

PART I: THE CRYSTAL AND MOLECULAR STRUCTURE  
OF TETRA-N-PROPYLPORPHINE

PART II: N-FORMYL-L-PHENYLALANINE:  
 $\alpha$ -CHYMOTRYPSIN COMPETITIVE INHIBITOR BINDING

Thesis for the Degree of Ph.D.  
MICHIGAN STATE UNIVERSITY  
PENELOPE WIXSON CODDING  
1971

This is to certify that the

thesis entitled

Part I. The Crystal and Molecular Structure of  
Tetra-n-propylporphine

Part II. N-Formyl-L-Phenylalanine:  $\alpha$ -Chymotrypsin  
Competitive Inhibitor Binding

presented by

Penelope Wixson Coddling

has been accepted towards fulfillment  
of the requirements for

Ph.D. degree in Chemistry

Q. Tulinsky

Major professor

Date September 3, 1971

ABSTRACT

PART I: THE CRYSTAL AND MOLECULAR STRUCTURE OF  
TETRA-N-PROPYLPORPHINE

PART II: N-FORMYL-L-PHENYLALANINE: $\alpha$ -CHYMOTRYPSIN  
COMPLEX, THE EFFECT OF BINDING OF A  
COMPETITIVE INHIBITOR SUBSTRATE-LIKE  
MOLECULE

By

Penelope Wixson Coddling

The structure of  $\alpha, \beta, \gamma, \delta$ -tetra-n-propylporphine was determined by three dimensional X-ray crystallographic techniques. The molecule crystallizes in space group  $P2_1/c$  with two molecules in the unit cell and cell dimensions of  $a = 5.078$ ,  $b = 11.59$ ,  $c = 22.39$  A, and  $\beta = 99.50^\circ$ . The three dimensional structure to 2.8 A resolution of a N-formyl-L-phenylalanine: $\alpha$ -chymotrypsin complex was studied. The complex is isomorphous with  $\alpha$ -chymotrypsin and crystallizes in space group  $P2_1$  with four molecules per unit cell and cell dimensions of  $a = 49.27$ ,  $b = 67.10$ ,  $c = 65.84$  A, and  $\beta = 101.75^\circ$ .

The structure of tetrapropylporphine was solved using Sayre's equation. The molecule is centrosymmetric with two independent pyrrole rings. A common structure for the free base macrocycle was obtained from this determination and the structures of porphine and tetraphenylporphine. The free base structure consists of opposite pyrrole rings possessing imino hydrogen atoms and pyrrole rings differing in like pairs. Small differences which are localized at the bridge carbon

[illegible]



atoms were found for the substituted compounds. In addition, the aliphatic substituent seems to decrease some of the differences between the two pyrrole rings. The nucleus of the tetrapropylporphine ring is essentially planar with some small deviations in planarity due to the close intermolecular distances between the planes stacked along a.

The 2.8 Å resolution difference electron density of the inhibitor:enzyme complex was calculated with the phases of  $\alpha$ -chymotrypsin. The N-formyl-L-phenylalanine molecule binds in the active site region by displacing a localized water molecule which was hydrogen bonded in the active site region. The inhibitor molecule has close contacts with the main chain peptide and side chain of methionine 192 and with the sequence tryptophan 215-serine 214-valine 213.

The substitution of N-formyl-L-phenylalanine produces several changes in the enzyme structure. The most important change is a movement of histidine 40 by about 0.5 Å to a position which is more favorable for hydrogen bonding to glycine 43. The two-fold equivalent residue, histidine' 40 is hydrogen bonded to glycine' 43 in the native enzyme structure. This change may be important for maintaining the conformation of the active site region. Another large change is observed in the region of favored uranyl binding. The change results from a movement of 0.4 Å of the side chain of arginine' 154 of one enzyme molecule in the asymmetric unit toward the side chain of glutamic acid 21 of the other molecule.

PART I: THE CRYSTAL AND MOLECULAR STRUCTURE OF  
TETRA-N-PROPYLPORPHINE

PART II: N-FORMYL-L-PHENYLALANINE: $\alpha$ -CHYMOTRYPSIN  
COMPLEX, THE EFFECT OF BINDING OF A  
COMPETITIVE INHIBITOR SUBSTRATE-LIKE  
MOLECULE

By

Penelope Wixson Coddling

A THESIS

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1971

To my parents, Mr. and Mrs. Robert M. Wixson, for their  
constant encouragement throughout my education and to  
Ed for his unlimited patience and understanding.

It  
patience

a teach

Th  
Discuss

due to  
tachymo

Th  
her thr

Su  
foundat

Th  
England

for he

### Acknowledgments

The author wishes to thank Dr. Alexander Tulinsky for his patience and assistance during this study and for his excellence as a teacher.

The author is also indebted to Mr. Richard L. Vandlen for helpful discussions and many computer programs. In addition, thanks are due to all those who contributed to the structure determination of  $\alpha$ -chymotrypsin which made the second half of this thesis possible.

The author is grateful for an NDEA Fellowship which supported her throughout this study.

Support from the Molecular Biology Section of the National Science Foundation is also gratefully acknowledged.

The author would like to thank Dr. Alan D. Adler of the New England Institute for supplying a sample of tetrapropylporphine and for helpful discussions.

## TABLE OF CONTENTS

### PART I: THE CRYSTAL AND MOLECULAR STRUCTURE OF TETRAPROPYLPORPHINE

I. Introduction . . . . .	1
1. General . . . . .	1
2. Influence of Porphyrin Side Chains . . . . .	3
3. Hydrogen Bonding . . . . .	4
4. The Central Hydrogen Atoms . . . . .	7
II. Experimental . . . . .	10
1. Crystals . . . . .	10
2. Space Group Determination . . . . .	10
3. Intensity Data Collection . . . . .	12
III. Solution of the Phase Problem . . . . .	21
1. The Phase Problem . . . . .	21
2. The Patterson Function . . . . .	21
3. Direct Methods . . . . .	23
IV. Structure Analysis . . . . .	30
1. The Method of Least Squares . . . . .	30
2. The Weighting Scheme . . . . .	31
3. Refinement . . . . .	32
4. Location of the Hydrogen Atoms . . . . .	35
5. Secondary Extinction Correction . . . . .	36
6. Evaluation of the Refinement . . . . .	37
7. The Gaussian Ellipsoid Approximation . . . . .	39
V. Results . . . . .	43
VI. Discussion . . . . .	54
PART II: N-FORMYL-L-PHENYLALANINE: $\alpha$ -CHYMOTRYPSIN COMPLEX, THE EFFECT OF BINDING OF A COMPETITIVE INHIBITOR SUBSTRATE-LIKE MOLECULE	
VII. Introduction . . . . .	65
1. Chymotrypsin . . . . .	65
2. N-Formyl-Phenylalanine (NFP) . . . . .	67



IX. Experimental . . . . .	76
1. Crystal and Derivative Preparation . . . . .	76
2. Intensity Data Collection . . . . .	77
3. Crystal Motion . . . . .	80
4. Lack of Balance and Twin Corrections . . . . .	83
5. Absorption Correction . . . . .	83
6. Decay Correction . . . . .	89
7. Data Processing . . . . .	90
X. Discussion of the Results . . . . .	95
1. Substitution in the Active Site Region . . . . .	95
2. Histidine 40 . . . . .	103
3. Other Similarities with pH Change . . . . .	105
4. Final Comments . . . . .	108
REFERENCES . . . . .	111
APPENDICES	
APPENDIX I The Reciprocal Lattice . . . . .	115
APPENDIX II The Observed and Calculated Structure Factors of Tetrapropylporphine . . . . .	116
APPENDIX III The Amino Acids and the Sequence of $\alpha$ -Chymotrypsin .	124



RE

I.

II

III

IV

V

VI

VII

VIII

IX

X

XI

XII

XIII

XIV

XV

XVI

# LIST OF TABLES

## TABLE

I. The Positions of $\bar{1}$ in $P2_1/c$ . . . . .	23
II. Statistical Distribution of $ E $ 's . . . . .	25
III. Weighting Scheme for TPrP Least Squares Refinement . . .	33
IV. Reflections Corrected for Extinction . . . . .	38
V. Final Atomic Parameters, Carbon and Nitrogen . . . . .	44
VI. The Mean Square Displacements (in $\text{\AA}^2$ ) for the Carbon and Nitrogen Atoms . . . . .	46
VII. Final Hydrogen Atom Parameters . . . . .	47
VIII. Atomic Deviations from the Least Squares Plane of Individual Pyrrole Rings . . . . .	48
IX. Data Ranges for the Enzyme . . . . .	78
X. The Cell Parameters for NFP: $\alpha$ -Chymotrypsin . . . . .	81
XI. Record of Crystal Motion . . . . .	84
XII. Ratios of the Intensities of Real and Twin Lattices . . .	85
XIII. Decay Correction Data . . . . .	91
XIV. $R_r$ Factors for the Redundant Reflections . . . . .	93
XV. Positions of Structural Changes of $\alpha$ -CHT with pH from Vandlen and Tulinsky (60) . . . . .	107
XVI. The Major Peaks in the NFP Difference Electron Density .	109

Figure

1.

