel within one degree to the plane of the ring. The conformation about the ethane bond C(1)-C(2) is almost perfectly trans and staggered. This corresponds to the minimum conformation energy expected as calculated by extended Hueckel molecular orbital theory. The possible conformations for C(3)-C(2)-C(1)-N(1) are shown in Figure III-6. A rotation angle of 180° between N(1) and C(3) corresponds to zero conformation energy and the trans- staggered configuration. This conformation angle in melatonin is 188.4°. The conformation of the nitrogen-containing side chain in melatonin is very close to that expected from Hueckel molecular orbital theory calculations, and the energy separation between the minimum (0.0kcal/mole), and that found for melatonin is 2.3kcal/mole. We may contrast this value with the very large separation energy found in 5-methoxytryptamine of 12.8 kcal/mole.

In 5-methoxytryptamine, the nitrogen side-chain moved into a higher conformational energy state in order to complete an intermolecular hydrogen bond between N(1) of mole-

cule (x,y,z) and N(2) of molecule (1+x,-y,1/2+z). The ability of N(1) to form an N-H-N hydrogen bond with N(2) is destroyed in melatonin by the acetyl group.

The acetyl group is planar as we can see from Table IV-7. The angles about C(12) and N(1) are all close to 120°. The planar acetyl group is nearly coplanar with the indole ring; the dihedral angle between the two planes is 7.7°. The acetyl group may be partially in resonance with the indole ring. The bonds C(2)-C(3) and C(1)-C(2) are significantly shorter than expected for a carbon-carbon single bond.[9] The tetrahedral angles about C(2) and C(1) are larger than normal by about 20 standard deviations for C(2) and 11 standard deviations for C(1).

Instead of an N-H-N hydrogen bond as found in 5-methoxytryptamine, melatonin forms an N-H-O hydrogen bond. The oxygen O(2) of molecule (x,y,z) is hydrogen bonded to N(2) in the indole ring of molecule (-x,1/2+y,1/2-z). The distance of separation is 2.903Å, and the angle N(2)-H(10)-O(2) is 163.6°. The oxygen O(2) of (x,y,z) is also only 2.966Å from N(1) of the side chain in (-x, 1/2+y,1/2-z), and the angle N(1)-H(1)-O(2) is 161.3°. Thus, both the nitrogen N(1) of the side chain and N(2) of the ring in molecule (x,y,x) are hydrogen-bonded to oxygen O(2) in the glide related molecule (-x,1/2+y, 1/2-z).

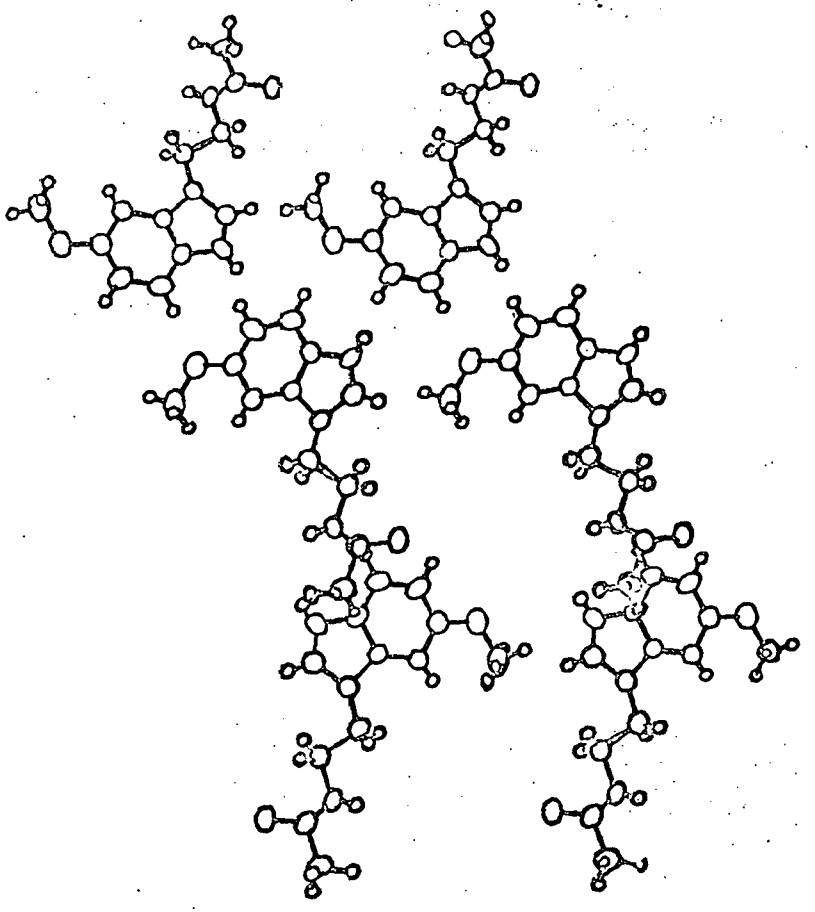
These distances and angles for the N-H-O hydrogen bond are in general agreement with those found in neutron diffraction work.[10] These hydrogen bonds are weaker than the N-H-O bonds of amino acids or proteins. The N-O distance for an N-H-O bond in these compounds is 2.79 ± .12Å. The ideal N-H-O angle is 180°, and for each 6° of deviation approximately .1 kcal/mole of strain is produced in the hydrogen bond.[11] These N-H-O hydrogen bonds in melatonin with distances of about 2.9Å and angles of about 160° are rather weak bonds. Also, the nitrogen-hydrogen distance is sometimes used as a criterion of hydrogen bond strength. If a strong hydrogen bond is formed, the nitrogen-hydrogen distance should be longer, as the hydrogen is donated to the oxygen. The distance N(2)-H(10) is .79Å, and the N(1)-H(1) distance is .84Å. These are the shortest intramolecular bonds involving hydrogen in the whole molecular structure, but least squares estimated standard deviations on these bond lengths are ± .03 Å. Since these estimated standard deviations are the minimum standard deviations for every structure, these short N-H bonds are probably not significant within the accuracy of the structure determination.