2.

3.

4.

5.

## LIST OF FIGURES

### Figure

1. Porphine (a), chlorophyll a (b), and the tetrapyrrole portion of vitamin B <sub>12</sub> (c) . . . . .	2
2. The visible spectra of tetraphenylporphine, tetrapropylporphine and porphine . . . . .	5
3. Three models for hydrogen bonding in porphine . . . . .	6
4. The path of the ring current (a) and two major resonance forms of porphine (b,c) . . . . .	8
5. Schematic of a four-circle diffractometer. (a) at $\chi = 90^\circ$ and (b) at $\chi = 0^\circ$ . . . . .	13
6. Flow diagram of automatic data collection from Vandlen and Tulinsky (19) . . . . .	18
7. The percent error curve for porphine from Chen (15). . . . .	34
8. Plots of the observed and calculated weighting schemes . . . . .	40
9. Labeling scheme for TPrP and the deviations from the NLS plane (in A) . . . . .	45
10. Bond distances (in A) and angles (in degrees) . . . . .	49
11. Computer plot of the TPrP molecule . . . . .	50
12. The composite electron density . . . . .	51
13. The packing along $a^*$ . . . . .	52
14. Dominant resonance structures of the porphine macrocycle (a,b). Average observed distances and angles of the independent pyrroles of TPrP (c) . . . . .	56
15. Average structure for free base porphine . . . . .	58
16. Scale representation of the two types of pyrrole rings . . . . .	59
17. Average structure for porphine and TPP with the differences ( $\Delta$ ) between the pyrrole rings . . . . .	61

18.

19.

20.

21.

22.

23.

24.

25.

26.

27.

28.

18.	Axial diffraction patterns for Native, NFP (2 weeks), and NFP (6 weeks) . . . . .	69
19.	Histogram of the number of reflections in each figure of merit range for $\alpha$ -CHT, from Tulinsky, <u>et al.</u> (53) . . .	75
20.	Plot of the intensity of the (0,18,0) <u>vs.</u> exposure time. .	82
21.	Plot of the relative absorption observed at $\chi = 90^\circ$ for crystals 1, 3, and 5 . . . . .	87
22.	Plot of the relative absorption observed at $\chi = 90^\circ$ for crystals 2 and 4 . . . . .	88
23.	The radial distribution of Native (solid line) and NFP:Native (dashed line) . . . . .	94
24.	(a) Electron density in the active site region viewed perpendicular to the local two-fold axis . . . . .	96
	(b) Distances (in A) from NFP in the electron density of Figure 24(a) . . . . .	97
25.	(a) Electron density in the active site region parallel to the two-fold axis . . . . .	99
	(b) Distances (in A) from NFP in the electron density shown in Figure 25(a) . . . . .	101
26.	Distances (in A) from the local two-fold axis (heavy line)	102
27.	The region near Histidine 40 . . . . .	104
28.	The uranyl substitution position . . . . .	106

PART I.

THE CRYSTAL AND MOLECULAR STRUCTURE OF TETRA-N-PROPYLPORPHINE

## 1. General

The p

complex r

porphyrin

processes

transport

porphyrin

Most

pyrrole

and vita

positio

through

and the

xiety

intere

Th

resona

(3) to

to de

units

stabi

percept

of a



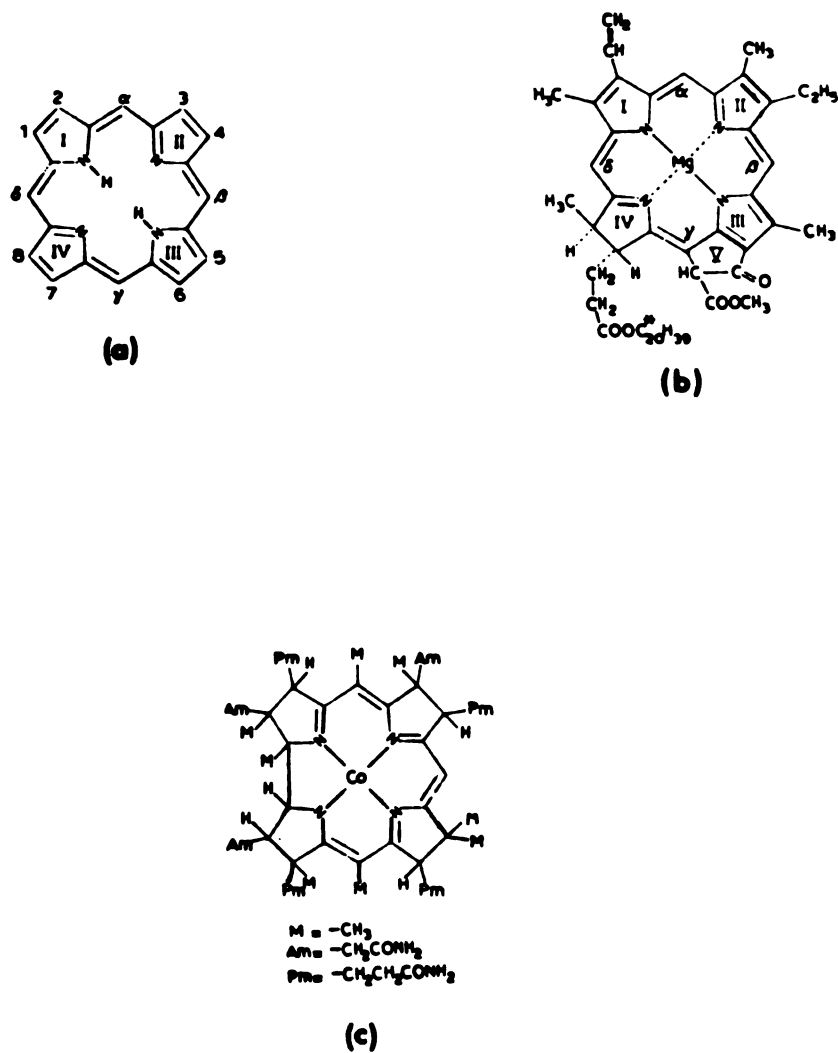
## I. Introduction

### 1. General

The porphyrins are tetrapyrrolic macrocycles with the ability to complex readily with many different metals. As metal chelates, porphyrin-like materials are important in a number of biological processes including photosynthesis, cellular respiration, and electron transport. Excellent reviews of the chemistry and function of porphyrins and their related compounds are available (1,2).

Most of the porphyrins of animal origin are substituted at the pyrrole positions 1-8 (Figure 1a); however, the chlorophylls (Figure 1b) and vitamin B<sub>12</sub> (Figure 1c) have side chains at the bridge carbon positions ( $\alpha, \beta, \gamma, \delta$ ). The porphyrins can have a wide range of functions through variations in the stereochemistry of the macrocyclic structure and the side chains. The correlation of the function of the porphyrin moiety with its stereochemistry makes these substances particularly interesting from a structural viewpoint.

The porphyrins are highly conjugated and intensely colored. The resonance stabilization for porphine was determined by Longo, *et al.*, (3) to be 419 kcal/mole. This large stabilization energy was attributed to delocalization of  $\pi$ -electrons available from the several identical units of 2-vinylpyrrole linked in conjugation. The strong resonance stability of porphyrins is indicated in the mass spectra of the porphyrins also. Adler, Green, and Mautner (4) found that the formation of a stable composition  $\pi$ -electron system with an effective number of



**Figure 1.** Porphine (a), chlorophyll a (b), and the tetrapyrrole portion of Vitamin B<sub>12</sub> (c).

electron

the first

monocycle

molecule.

effects of

1. Influ

The

electron

are a cl

best ba

region w

compound

of comb

spectrum

of visib

and the

Th

Spectra

Antenna

result

in the

listed

forbid

transit

displac

sensit.

ring.

$\pi$ -electrons equal to  $4S + 2$  (where  $S$  is an integer) was the major factor in the fragmentation of the porphyrin ion. Hence, the charge on the macrocycle is positioned so as to maintain the aromaticity of the basic molecule. Such strong aromatic character should temper the electronic effects of any substituent on the ring.

## 2. Influence of Porphyrin Side Chains

The highly conjugated nature of the porphyrins causes their electronic absorption spectra to be particularly useful. The porphyrins have a characteristic intense absorption band at 400 m $\mu$  called the Soret band; in addition, there are two to four bands in the visible region whose wavelengths and relative intensities depend on the specific compound. Lemberg and Falk (5) found a correlation between the nature and number of the substituents on the porphyrin molecule and its visible spectrum. For the naturally occurring porphyrins there are four types of visible spectra which are related to the positions of the substituents and their electronegativities.

This correlation of substituent effects to changes in the visible spectra is explained by the theory of porphyrin spectra developed by Gouterman (6). For a square planar porphyrin the visible bands are a result of transitions from filled  $A_{2u}$  orbitals to vacant  $E_g$  orbitals. In the free base the symmetry is lowered to  $D_{2h}$  and the degeneracy is lifted producing four bands instead of two which arise from the forbidden (0-0) and the allowed (0-1) vibrational transitions. The transitions producing these bands are associated with electronic displacement to the periphery of the ring, thus accounting for the sensitivity of these absorptions to changes on the exterior of the ring.

The s

monstra

(27), an

re subst

respecti

for the

The

their nu

rising

case th

to low f

found to

pyroles

in the b

the porp

1. Hydr

In

hydroge

N-H...N

frequen

the N-H

1500 cm

binding

signifi

intramo

The

Waven (

The sensitivity of the visible bands to changes in side chains is demonstrated in the visible spectra of porphine, tetraphenylporphine (TPP), and tetrapropylporphine (Figure 2), where the  $\alpha, \beta, \gamma, \delta$  positions are substituted with hydrogen atoms, phenyl groups, and n-propyl groups, respectively. The bands vary in relative intensity and in wavelength for the three compounds.

The effect of substitution on the porphyrins is also observed in their nuclear magnetic resonance (NMR) spectra (7). The ring currents, arising from the delocalized  $\pi$ -electrons, in porphyrins are large and cause the peaks for protons on the periphery of the ring to be shifted to low field. The peak from the methine (bridge) hydrogen atom was found to be very sensitive to the symmetry of the substitutions on the pyrroles. Abraham, *et al.*, (8) found that substitution of methyl groups on the bridge positions decreased the ring current or aromaticity of the porphyrin by approximately 10%.

### 3. Hydrogen Bonding

In the free base porphyrins the close proximity of the central hydrogen atoms to all four nitrogen atoms suggests the possibility of N-H...N hydrogen bonding. Badger, *et al.*, (9) found the N-H stretching frequency to be  $3320 \text{ cm}^{-1}$  for a number of free base porphyrins whereas the N-H stretching frequency in pyrrole, indole, and carbazole is  $3500 \text{ cm}^{-1}$ . This decrease in stretching frequency indicates hydrogen bonding; furthermore, the N-H stretching frequency does not change significantly for solids suggesting that the hydrogen bonding is intramolecular.

Three possible models for the hydrogen bonds were suggested by Mason (10); these are given in Figure 3. The symmetrical model

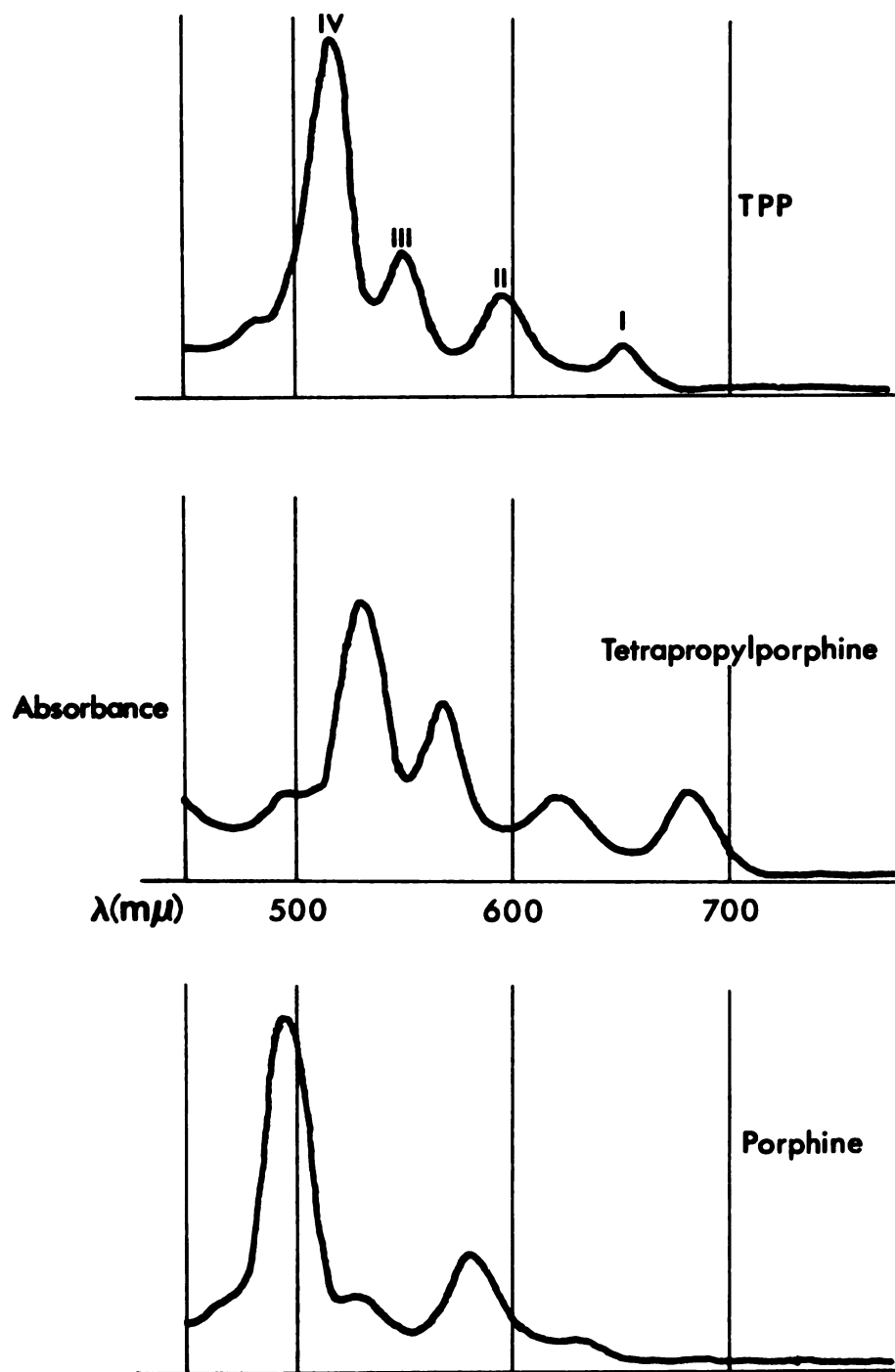


Figure 2. The visible spectra of tetraphenylporphine, tetrapropylporphine and porphine. Solvent for all three was benzene.