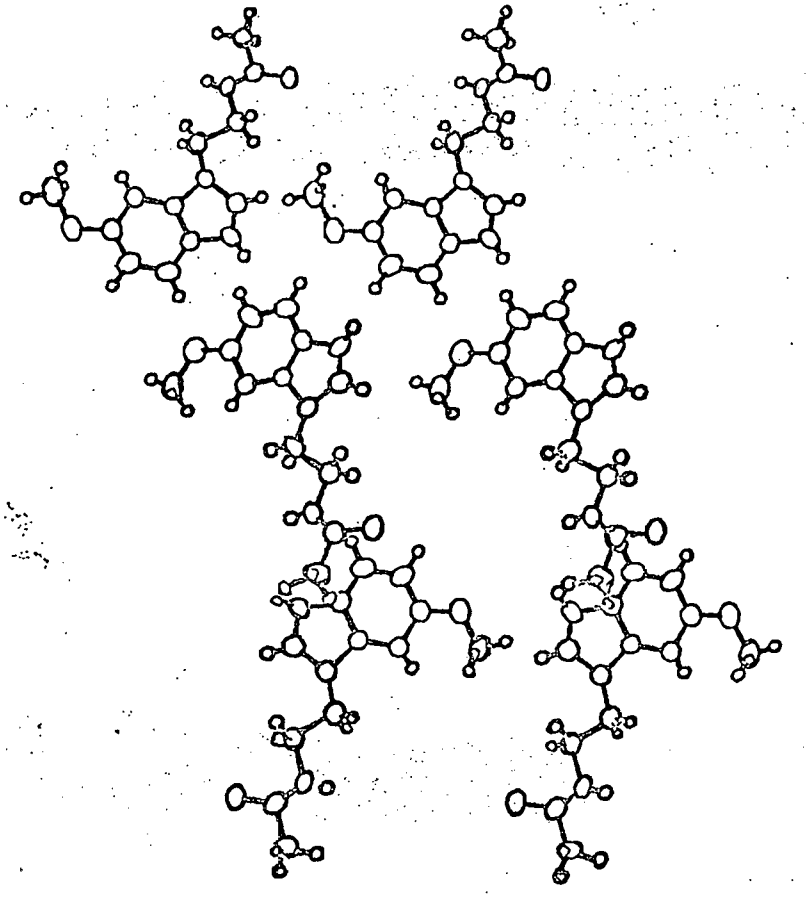
The crystal structure of melatonin is held together by a rather weak network of hydrogen bonds. The indole ring makes an angle of 73.2° with the ac plane, and the side chain in the molecule at (x,y,z) is hydrogen bonded to the

ring and side chain of the glide-related molecule at (-x,y+1/2, 1/2-z). The closest non-bonded approach in the structure not involving hydrogen is 3.758 Å. This interaction is between O(2) of (x,y,z) and C(13) of (-x,y+1/2,1/2-z). A diagram of intermolecular packing in melatonin is shown in Figure IV-3. A list of final observed and calculated structure factors for melatonin is given in Table IV-8.

Figure IV-3

Intermolecular Packing in Melatonin.





XBL 705-950

Table IV-8

Observed and Calculated Structure Factors for Melatonin.

TABLE OF OBSERVED AND CALCULATED STRUCTURE FACTORS FOR PLATONEA

FC(10,0,0) = 446

XBL 705-936

E . Conclusion

The molecular structure of melatonin agrees in many aspects with the crystal structure of 5-methoxytryptamine which was related in Section III. The indole ring in melatonin is not planar, although the benzene ring is. The atoms C(2) and C(10) of the pyrrole ring are warped out of the plane of the indole ring by 1.8° . Short bonds in the indole ring correlate with a large frontier electron density as calculated by Hueckel molecular orbital theory. The shortest bond correlates with the highest pi electron density.