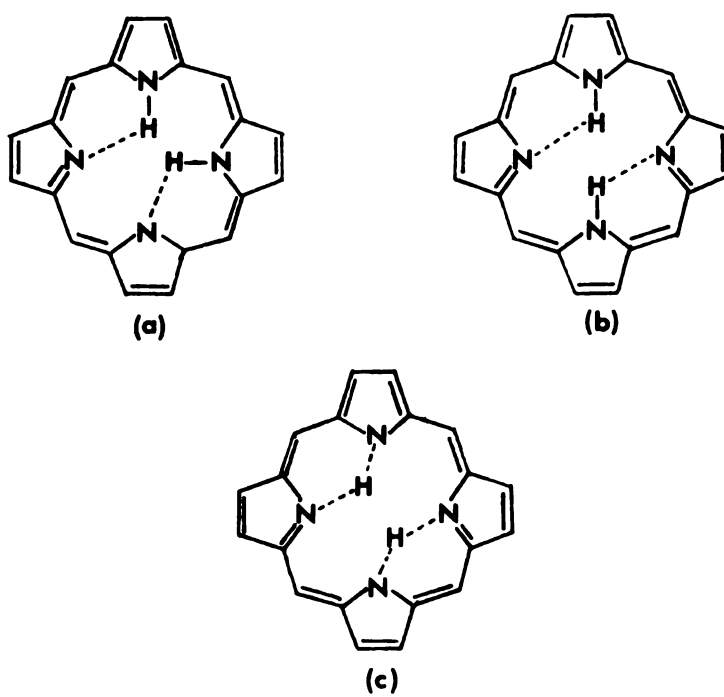
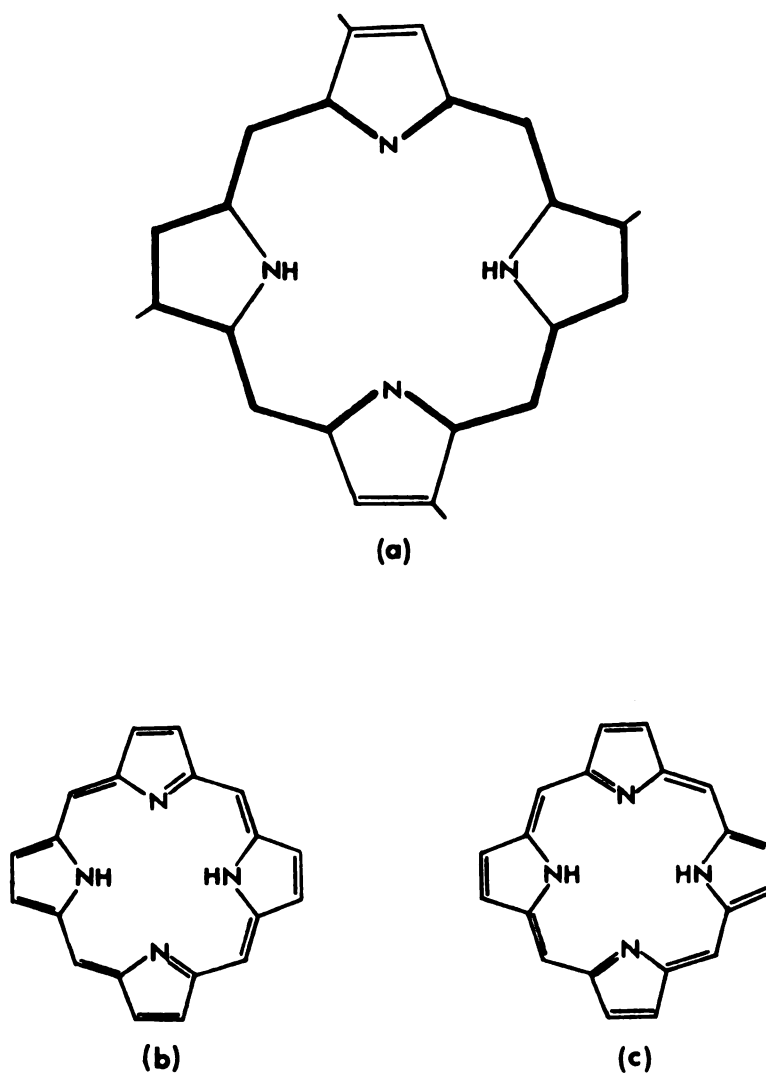


Figure 3. Three models for hydrogen bonding in porphine.

(Figure 3c) was rejected since the N-N distance of 2.65 Å (from phthalocyanine) would result in the N-H bond being too long; since the interior of porphine rings tend to be larger than for phthalocyanines this model becomes even less likely. Mason identified three N-H vibrations in the infrared spectrum of porphine which correspond to the  $D_{2h}$  symmetry of the opposite tautomer (Figure 3b); the adjacent tautomer with  $C_{2v}$  symmetry would have five bands. The conclusion from the infrared studies was that porphyrins have the imino hydrogen atoms bonded to opposite pyrrole rings and that intramolecular hydrogen bonding is present.

#### 4. The Central Hydrogen Atoms

With the acceptance of the opposite arrangement of the imino hydrogen atoms attention was directed to the possibility of rapid tautomerism of the N-H protons. Becker, Bradley, and Watson (11) studied the NMR spectrum of coproporphyrin-1 which has methyl groups substituted on the pyrrole rings in positions 1,3,5,7 (Figure 1a). In Figure 4 the path of the ring current for the opposite arrangement of imino hydrogen atoms is given along with the two major resonance structures which contribute to the hybrid. Methyl groups in positions 1 and 5 are further from the ring current than the other two methyl groups; hence, two methyl hydrogen NMR lines were expected. Becker, et al., (11), observed one single peak for the methyl hydrogen atoms; this result was explained by rapid tautomerism of the imino protons which would make all four pyrrole rings magnetically equivalent. At  $-63^{\circ}\text{C}$  no broadening of the  $\text{CH}_3$  line was observed which indicated that the tautomerism rate was greater than 200 exchanges per second.



**Figure 4.** The path of the ring current (a) and two major resonance forms (b,c) of porphine.

In the crystal structure of porphine reported by Webb and Fleischer (12) the molecule is observed to have  $D_{4h}$  symmetry with half-hydrogens bonded to each of the four nitrogen atoms. This result was explained as due to the tautomerism of the N-H protons.

In the crystal structure of triclinic tetraphenylporphine (TPP) (13) the imino hydrogen atoms are bonded to opposite nitrogen atoms and there are slight differences in the two types of pyrrole rings resulting in a macrocycle with  $D_{2h}$  symmetry and no evidence of tautomerism. The tetragonal form of TPP (14) is required to have  $D_{4h}$  symmetry because of the space group and cannot therefore provide any information about tautomerism.

The structural investigation of  $\alpha,\beta,\gamma,\delta$ -tetra-n-propylporphine (TPrP) was undertaken for two reasons: (1) to obtain an accurate structure of the free base macrocycle and (2) to determine the effect, if any, of the aliphatic substituents on the aromaticity of the porphyrin moiety. The TPP structural determination has indicated a macrocyclic structure which corresponds to a hybrid of the two major resonance structures (see Figure 4). Since the time the TPrP structure determination was begun, Chen (15) has redetermined the structure of porphine and found it to be similar to TPP. Thus there were two examples of the hybrid structure of the macrocycle, both the parent compound and one with aromatic side chains. The aliphatically substituted molecule (TPrP) could provide more evidence for the common macrocycle if its structure is similar. Also the structural analysis of TPrP could determine if the electronic effects which produce differences in the visible spectra of TPP, TPrP, and porphine (Figure 2) also produce differences in their structures.

1. Gr

The

of test

plate)

approx

the X-

diffin

flotta

be 1.

1. S

T

the 1

conv

recip

lact

used

thei

lips

be r

lact

recip

## II. Experimental

### 1. Crystals

The crystals were obtained by slow evaporation of a benzene solution of tetrapropylporphine in the dark; the crystals are purple hexagonal platelets. The crystal which was used for all X-ray measurements was approximately 0.15 mm in each dimension. This crystal was smaller than the X-ray beam and of sufficient size and quality to produce a strong diffraction pattern. The density of a TPrP crystal, measured by flotation in an aqueous silver nitrate solution, was determined to be  $1.22(2) \text{ g cm}^{-3}$ .

### 2. Space Group Determination

The geometry of the diffraction pattern must be determined before the intensity measurements can be made. Photographic methods are a convenient way to determine the size and symmetry of the direct and reciprocal lattices (see Appendix I for a discussion of the reciprocal lattice). The oscillation, Weissenberg and precession methods were used in this study. Detailed treatments of photographic methods and their uses have been presented by Stout and Jensen (16) and Henry, Lipson, and Wooster (17).

In the oscillation method the crystal is mounted such that it can be rotated about a direct lattice line and oriented so that this lattice line is perpendicular to the incident X-ray beam. The reciprocal lattice planes which are perpendicular to the lattice line

can then be brought into reflecting position through oscillation of the crystal. The diffraction pattern is recorded on photographic film which is placed in a cylindrical camera around the oscillation axis. The pattern appears as a series of separated layer lines each made up of individual spots. As well as being useful for aligning the crystal, the oscillation photograph can provide information regarding the symmetry of the lattice line and the spacing of the lattice along the oscillation line. A  $10^\circ$  oscillation photograph of the TPrP crystal showed no symmetry which indicated that the oscillation axis was not a symmetry axis; in addition, the separation of the layer lines indicated that the lattice spacing was 5-6 Å.

In the Weissenberg film method, the cylindrical camera is used to record only one layer line, which is selected by means of a slotted screen. As the crystal is rotated the film is synchronously translated past the slot giving a two dimensional picture of the reciprocal lattice plane. Zero and first layer Weissenberg photographs were taken of the TPrP crystal. These photographs showed two axes  $90^\circ$  apart, one of which was not perpendicular to the rotation axis. However, neither the identification of the crystal class nor the indexing of the Weissenberg photographs could be accomplished until the oscillation axis had been identified.

The precession method provides a means for observing the reciprocal lattice planes parallel to the oscillation axis. The precession motion is achieved by two oscillations: one about the oscillation axis of the crystal and the other about a line  $90^\circ$  from the oscillation axis and parallel to the X-ray beam. The flat photographic film is tilted to be parallel to the reciprocal lattice planes being photographed and follows the same precession motion as the crystal. A screen with an

annular opening is used to select the desired reciprocal lattice plane to be photographed. Precession photographs for TPrP were taken of the planes defined by the oscillation axis and the two axes previously identified in the Weissenberg photographs. These photographs indicated that the oscillation axis was not a cell axis and subsequent indexing revealed that this axis was the  $(h0\bar{h})$  lattice line. The crystal class was determined to be monoclinic and the unique axis,  $\underline{b}$ , and the  $\underline{c}$  axis were the two axes observed in the Weissenberg photographs. When indexed the Weissenberg photographs revealed systematic absences for the  $(0k0)$  reflections for odd values of  $k$ . These absences indicate the presence of a two-fold screw axis along the  $\underline{b}$  axis (a two-fold rotation coupled with a translation of  $b/2$  along the rotation axis). The precession photographs of the  $(h0\ell)$  and  $(h\ell\ell)$  zones showed another systematic absence for reflections of the type  $(h0\ell)$  whenever  $\ell$  was odd. This type of extinction is characteristic of a  $\underline{c}$  axis glide plane which corresponds to reflection through a plane parallel to the  $\underline{c}$  axis and a translation by  $c/2$  along the  $\underline{c}$  axis.

The fact that TPrP was monoclinic along with the observation of systematic extinctions in the Weissenberg and precession photographs fixed the space group to be uniquely determined as centrosymmetric  $P2_1/c$ , space group no. 14, subgroup number 2, as tabulated in the International Tables of X-ray Crystallography, Volume I (18).

### 3. Intensity Data Collection

The crystal was mounted on a 4-circle diffractometer with the  $(h0\bar{h})$  lattice line coincident with the  $\phi$  axis of the diffractometer (Figure 5 gives the geometry of the four circle diffractometer). The  $\underline{b}$  axis was thus located in the equatorial plane ( $\chi = 0^\circ$ ). The





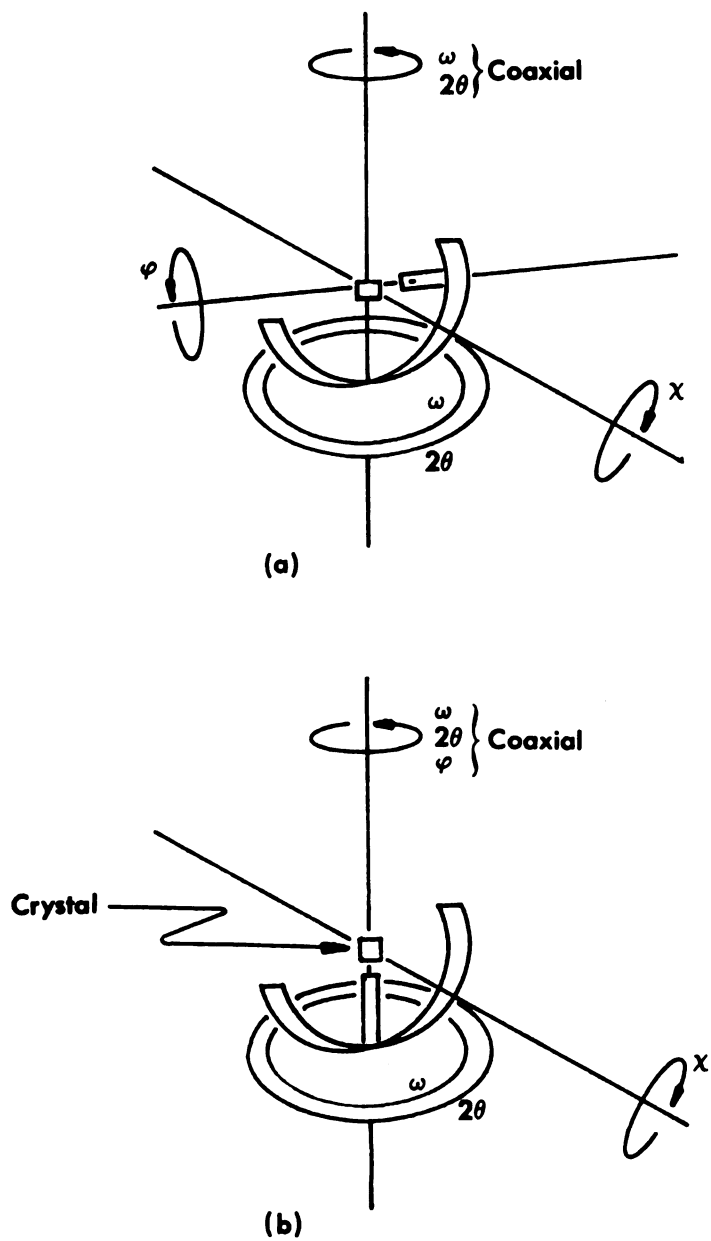


Figure 5. Schematic of a four-circle diffractometer. (a) at  $\chi = 90^\circ$  and (b) at  $\chi = 0^\circ$ .

unit c

of two

distr

orien

coord

are a

per u

obser

posi

are

situ

can

of t

who

diff

to b

mos

mea

pea

ang

ste

(36

mea

wid

que

unit cell dimensions were determined by measuring the angular positions of twelve centered, moderately intense reflections which were well distributed in reciprocal space. The unit cell parameters and the orientation parameters of the crystal were obtained from the angular coordinates by the method of least squares. The unit cell parameters are  $a = 5.078(5)$ ,  $b = 11.59(2)$ ,  $c = 22.39(3)$  Å and  $\beta = 99.50(5)^\circ$ .

The calculated density of the crystal on the basis of two molecules per unit cell is  $1.223 \text{ g cm}^{-3}$  which compares favorably with the observed density of  $1.22 \pm .02 \text{ g/cm}^3$ . Since there are four equivalent positions in each unit cell for space group  $P2_1/c$ , the fact that there are only 2 molecules/unit cell fixes each molecule of TPrP to be situated on a center of symmetry.

The quality of the crystal and the quadrant for data collection can be determined by examination of the mosaic spreads of the intensities of the reflections. The spread arises because the small mosaic blocks which comprise the crystal are not perfectly aligned so each one diffracts from a slightly different angle. If the intensity data are to be obtained with a non-scanning technique, it is necessary that the mosaic spreads be narrow enough that all of the diffracted beam can be measured at one time thus obtaining the integrated intensity of the peak.

The mosaic spreads were obtained by centering the peak with the angles  $\phi$ ,  $\chi$ , and  $2\theta$  and then measuring the intensity of the peak at steps of  $0.02^\circ$  in  $\omega$  across the peak. The intensities for the  $(060)$ ,  $(0\bar{6}0)$ ,  $(00,12)$ ,  $(00,\bar{1}2)$ ,  $(100)$ , and  $(20\bar{2})$  reflections were measured. The mosaic spreads were symmetrical, single-peaked, and had widths less than  $0.3^\circ$  from background to background. Therefore, the quadrant choice was based on convenience and was defined by the

$+c^*$  and  $-c^*$  axes including the  $+b^*$  axis and the  $+a^*$  axis. To avoid  $K_{\alpha_1}$ ,  $K_{\alpha_2}$  splitting, the intensity data were collected to  $\sin \theta/\lambda \leq 0.55$  corresponding to a minimum spacing of 0.94 Å.

The intensity data were collected using  $\text{CuK}_{\alpha}$  radiation ( $\lambda = 1.5418$  Å) and a Picker Four Circle Automatic X-ray Diffractometer controlled by a Digital Equipment Corp. (DEC) 4K PDP-8 computer (FACS-I System) coupled to a DEC 32K Disc File.

Balanced Ni and Co filters were used to ensure that the incident radiation would be monochromatic within a fixed window width of 0.12 Å. Monochromatic radiation improves the peak to background ratio of the diffraction pattern. These filters consist of a metal foil of an element which has an absorption edge at the wavelength of the unwanted X-rays. An absorption edge occurs when the X-radiation is energetic enough to excite an electron out of an atomic orbital. For a filter of an element with atomic number  $Z-1$  the wavelength of the absorption edge lies between the  $K_{\alpha}$  and  $K_{\beta}$  lines of the target element with atomic number  $Z$ .

A system of balanced filters provides a means of obtaining X-radiation having a very narrow range of wavelengths which may be used to measure the peak intensity of a particular reflection. In this manner the peak is first measured using a  $K_{\beta}$  filter (Ni in the case of  $\text{CuK}_{\alpha}$  radiation) then the background is measured with a balanced  $K_{\alpha}$  filter (Co) which transmits the same amount of  $K_{\beta}$  radiation as the  $K_{\beta}$  filter while removing some of the  $K_{\alpha}$  radiation (about 10% transmission of  $K_{\alpha}$ ). The final intensity is calculated from the difference between the two measurements.