In contrast to 5-methoxytryptamine, melatonin assumes a conformation in the solid state very close to that expected from quantum mechanical calculations. Acetylation of the primary amine group destroys the ability of 5-methoxytryptamine to form a strong N-H-N hydrogen bond. The N-H-O hydrogen bonds formed in melatonin are very weak, and the molecule is able to assume the minimum energy conformation calculated for the isolated molecule with no intermolecular interactions. Partial conjugation of the acetyl group with the indole ring may help stabilize this preffered conformation.

Since 5-methoxytryptamine forms such a strong intermolecular bond in the solid state, it seems plausible that serotonin, which differs only by a methyl group in the 5 position should also form a strong hydrogen bond. Furthermore, it is reasonable to assume that serotonin could be bound in the "granules" by hydrogen bonding.[12] This idea is partially supported by the fact that serotonin forms an addition complex with ATP.[13] It is possible that this addition complex is formed by N-H-N hydrogen bonds of serotonin with adenine.

The dual conformation theory of Gaddum is completely consistent with the x-ray work.[14] Serotonin could assume a higher energy conformation in a biological environment conducive to the formation of hydrogen bonds, and a lower energy conformation similar to that calculated for the isolated molecule in a biological environment where formation of hydrogen bonds would be difficult.

It is hoped that these molecular structures will help those who are trying to piece together the truth of the serotonin metabolism.

Bibliography

1. A.B. Lerner, J.D. Case, and Y. Takahashi, J. Biol. Chem., 235,1992(1960).
2. J. Axelrod, and H. Weissbach, Science, 131, 1312(1960).
3. W.M. McIsaac, and I.H. Page, Science, 128, 537(1958).
4. H.S. Maddox, and M.L. Maddox, unpublished program.
5. D. Sayre, Acta Cryst., 5, 60(1952).
6. H. Hauptman, and J. Karle;"Solution of the Phase Problem I. The Centrosymmetric Crystal" ACA Monograph No. 3. Brooklyn Polycrystal Book Service 1953.
7. Robert Everson Long, Phd. Thesis, UCLA, (1965).
8. Lemont B. Kier, J. Pharm. Sci., 57,(7), 1188(1968).
9. International Tables for X-ray Crystallography, 3,276, Kynoch Press, Birmingham, England,(1962).
10. Walter C. Hamilton, and James A. Ibers, "Hydrogen Bonding in Solids," W.A. Benjamin, New York, (1968).
11. Linus Pauling, "Nature of the Chemical Bond," Cornell University Press, Ithaca, New York,(1960).
12. E.A. Carlini, and J.P. Green, Brit. J. Pharmacol., 20, 264(1963).
13. G.V.R. Born, G.I.C. Ingram, and R.S. Stacey, Brit. J. Pharmacol.,13,62(1958).
14. J. H. Gaddum, and K.A. Hameed, Brit. J. Pharmacol.,9 , 240(1954).

Acknowledgements

I wish to thank Professor David Templeton for his suggestions, support, and guidance. I thank Dr. Allan Zalkin for the use of his Fortran computer programs, and for essential help with the PDP-8I computer. I thank Fred Hollander for his assistance and consideration. I wish to thank Siddhartha Kay, Mark Fisher, and Mrs. Helena Ruben for valuable discussions and suggestions.

I thank Professor Quay of the physiology department at UC Berkeley for crystals of 5-methoxytryptamine, and melatonin. I wish to thank the photographic section of the graphic arts department here at LRL-Berkeley for their excellent service and fine workmanship.

I wish to thank Cecilia Smilingtrout for the real sharp typing job she did on the manuscript.

I also wish to thank the Lawrence Radiation Laboratory, Berkeley, and the Atomic Energy Commission for their support of this work.

Appendix A

1. Derivation of the Orientation Matrix $\underline{a}_{ij}$

The orientation matrix $\underline{a}_{ij}$ for the trial and error computer program is a function only of the cell dimensions of the crystal studied and the direction cosines of the normal to the molecular plane relative to the unitary monoclinic base vectors, $\underline{u}_i$, mentioned in Section II. The derivation of the components of this matrix relies heavily upon use of the dot and cross vector products. Since in many cases it is possible to deduce independently the orientation of the molecular plane in terms of the Miller indices of a Bragg plane which the molecule is nearly coincident with, the first step in the derivation of the orientation matrix is the development of an expression for this Bragg normal in terms of its direction cosines.

Knowledge of the principles of the reciprocal lattice will be assumed to permit brevity of development. Several good books which elucidate the concepts of reciprocal space are available, and a few of them are listed in the bibliography.[1,2,3]

The reciprocal lattice vector $\underline{g}(h_i) = \Sigma h_i \underline{s}_i$ $(i = 1,3)$ where h_i represents the ith Miller index and $\underline{s}_i$ represents the ith reciprocal space base vector, may be dotted with the direct space monoclinic base vector set $\underline{b}_i$ since the direct

space and the reciprocal space vector sets have the same origin. The result of this is (1)

$$\cos \varepsilon_i = h_i/|\underline{\sigma}(h_i)|\ |\underline{b}_i| \qquad (i = 1,3) \qquad (1)$$

The direction cosines, $\cos \varepsilon_i$, of a Bragg plane normal relative to the direct space vector basis set $\underline{b}_i$ is a function only of h_i, the ith Miller index, and the length of $\underline{b}_i$ and $\underline{\sigma}(h_i)$. The length of $\underline{\sigma}(h_i)$ is a well-known property of reciprocal space, and $|\underline{\sigma}(h_i)| = 1/d(h_i)$ where $d(h_i)$ is the perpendicular distance between the Bragg planes involved.

It is convenient to normalize the $\underline{b}_i$ such that $\underline{\nu}_i = \underline{b}_i/|\underline{b}_i|$. These $\underline{\nu}_i$ are the unitary monoclinic base vectors mentioned in Section II. The vector $\underline{\sigma}(h_i)$ may also be expressed in this vector basis, and $\underline{\sigma}(h_i) = \Sigma n_i \underline{\nu}_i$ $(i = 1,3)$. The components, n_i, of $\underline{\sigma}(h_i)$ may be expressed in terms of direction cosines. This result follows if we dot $\underline{\sigma}(h_i)$ with $\underline{\nu}_i$ and solve the resulting set of linear equations remembering that the $\underline{\nu}_i$ are normalized direct space monoclinic base vectors. The n_i are given in (2).

$$\begin{aligned} n_1 &= [(\cos\varepsilon_1 - \cos\varepsilon_3 \cos\beta)/\sin^2\beta]|\underline{\sigma}(h_i)| \\ n_2 &= [\cos\varepsilon_2]|\underline{\sigma}(h_i)| \qquad (2) \\ n_3 &= [(\cos\varepsilon_3 - \cos\varepsilon_1 \cos\beta)/\sin^2\beta]|\underline{\sigma}(h_i)| \end{aligned}$$

The expression for the n_i may be normalized with division by $|\underline{\sigma}(h_i)|$. The normal $\underline{\sigma}(h_i)$ to the Bragg plane h_i has now been normalized and expressed as a function of its direction cosines and the cell constants of the crystal. This normal is one vector of an orthonormal vector set into which the geometry of a planar molecule can be graphed. Another vector $\underline{e}_1 = \Sigma\lambda_i\underline{\nu}_i$ $(i = 1,3)$ can be defined. This vector can be dotted with $\underline{\sigma}(h_i)$ and the result set equal to zero. This will give us after $\underline{e}_1$ is normalized the expression (3)

$$\lambda_1 = 1/|\underline{e}_1|$$

$$\lambda_2 = 1/|\underline{e}_1| \qquad (3)$$

$$\lambda_3 = -[(\cos\varepsilon_1 + \cos\varepsilon_2)/\cos\varepsilon_3]/|\underline{e}_1|$$

$$|\underline{e}_1| = \{2+[(\cos\varepsilon_1+\cos\varepsilon_2)/\cos\varepsilon_3]^2-2(\cos\varepsilon_1+\cos\varepsilon_2)\cos\beta/\cos\varepsilon_3\}^{1/2}$$

If $\cos\varepsilon_3 = 0$, the expression for λ_i becomes infinite. This difficulty is overcome by deriving analogous expressions for λ_i having different $\cos\varepsilon_i$ in the denominators of the terms involved and branching the computer program to avoid division by zero.

The $\lambda_i = a_{i1}$ $(i = 1,3)$ of the orientation matrix. We can define a third vector $\underline{e}_2 = \Sigma\mu_i\underline{\nu}_i$ $(i = 1,3)$ such that $\underline{\sigma}(h_i) \times \underline{e}_1 = \underline{e}_2$. Solution of this vector equation, expression of the answer in terms of the $\underline{\nu}_i$, and normalization of

$\underline{e}_2$ gives (4).

$$\mu_1 = \xi - \eta \cos\beta/\sin\beta$$

$$\mu_2 = \Omega \sin\beta \qquad (4)$$

$$\mu_3 = \eta - \xi \cos\beta/\sin\beta$$

where

$$\xi = (n_2\lambda_3 - n_3\lambda_2)$$

$$\Omega = (n_3\lambda_1 - n_1\lambda_3)$$

$$\eta = (n_1\lambda_2 - n_2\lambda_1)$$

and the required matrix $\underline{a}_{ij}$ is represented in (5).

$$\underline{a}_{ij} = \begin{vmatrix} \lambda_1 & \mu_1 \\ \lambda_2 & \mu_2 \\ \lambda_3 & \mu_3 \end{vmatrix} \qquad (5)$$