The intensity measurements were made under the control of a computer program using an  $\omega$  step scan procedure (19). This is essentially a

stationary crystal-stationary counter technique which measures the intensity of six steps in  $\omega$ , in  $0.03^\circ$  increments, through the calculated position of the peak. At each step the intensity is measured for four seconds, using a balanced Ni filter, and the four largest measurements are summed to give the integrated intensity (count-six-drop-two; Wyckoff, et al., (20)). At the  $\omega$  value of the step with the highest intensity, the background is measured for four seconds, using the Co balanced filter, and the intensity is multiplied by four to give the total background. If the crystal has undergone some misalignment the calculated value of  $\omega$  will not correspond to the observed peak. To compensate for these small changes in  $\omega$ , the program checks the  $\omega$  value of the step corresponding to the highest intensity and the value of the intensity and if this step is not at the center of the range and if the intensity is greater than some preset value, then one or two additional steps are made depending on the position of the highest intensity. This "wandering"  $\omega$  step scan procedure allows the computer to continue to collect useful data even when some small misalignment has occurred.

Throughout the intensity data collection, the intensities of three reflections were measured every 100 reflections (about 1 hour) as a check on the alignment of the crystal. The computer program (19) automatically detects any misalignment from a decrease in the intensities of the monitor reflections and realigns the crystal by remeasuring the angular settings of the orienting reflections and obtaining a new orientation matrix through least squares analysis. The monitor reflections were chosen to give a description of the orientation of the reciprocal lattice; they were the (060) reflection at  $\chi = 0^\circ$  and the  $(20\bar{2})$  reflection at  $\chi = 90^\circ$  measured at two  $\phi$  values  $90^\circ$  apart. A



given

**F**

22

recd.

27e

1000

the

22

12

11.

rel

for

522

22

 $\dot{I} =$ 

20

Re

CC:

ie

72

11

;



Si

flow diagram of the entire automated process of data collection is given in Figure 6.

Before the intensity data can be converted to structure factor amplitudes and used to determine the crystal structure, it may be necessary to apply several corrections to the data. These corrections are absorption, lack of balance, decay, and Lorentz-polarization.

The amount of absorption of X-rays by a crystal depends on the content of the unit cell and the path lengths through the crystal of the incident and diffracted beams. The absorption by one unit cell is determined by the mass absorption coefficient at the appropriate X-ray wavelength(18) of each element comprising the unit cell. The linear absorption coefficient of  $9.11 \text{ cm}^{-1}$  for the TPrP unit cell was relatively small. Calculation of the path length of the X-ray beam for all the reflections used in a structure analysis would be difficult, especially for a crystal of general shape. Furnas (21) has suggested that the variation of the absorption in the crystal can be measured by the variation in the intensity of a reflection with  $\phi$  at  $\chi = 90^\circ$ . In this orientation the normal to the reflecting planes is coincident with the  $\phi$ -axis so that a rotation about  $\phi$  does not move the planes from the reflecting position and the intensity should remain constant. Thus, any variation in the intensity of the reflection would be due to a difference in the X-ray absorption or pathlength. This variation would be a measure of the relative absorption by the crystal. For TPrP the variation of the intensity of the  $(20\bar{2})$  reflection with  $\phi$  was measured for the  $\phi$  values used in the data collection. Due to the small value of the absorption coefficient and the nearly isotropic shape of the crystal a reliable variation of absorption could not be



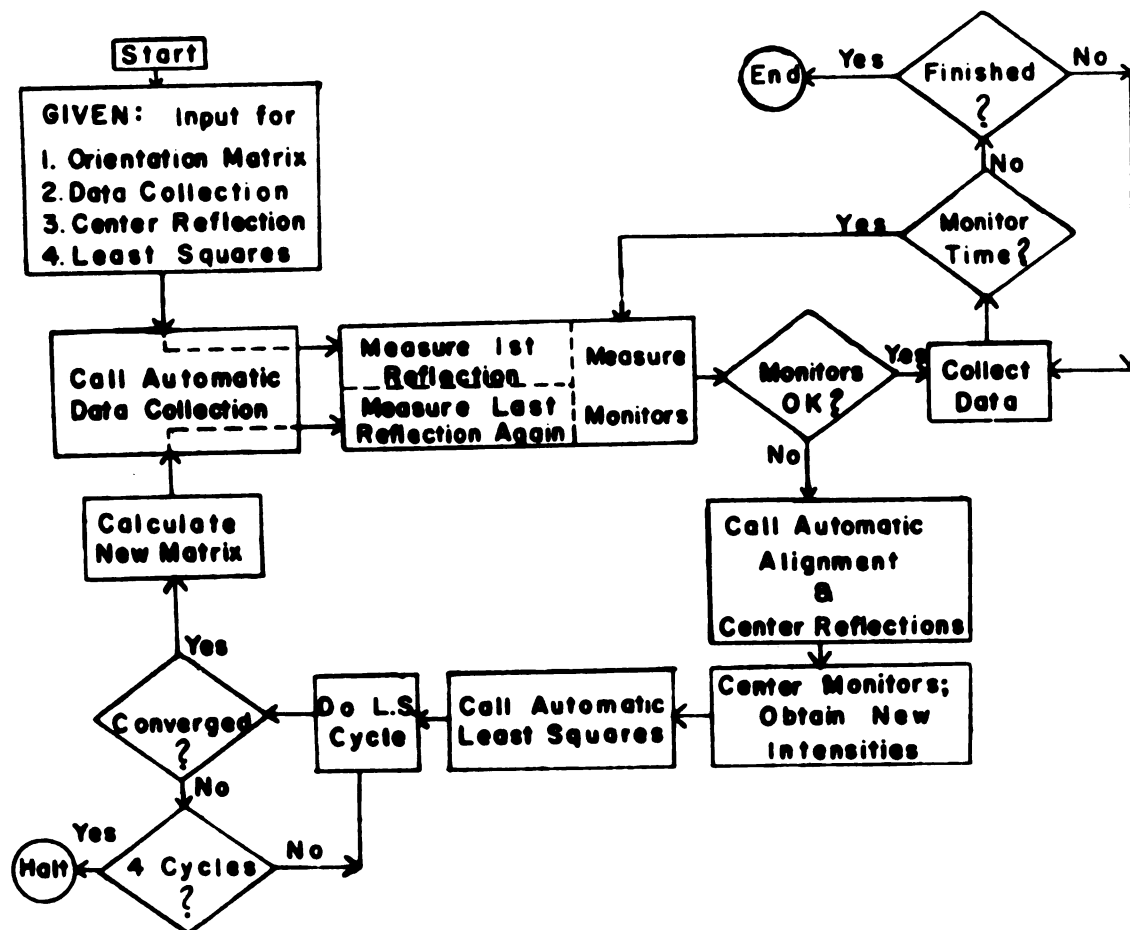


Figure 6. Flow diagram of automatic data collection from Vandlen and Tulinsky (19).

2255

2256

2257

2258

2259

2260

2261

2262

2263

2264

2265

2266

2267

2268

2269

2270

2271

2272

2273

2274

2275

2276

measured; therefore, no absorption correction was applied to the intensity data.

The Ni/Co filter pair which was used for data collection was not balanced exactly. This lack of balance was determined as a function of the scattering angle ( $2\theta$ ) by measuring the general background with both the Ni and Co filters. A plot of the differences between these background measurements was used to correct the background intensities for lack of balance. This correction was small for the TPrP data crystal and decreased to zero at about  $2\theta = 40^\circ$ .

The intensities of the monitor reflections were compared to determine if the crystal had sustained any damage due to X-ray exposure. The monitor reflections were measured 22 times over a period of approximately 20 hours and the percent standard deviations from the average intensity were 0.5, 0.9, and 1.0% for the (060) and the two ( $20\bar{2}$ ) reflections, respectively. Thus there was no decrease in intensity due to X-ray exposure and no correction for decay was applied.

In all, the intensities of 1744 independent reflections were measured. In addition, a redundant set of ( $0k\ell$ ) reflections were measured and used to calculate  $R_{OBS}$  which is

$$R_{OBS} = 2 \sum |I_{0k\ell} - I_{0k\bar{\ell}}| / \sum (I_{0k\ell} + I_{0k\bar{\ell}}).$$

The resulting  $R_{OBS}$  value of 0.011 for 382 redundant reflections indicates that the data are highly reproducible. The average value of the measured intensities for the ( $h0\ell$ ) systematically absent reflections were used to fix an observable limit of 15 counts/16 seconds. This limit gave 331 accidentally absent reflections and 1413 (81%) reflections were taken to be observable.

Since the absorption and decay corrections were negligible, the data were converted to structure amplitudes by simply applying Lorentz-polarization ( $L_p$ ) and lack of balance corrections.