Appendix A

2. Fortran Listings

```
      PROGRAM NORMA(INPUT,OUTPUT,TAPE2=INPUT,TAPE3=OUTPUT,TAPE4)
      DIMENSION HH(50),HK(50),HL(50),RDHKL(50),DHKL(50)
      DIMENSION COSA(50),COSB(50),COSC(50),A(50),B(50),C(50)
      DIMENSION BAT(50),ANGA(50),ANGB(50),ANGC(50)
      DIMENSION A1(50),A2(50),A3(50),SUM(50),BVECT(50)
      DIMENSION BSQ(50),B1(50),B2(50),B3(50),R(50),S(50),T(50),C1(50)
      DIMENSION C2(50),C3(50)
      IN=2
      LOUT=3
      NPUN=4
      READ(IN,300) AA,BB,CC,ALPHA,BETA,GAMMA,NHKL
      WRITE(LOUT,300) AA,BB,CC,ALPHA,BETA,GAMMA,NHKL
      READ(IN,310) (HH(I),HK(I),HL(I),I=1,NHKL)
      WRITE(LOUT,310) (HH(I),HK(I),HL(I),I=1,NHKL)
      PI=3.14159265
      TWOPI=2.*PI
      RADCON=PI/180.
      CONRAD=180./PI
      ASTAR=1.0/(AA*SIN(RADCON*BETA))
      BSTAR=1.0/BB
      CSTAR=1.0/(CC*SIN(RADCON*BETA))
      BETAS=180.-BETA
      DO 10 I=1,NHKL
      BAT(I)=(HH(I)*ASTAR)**2+(HK(I)*BSTAR)**2+(HL(I)*CSTAR)**2
     1+2.0*HH(I)*HL(I)*ASTAR*CSTAR*COS(BETAS*RADCON)
      RDHKL(I)=SQRT(BAT(I))
      DHKL(I)=1.0/RDHKL(I)
      COSA(I)=(HH(I)*DHKL(I))/AA
      COSB(I)=(HK(I)*DHKL(I))/BB
      COSC(I)=(HL(I)*DHKL(I))/CC
      A(I)=ACOS(COSA(I))
      B(I)=ACOS(COSB(I))
      C(I)=ACOS(COSC(I))
      ANGA(I)=A(I)*CONRAD
      ANGB(I)=B(I)*CONRAD
      ANGC(I)=C(I)*CONRAD
   10 CONTINUE
      WRITE(LOUT,400)
      WRITE(LOUT,410) (HH(I),HK(I),HL(I),DHKL(I),COSA(I),COSB(I),
     1COSC(I),ANGA(I),ANGB(I),ANGC(I),I=1,NHKL)
      DO 20 J=1,NHKL
      A1(J)=(COSA(J)-COSC(J)*COS(BETA*RADCON))/(SIN(BETA*RADCON))**2.
      A2(J)=COSB(J)
      A3(J)=(COSC(J)-COSA(J)*COS(BETA*RADCON))/(SIN(BETA*RADCON))**2.
      IF(COSC(J).EQ.0.) 12,15
   12 SUM(J)=COSB(J)+COSC(J)
      BSQ(J)=(SUM(J)/COSA(J))**2+1.+1.-2.*(SUM(J)/COSA(J))*
     1 COS(BETA*RADCON)
      BVECT(J)=SQRT(BSQ(J))
      WRITE(LOUT,395) BVECT(J)
      B1(J)=-(SUM(J)/COSA(J))/BVECT(J)
      B2(J)=1./BVECT(J)
      B3(J)=1./BVECT(J)
      GO TO 16
   15 SUM(J)=COSA(J)+COSB(J)
      BSQ(J)=1.+1.+(SUM(J)/COSC(J))**2-2.*(SUM(J)/COSC(J))*
     1 COS(BETA*RADCON)
      BVECT(J)=SQRT(BSQ(J))
      WRITE(LOUT,395) BVECT(J)
      B1(J)=1.0/BVECT(J)
      B2(J)=1.0/BVECT(J)
```

```
      B3(J)=-(SUM(J)/COSC(J))/BVECT(J)
   16 CONTINUE
      R(J)=A2(J)*B3(J)-A3(J)*B2(J)
      S(J)=A3(J)*B1(J)-A1(J)*B3(J)
      T(J)=A1(J)*B2(J)-A2(J)*B1(J)
      C1(J)=(R(J)-T(J)*COS(BETA*RADCON))/(SIN(BETA*RADCON))
      C2(J)=S(J)*SIN(BETA*RADCON)
      C3(J)=(T(J)-R(J)*COS(BETA*RADCON))/(SIN(BETA*RADCON))
   20 CONTINUE
      WRITE(LOUT,320)
      WRITE(LOUT,330) (A1(J),A2(J),A3(J),J=1,NHKL)
      WRITE(LOUT,340)
      WRITE(LOUT,350) (B1(J),B2(J),B3(J),C1(J),C2(J),C3(J),HH(J),HK(J),
     1HL(J),J=1,NHKL)
      WRITE(NPUN,500) (B1(J),B2(J),B3(J),C1(J),C2(J),C3(J),HH(J),HK(J),
     1HL(J),J=1,NHKL)
      END FILE NPUN
      REWIND NPUN
      WRITE(LOUT,475) NHKL
  300 FORMAT(6F10.5,I5)
  310 FORMAT(3F5.0)