### III. Solution of the Phase Problem

#### 1. The Phase Problem

Due to the periodicity of the molecular arrangement in a crystal the electron density,  $\rho(x,y,z)$ , may be represented as a Fourier series

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_{\ell=-\infty}^{+\infty} F(hk\ell) \exp\{-2\pi i(hx+ky+\ell z)\} \quad (1)$$

The Fourier coefficients, known as structure factors, are complex quantities related to the observed intensity data. Unfortunately, the phase relations among the diffracted waves can not be observed so only the intensities of the waves are obtained. Thus, the structure amplitudes,  $|F(hk\ell)|$ , can be obtained from the data but the phases must be deduced. This is the phase problem in crystallography and the utility of the method depends upon its solution.

#### 2. The Patterson Function

The Patterson function,  $P(u,v,w)$ , is defined as the product of two electron density functions:

$$P(u,v,w) = V \int_0^1 \int_0^1 \int_0^1 \rho(x,y,z) \rho(x+u,y+v,z+w) dx dy dz. \quad (2)$$

This function will be nonzero only when the electron density at the points  $(x,y,z)$  and  $(x+u,y+v,z+w)$  is nonzero; hence, the Patterson function will have positive regions at the positions representing the vectors between atoms. This function is useful in structure analysis

because it is a phaseless Fourier series which can be calculated directly from the observations. The Patterson function in this form is

$$P(u,v,w) = \frac{1}{V} \sum_h \sum_k \sum_{\ell=-\infty}^{+\infty} |F(hk\ell)|^2 \cos 2\pi(hu+kv+\ell w), \quad (3)$$

which is also a centrosymmetric function.

The three dimensional Patterson function for TPrP was computed using the squares of the observed structure amplitudes as coefficients. The positive regions of this function were concentrated in planes parallel to the (114) plane. If the vectors between atoms lie close to this plane, the atoms of the molecule must also lie near the plane. As further support for the conclusion that the (114) planes are parallel to the approximate molecular plane the three largest intensities are those of the (114), (104), and (112) reflections.

The tilt of the molecule was thus defined by the requirements that the molecular plane be parallel to the (114) planes and the translation by the molecule being situated on a center of symmetry. The angular orientation of the molecule could be determined by systematically rotating a model structure in the plane until the best fit of the observed diffraction pattern was obtained.

The angular orientation was not determined in the manner described above because the molecule was originally positioned at the wrong origin. In space group  $P2_1/c$  there are four independent sets of inversion centers (these are given in Table I) and the rotational fit was attempted only for the set (0,0,0); (0,1/2,1/2). The extra symmetry of the Patterson function and the presence of cross-vectors between the two molecules in the unit cell obscured the real centers of symmetry of (0,0,1/2); (0,1/2,0) which were subsequently determined by direct methods.

Table I. The Positions of  $\bar{1}$  in  $P2_1/c$ 

1. (0,0,0);(0,1/2,1/2)
2. (1/2,0,0);(1/2,1/2,1/2)
3. (0,0,1/2);(0,1/2,0)
4. (1/2,0,1/2);(1/2,1/2,0)

### 3. Direct Methods

Since the Patterson function can be used to deduce the phases indirectly (which are calculated using the atomic coordinates obtained from the vectors) employing only the structure amplitudes, information of the structure must be contained in the intensities or their relative distribution. Direct methods are objective mathematical procedures for obtaining the phases from the intensities and are independent of the structure. The method makes use of one or more mathematical expressions which predict the phase of a reflection from the phases and magnitudes of combinations of other reflections.

In these methods the intensity data are reduced to a form which describes point atoms with all temperature effects removed. The structure amplitudes are corrected for vibrational motion and placed on an absolute scale, then converted to normalized structure factors,  $|E|$ 's. The conversion to  $|E|$ 's is necessary because the probability function for the phase of E is independent of the complexity of the structure; thus, the mathematical expressions are valid regardless of the number of atoms in the unit cell. E is defined by

$$|E(hkl)|^2 = \frac{|F(hkl)|^2}{\sum_{i=1}^N f_i^2(hkl)}, \quad (4)$$

when

fact

and

for

dist

trye

The

abs

will

dist

sim

the

red

and

and

prop

det

equ

ser

and

rel

con



where  $N$  is the number of atoms in the unit cell,  $f_i(hkl)$  is the scattering factor of atom  $i$  at the  $\theta$  value associated with the  $(hkl)$  reflection, and  $c$  is a symmetry factor which accounts for the multiplicity of  $|F(hkl)|$  in symmetry operations in reciprocal space. The theoretical distribution of  $|E|$ 's for both centrosymmetric and noncentrosymmetric crystals may be obtained from the probability function for  $E$  (22). The TPrP data were converted to normalized structure factors with the absolute scale and average temperature factor ( $3.5 \text{ \AA}^2$ ) determined by Wilson's method (23). The  $|E|$ 's were scaled so that  $\langle E^2 \rangle = 1.0$ . The distribution of  $|E|$ 's, given in Table II, is as expected: they are similar to that of a centrosymmetric crystal.

In a centrosymmetric space group, the center of inversion causes the positions  $x, y, z$  and  $-x, -y, -z$  to be equivalent. This condition reduces the geometrical part of the structure factor to a cosine term and since the sine term is zero the phase angles must be 0 or  $\pi$  and the corresponding cosine terms must be  $\pm 1$ . Since the space group for TPrP is centrosymmetric, the phase problem is reduced to the determination of the signs of the  $F(hkl)$ 's.

The mathematical phase relationship used for TPrP is Sayre's equation (24) which is derived from the properties of squared Fourier series with the assumption that the electron density is non-negative and contains resolved atoms. This equation is similar to the  $\sum_2$  relationship developed by Hauptman and Karle (25) from probability considerations and may be expressed with normalized structure factors as

$$S(\underline{E}_{\underline{A}}) = S\left(\sum_{\underline{A}=\underline{B}+\underline{C}} \underline{E}_{\underline{B}} \underline{E}_{\underline{C}}\right), \quad (5)$$

Table II. Statistical Distribution of  $|E|$ 's

	<u>TPrP</u>	<u>Centric</u>	<u>Theoretical</u> <u>Acentric</u>
$\langle  E  \rangle$	0.83	0.80	0.89
$\langle E^2 \rangle$	1.00	1.00	1.00
$\langle  E^2 - 1  \rangle$	0.90	0.97	0.74
% > 1.0	31.08	32.00	37.00
% > 2.0	3.66	5.00	1.80
% > 3.0	0.50	0.30	0.01

where  $S(\ )$  means "the sign of",  $\underline{A}$ ,  $\underline{B}$ , and  $\underline{C}$  represent vectors of Miller indices  $(hkl)$ , and the sum is over all combinations for which  $\underline{A} = \underline{B} + \underline{C}$ . The sign of the sum in Equation (5) can be determined from a small number of terms if these terms are significantly larger than the rest. Thus, if the set of  $|E|$ 's is restricted to those with large values, i.e.,  $|E| > 1.5$  (26), only a partial sum is necessary. The probability function,  $P_+(E_{\underline{A}})$ , which gives the probability that the sign of  $E_{\underline{A}}$  is positive is defined by Cochran and Woolfson (27) to be

$$P_+(E_{\underline{A}}) \sim 1/2 + (1/2) \tanh \left( \frac{\sigma_3}{\sigma_2^{3/2}} \right) |E_{\underline{A}}| \sum_{\underline{A}=\underline{B}+\underline{C}} E_{\underline{B}} E_{\underline{C}} \quad (6)$$

where  $\sigma_m = \sum_{i=1}^N Z_i^m$ . The  $Z_i$  is the atomic number of atom  $i$  and  $N$  is the total number of atoms in the unit cell.

Sayre's equation requires knowledge of the signs of some reflections before it can be used to predict new signs. A set of starting signs can be obtained by assigning the origin which is accomplished by choosing the phases of certain reflections. In a centrosymmetric monoclinic crystal there are eight possible origins (Table I) and the choice of any origin ( $\epsilon_1, \epsilon_2, \epsilon_3$ ) affects the structure factors in the following way:

$$F'(hkl) = F(hkl) \times (-1)^{2(h\epsilon_1 + k\epsilon_2 + l\epsilon_3)} \quad (7)$$

The signs of structure factors for reflections of the type  $(hkl) = (\text{even}, \text{even}, \text{even})$  are independent of the origin and are determined only by the structure. Reflections of other parities can be arbitrarily assigned phases which will determine the origin of the unit cell. For a monoclinic crystal three reflections are necessary to define the origin.

The origin-defining reflections should be chosen from the largest values of  $|E|$  and should have a large number of interactions in the

phase determining equation. In addition, these reflections must also be independent, i.e., parity of  $(hkl)$  must be different for each reflection. Each reflection must form a large number of useful terms in the Sayre's equation; thus, the reflection must have a large number of interactions to form products  $E_B E_C$ . The three reflections which were used for the TPrP sign determinations were the  $(112)$ ,  $(23, \overline{17})$  and  $(16, \overline{10})$  reflections with  $|E|$ 's of 2.55, 3.14, and 2.57, respectively. The number of interactions in the Sayre's equation were 104, 59, and 55, respectively. By assigning all three of these reflections positive signs the origin was defined as  $(0,0,0)$  (25).