  320 FORMAT(41HTHE COMPONENTS OF THE NORMAL FOLLOW BELOW)
  330 FORMAT(3F10.4)
  340 FORMAT(5X,2HB1,10X,2HB2,10X,2HB3,10X,2HC1,10X,2HC2,10X,2HC3)
  350 FORMAT(6F10.4,1X,3F3.0)
  395 FORMAT(1F10.3)
  400 FORMAT(4X,1HH,3X,1HK,3X,1HL,3X,4HDHKL,5X,4HCOSA,5X,4HCOSB,
     1 5X,4HCOSC,5X,4HANGA,5X,4HANGB,5X,4HANGC)
  410 FORMAT(3F5.0,4F10.4,3F10.1)
  475 FORMAT(10HA TOTAL OF,I5,23HCARDS HAVE BEEN PUNCHED)
  500 FORMAT(6F10.5,1X,3F3.0)
      STOP
      END
```

```
      PROGRAM OMOO(INPUT,OUTPUT,TAPE2=INPUT,TAPE3=OUTPUT)
C     IN ITS PRESENT FORM THIS PROGRAM WORKS FOR MONOCLINIC OR HIGHER
C     SYMMETRY.  THE PROGRAMMER MUST WRITE THE SIN AND COSINE FORMULAS
C     OF HIS SPACE GROUP FROM THE INTERNATIONAL TABLES AND INSERT THEM
C     IN SUBROUTINE FANG.  IF MORE THAN 200 DATA OR 30 ATOMS ARE READ IN
C     THE PROGRAMMER MUST EXPAND THE DIMENSION STATEMENTS OF THOSE
C     VARIABLES.
C     CARD 1(6F10.5,2I5) A,B,C,ALPHA,BETA,GAMMA,NDATA,NATOM
C     AXIS LENGTHS IN ANGSTROMS AND ANGLES IN DEGREES.  NDATA IS  THE
C     NUMBER OF DATA READ IN.  NATOM IS THE NUMBER OF ATOMS YOU READ IN
C     CARD 2. (5F10.3,I10) XOR,YOR,ZOR,ANGMAX,DELANG,NCOS
C     XOR,YOR,ZOR ARE CRYSTAL COORDINATES OF CHOICE OF ORIGIN FOR AN
C     ATOM WITHIN YOUR MOLECULE.  GRAPH THE MOLECULE TO SCALE ON A
C     CARTESIAN AXIS SYSTEM.  PLACE ONE OF YOUR ATOMS AT THE ORIGIN OF
C     THE CARTESIAN SET. MAKE ANOTHER ATOM COINCIDE WITH  SOME POINT
C     ON ONE OF THESE CARTESIAN AXES.   THESE TWO ATOMS DESCRIBE
C     A VECTOR WHICH THE PROGRAM WILL USE TO ROTATE THE MOLECULE ABOUT
C     THE PLANE NORMAL.  THE MOLECULE WILL ROTATE ABOUT THE PLANE NORMAL
C     FROM THE ANGLE WHICH YOU DEFINE WITH YOUR GRAPH AS ZERO TO A
C     MAXIMUM ANGLE ANGMAX IN STEPS OF DELANG DEGREES.
C     NCOS IS THE EXACT NUMBER OF CARDS 4.
C     CARD 2A. (3F10.3,3I10) DELX,DELY,DELZ,NX,NY,NZ
C     THE MOLECULAR ORIGIN IS MOVED THROUGH THE CELL IN STEPS OF DELX,
C     DELY,DELZ, FOR NX,NY,NZ TIMES
C     CARDS 3. TAKE YOUR SCATTERING FACTOR TABLES AND AVERAGE THEM
C     TYPE THEM IN FORMAT(7(F4.2,F6.3))
C      CARDS 4 CONTAIN INFORMATION THE PROGRAM NEEDS TO PRODUCE CRYSTAL
C     COORDINATES FOR YOUR ATOMS FROM YOUR GRAPHED GEOMETRY.   THIS
C     INFORMATION IS PUNCHED OUT BY PROGRAM NORMA IN THE PROPER FORMAT
C     AND MAY BE USED DIRECTLY. THERE ARE NCOS OF THESE CARDS.
C     CARDS 5. THERE ARE NDATA OF THESE. HAVE YOUR DATA TAPE PUNCHED
C     FROM THIS SELECT THE CARDS YOU NEED
C     CARDS 6. NATOM OF THEM.  R,THETA,FORMAT(2F10.4) R IS THE DISTANCE
C     IN ANGSTROMS FROM YOUR ORIGIN ATOM TO ANY ATOM WITHIN YOUR MOLE-
C     CULE.  YOU GET THIS FROM YOUR GRAPH.  THETA IS THE ANGLE THAT
C     THE NTH ATOM MAKES WITH THE ORIENTATION VECTOR WHICH YOU GRAPHED
C     COINCIDENT WITH ONE OF THE CARTESIAN AXES EARLIER.
      COMMON/CELL/AA,BB,CC,ALPHA,BETA,GAMMA,NDATA,NATOM
      COMMON/SCATER/SLAM(21),FTBL(21)
      COMMON/HKL/HH(200),HK(200),HL(200),FOBS(200)
      DIMENSION XEX(30),YEY(30),R(30),THETA(30),OMEGA(36)
      DIMENSION XU(30),YU(30),ZU(30),YO(30)
      DIMENSION B1(50),B2(50),B3(50),C1(50),C2(50),C3(50)
      DIMENSION NNN(50),MMM(50),LLL(50),XO(20),ZO(20)
      IN=2
      LOUT=3
      READ(IN,300) AA,BB,CC,ALPHA,BETA,GAMMA,NDATA,NATOM
      READ(IN,330) XOR,YOR,ZOR,ANGMAX,DELANG,NCOS
      READ(IN,295)    DX,DY,DZ,NX,NY,NZ
      READ(IN,320) (SLAM(K),FTBL(K),K=1,21)
      READ(IN,310) (B1(L),B2(L),B3(L),C1(L),C2(L),C3(L),NNN(L),MMM(L),
     1LLL(L),L=1,NCOS)
      READ(IN,340) (HH(I),HK(I),HL(I),FOBS(I),I=1,NDATA)
      READ(IN,350) (R(N),THETA(N),N=1,NATOM)
      WRITE(LOUT,300) AA,BB,CC,ALPHA,BETA,GAMMA,NDATA,NATOM
      WRITE(LOUT,330) XOR,YOR,ZOR,ANGMAX,DELANG,NCOS
      WRITE(LOUT,295) DX,DY,DZ,NX,NY,NZ
      WRITE(LOUT,320) (SLAM(K),FTBL(K),K=1,21)
      WRITE(LOUT,310) (B1(L),B2(L),B3(L),C1(L),C2(L),C3(L),NNN(L),
```

```
     1MMM(L),LLL(L),L=1,NCOS)
      WRITE(LOUT,340) (HH(I),HK(I),HL(I),FOBS(I),I=1,NDATA)
      WRITE(LOUT,350) (R(N),THETA(N),N=1,NATOM)
      PI=3.14159265
      TWOPI=2.0*PI
      RADCON=PI/180.
      NANG=ANGMAX/DELANG
      XO(1)=XOR
      YO(1)=YOR
      ZO(1)=ZOR
      OMEGA(1)=0.0
      WRITE(LOUT,430) NNN(1),MMM(1),LLL(1)
      WRITE(LOUT,440)     XO(1)
      WRITE(LOUT,450)     YO(1)
      WRITE(LOUT,460)     ZO(1)
      DO 40 L=1,NCOS
      DO 35 N=1,NZ
      DO 30 I=1,NY
      DO 25 M=1,NX
      DO 20 J=1,NANG
      DO 10 K=1,NATOM
      XEX(K)=R(K)*COS((THETA(K)+OMEGA(J))*RADCON)
      YEY(K)=R(K)*SIN((THETA(K)+OMEGA(J))*RADCON)
      XU(K)=(XEX(K)*B1(L)+YEY(K)*C1(L))+(XO(M)*AA)
      YU(K)=(XEX(K)*B2(L)+YEY(K)*C2(L))+(YO(I)*BB)
      ZU(K)=(XEX(K)*B3(L)+YEY(K)*C3(L))+(ZO(N)*CC)
C     XU,YU,ZU ARE COORDINATES OF EACH ATOM REFFERRED TO MONOCLINIC
C     AXES WHICH ARE ONE ANGSTROM LONG
   10 CONTINUE
      WRITE(LOUT,410) OMEGA(J)
      CALL FANG(XU,YU,ZU,11)
      OMEGA(J+1)=OMEGA(J)+DELANG
   20 CONTINUE
      XO(M+1)=XO(M)+DX
      WRITE(LOUT,440)    XO(M+1)
   25 CONTINUE
      YO(I+1)=YO(I)+DY
      WRITE(LOUT,450)   YO(I+1)
   30 CONTINUE
      ZO(N+1)=ZO(N)+DZ
      WRITE(LOUT,460)   ZO(N+1)
   35 CONTINUE
      WRITE(LOUT,430)   NNN(L+1),MMM(L+1),LLL(L+1)
   40 CONTINUE
  295 FORMAT(3F10.3,3I10)
  300 FORMAT(6F10.5,2I5)
  310 FORMAT(6F10.5,1X,3I3)
  320 FORMAT(7(F4.2,F6.3))
  330 FORMAT(5F10.3,I10)
  340 FORMAT(3F5.0,1F10.2)
  350 FORMAT(2F10.4)
  410 FORMAT(8HOMEGA IS,1F10.1)
  430 FORMAT(19HTHIS IS ORIENTATION,3I5)
  440 FORMAT(4HX IS,1F5.2)
  450 FORMAT(4HY IS,1F5.2)
  460 FORMAT(4HZ IS,1F5.2)
  500 STOP
      END
      SUBROUTINE FANG(XU,YU,ZU,J)
      DIMENSION XU(J),YU(J),ZU(J)
```

```
      DIMENSION X(30),Y(30),Z(30)
      COMMON/CELL/AA,BB,CC,ALPHA,BETA,GAMMA,NDATA,NATOM
      COMMON/SCATER/SLAM(21),FTBL(21)
      COMMON/HKL/HH(200),HK(200),HL(200),FOBS(200)
      DIMENSION SUM1(30),SUM2(30),A(30),B(30)
      DIMENSION FAT(200),SINLSQ(200),BAT(200)
      DIMENSION SUMFO(201),SUMFC(201),SFOBS(201),DELF(201),SUMDELF(201)
      DIMENSION TEMP(200),SINL(200),SCAT(200),FCALC(200)
      DIMENSION SUMSFOB(201),ASUM(31),BSUM(31)
      PI=3.14159265
      TWOPI=2.*PI
      RADCON=PI/180.
      ASTAR=1.0/(AA*SIN(RADCON*BETA))
      BSTAR=1.0/BB
      CSTAR=1.0/(CC*SIN(RADCON*BETA))
      BETAS=(180.0-BETA)*RADCON
C     THIS LOOP CALCULATES ISOTROPIC TEMP FACTOR
      DO 5 I=1,NDATA
      BAT(I)=(HH(I)*ASTAR)**2+(HK(I)*BSTAR)**2+(HL(I)*CSTAR)**2
     1+2.0*HH(I)*HL(I)*ASTAR*CSTAR*COS(BETAS)
      SINLSQ(I)=BAT(I)/4.0
      TEMP(I)=EXP(-3.0*SINLSQ(I))
      SINL(I)=SQRT(SINLSQ(I))
    5 CONTINUE
      I=1
    6 CONTINUE
      DO 10 K=1,21
      KK=K
    3 IF(SLAM(K+1).GT.SINL(I)) GO TO 12
   10 CONTINUE
   12 SCAT(I)=FTBL(KK)+(FTBL(KK+1)-FTBL(KK))*((SINL(I)-SLAM(KK))/
     1(SLAM(KK+1)-SLAM(KK)))
      I=I+1
      IF(I-NDATA) 6,6,14
C     SCAT IS SCATTERING FACTOR AND TEMP IS ISOTROPIC TEMP FACTOR
   14 CONTINUE
C     TEMP IS AN AVERAGE ISOTROPIC TEMPERATURE FACTOR FOR ALL ATOMS
      DO 20 I=1,NDATA
      DO 15 J=1,NATOM
      X(J)=XU(J)/AA
      Y(J)=YU(J)/BB
      Z(J)=ZU(J)/CC
      SUM1(J)=HH(I)*X(J)+HL(I)*Z(J)+(HK(I)+HL(I))/4.
      SUM2(J)=HK(I)*Y(J)-(HK(I)+HL(I))/4.
      A(J)= 4.*COS(TWOPI*SUM1(J))*COS(TWOPI*SUM2(J))
      B(J)=0.0
      ASUM(1)=0.0
      BSUM(1)=0.0
      ASUM(J+1)=A(J)+ASUM(J)
      BSUM(J+1)=B(J)+BSUM(J)
   15 CONTINUE
      SUMA=ASUM(J+1)
      SUMB=BSUM(J+1)
C     THIS IS THE UNITARY PART OF THE STRUCTURE FACTOR FOR NATOM
      FAT(I) = SQRT(SUMA**2 + SUMB**2)
C     SCAT IS THE AVERAGE SCATTERING FACTOR FOR NATOM
C     TEMP IS THE AVERAGE ISOTROPIC TEMPERATURE FACTOR FOR NATOM
      FCALC(I)= FAT(I)*SCAT(I)*TEMP(I)
   20 CONTINUE
```

```
      DO 25 I=1,NDATA
      SUMFO(1)=0.0
      SUMFC(1)=0.0
      SUMFO(I+1)=FOBS(I)+SUMFO(I)
      SUMFC(I+1)=FCALC(I)+SUMFC(I)
      IF(I-NDATA) 25,30,30
   25 CONTINUE
   30 CONST=SUMFC(I+1)/SUMFO(I+1)
      WRITE(3,350) CONST
      DO 40 L=1,NDATA
      SFOBS(L)=CONST*FOBS(L)
      DELF(L)=ABS(SFOBS(L)-FCALC(L))
      SUMDELF(1)=0.0
      SUMSFOB(1)=0.0
      SUMDELF(L+1)=DELF(L)+SUMDELF(L)
      SUMSFOB(L+1)=SFOBS(L)+SUMSFOB(L)
      IF(L-NDATA) 40,50,50
   40 CONTINUE
   50 R1=SUMDELF(L+1)/SUMSFOB(L+1)
      WRITE(3,370) R1
      WRITE(3,400)  (X(J),Y(J),Z(J),J=1,NATOM)
  350 FORMAT(28HTHE SCALE FACTOR FOR FOBS IS,1F10.5)
  370 FORMAT(15HTHE R FACTOR IS,1F10.5)
  400 FORMAT(3F10.2)
      RETURN
      END
```

Bibliography to the Appendix

1. Martin Buerger, "X-ray Crystallography," John Wiley, New York, 1942.
2. W. H. Zachariasen, "Theory of X-ray Diffraction in Crystals," Dover Publications, New York, second edition, 1967.
3. A. L. Patterson, Nature, 125, 238 (1930).

TECHNICAL INFORMATION DIVISION
LAWRENCE RADIATION LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720