Additional signs can be generated through the phase relationships of the space group. For  $P2_1/c$  these relationships are:

$$\begin{aligned} E(hkl) &= E(\overline{h}\overline{k}\overline{l}) \\ E(hkl) &= (-1)^{k+l} E(\overline{h}\overline{k}l) \\ E(\overline{h}\overline{k}l) &= (-1)^{k+l} E(hkl) \end{aligned} \quad (8)$$

Signs may also be obtained from the symbolic addition method in which reflections with large  $|E|$ 's and many interactions are assigned symbols and then phases based on these  $n$  symbols are generated. Since each symbol may be + or - there will be  $2^n$  possible solutions to the structure.

The computer program (28) which was used to calculate the signs for the TPrP structure generates all the possible sets of phases by applying Sayre's equation to the  $2^n$  starting sets. The starting set is used to determine additional signs which are then used in an iterative application of the Sayre's equation which is terminated when no new additions or changes are made to the list of signs.

Once all the possible sets of phases are calculated one or the true solution must be determined. One method is to calculate all  $2^n$  electron

pens

then

cons

inde

when

then

sum

exce

dev

bal

has

min

sta

gre

wer

thr

(11

sol

sol

wit

in

sol

density maps. The true solution will produce a map which gives the chemically correct structure. A more efficient method is to use a consistency index as a test for the correct structure. The consistency index,  $C$ , is defined as

$$C = \frac{\langle |E_{\underline{A}} \sum_{\underline{A}=\underline{B}+\underline{C}} E_{\underline{B}} E_{\underline{C}}| \rangle}{\langle |E_{\underline{A}}| \sum_{\underline{A}=\underline{B}+\underline{C}} |E_{\underline{B}}| |E_{\underline{C}}| \rangle}, \quad (9)$$

where the average is taken over all values of  $\underline{A}$ . If  $C$  is equal to one, then for each reflection there are no sign contradictions in the Sayre's sum. The true solution is usually the most consistent one. There are exceptions, however, when the phase relationships for a space group never change the sign of  $E$ . For these cases,  $P\bar{1}$  is an example, one false solution will have  $C = 1.0$ . The computer program used for TPrP has two other useful tests: the true solution always requires the minimum number of cycles and for the true solution the signs of the starting set are always predicted to be their starting values.

In the TPrP structure determination the 202  $|E|$ 's with values greater than 1.3 (giving a ratio of [number of signs/atom]  $\sim 10$  (26)) were used in the sign determination program.

A starting set of seven signs was used for TPrP; these were the three origin-defining reflections and four general reflections: the (114), (229), (116) and (14 $\bar{8}$ ). In each of the sixteen possible solutions all 202 signs were determined and in four of these the solutions were obtained in three cycles. Of these four, the solution with the highest consistency index,  $C = 0.828$ , also had no contradictions in the determinations of the starting signs and was taken as the true solution.

The phases of this set along with the  $|E|$ 's were used to calculate an E-map which is a synthesis of the electron density with  $E(hkl)$ 's as the Fourier coefficients. Since the  $|E|$ 's describe point atoms with no vibrational motion the resulting map is characterized by sharp peaks and a low background. The E-map for TPrP revealed the positions of all eighteen non-hydrogen atoms in the asymmetric unit. The atomic positions had an average peak height of 237 with a fluctuating background of  $\pm 70$ .

The trial structure obtained from the map with an average thermal parameter for all atoms was then used to calculate structure factors. The R-factor for this calculation was 0.31. R is defined as

$$R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \quad (10)$$

where  $|F_o|$  and  $|F_c|$  are the observed and calculated structure amplitudes, respectively, and the sum is over all of the observed reflections. The choice of this structure as the true one was subsequently confirmed by the successful refinement of the structure. At the completion of the refinement, the phases of the final structure (which are given in Appendix II) were compared to those obtained from direct methods and all of the 202 signs which were originally determined with Sayre's equation were correct.

#### IV. Structure Analysis

##### 1. The Method of Least Squares

After a trial structure consisting of all the non-hydrogen atoms has been obtained, the preliminary parameters can be improved by refinement. The method of least squares was used to refine the trial structure of TPrP.

The principle of least squares is to find the parameters which best describe the observations by minimizing the weighted sum of the squares of the differences between observed and calculated values. The structure factor, which is related to the observations in crystal structure analysis, can be represented as a function of the atomic parameters by

$$F(hkl) = \sum_{i=1}^N f_i(hkl) T_i(hkl) \exp\{2\pi i(hx_i + ky_i + lz_i)\} \quad (11)$$

where  $f_i(hkl)$  is the atomic scattering factor,  $T_i(hkl)$  is the thermal parameter,  $x_i$ ,  $y_i$ ,  $z_i$  are the coordinates of atom  $i$ , and  $N$  is total number of atoms in the unit cell.

The structure factor is not a linear function of the atomic parameters so the relationship is linearized by expanding Equation 11 in a Taylor series and truncating the series after the first term. The trial parameters have to be close to the true values for this approximation to be successful. The function which is minimized is

$$D = \sum_{i=1}^m \omega_i (|F_{oi}| - |F_{ci}|)^2 \quad (12)$$



where the sum is over all of the observed reflections and  $\omega_i$  is the individual weight of the observed structure factor. The weight of the reflection is related to the standard deviation ( $\sigma_i$ ) of  $|F_o|$  by  $\omega_i = 1/\sigma_i^2$ . The method will usually converge even though the equation for  $F_c$  is approximate and the resulting adjustments are not exact. A system of iterative cycles of refinement is used until the changes in the parameters are insignificant with respect to the standard deviations of the parameters.

In the least squares refinement, the overall scale constant, the three coordinates of each atom, and the individual temperature factors are the parameters which are varied. The individual temperature factor is used to describe the thermal motion of the atoms and is related to the mean-square amplitude ( $\overline{u^2}$ ) of atomic vibration. Each atom may be assumed to have an isotropic temperature correction,  $T(hkl) = \exp(-B \sin^2 \theta / \lambda^2)$  with one parameter,  $B = 8\pi^2 \overline{u^2}$ , to be determined or to have an anisotropic thermal correction of the form

$$T(hkl) = \exp[-(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)]. \quad (13)$$

These two forms of the thermal parameter are interrelated by  $\beta_{11} = B a^{*2}/4$ ,  $\beta_{12} = B a^* \cdot b^* \cos \gamma^*/4$ , etc.

The scattering factors for the carbon and nitrogen atoms were taken from Cromer and Waber (29). A spherical scattering factor (30) was used for the hydrogen atoms.

## 2. The Weighting Scheme

A weighting scheme for least squares refinement is based on the estimated errors of the intensity measurements. The weighting scheme, which was chosen to be similar to that of Hughes (31), is given in

Table III. The percent error curve which was used for the small intensities was constructed to be similar to that used in porphine (Figure 7 (15)) with limits of 100% error for  $I = 15$  counts/16 seconds and 4% error for  $I = 500$  counts/16 seconds. For the intermediate range of intensities the  $\sigma(hkl)$  was assumed to be a constant percentage of the magnitude of the structure amplitude. For the other ranges the  $\sigma(hkl)$  values reflected the increased error in the measurement of very small and very large reflections.

### 3. Refinement

A set of 1100 signs were obtained from a structure factor calculation based on the E-map coordinates and were used with the observed structure amplitudes to calculate an electron density. Estimates of individual thermal parameters and improved atomic coordinates were obtained from this map and were used as input for the least squares refinement.

Full matrix least squares refinement with unit weights and isotropic thermal parameters was initiated using the computer program ORFLS (32). Three cycles of refinement varying the coordinates of the ten atoms of the two pyrrole rings, the coordinates of the methine bridge and propyl side chain atoms, and all of the isotropic thermal parameters reduced  $R$  to 0.150.

At this stage, the isotropic thermal parameters were converted to the anisotropic form and the weighting scheme discussed above was included. The anisotropic thermal parameters were varied for six atoms per cycle. The first two cycles were with respect to the atoms of one pyrrole ring and one methine carbon atom while the last cycle varied the parameters of the six propyl carbon atoms. Refinement of the thermal parameters, coordinates of all the atoms, and the overall scale

Table III. Weighting Scheme for TPrP Least Squares Refinement.

Intensities in counts/16 seconds.

$$I > 20,000 \quad \omega = \left( \frac{20,000}{I} \right) \times \frac{1}{(0.04 |F_o|)^2}$$

$$500 < I \leq 20,000 \quad \omega = \frac{1}{(0.04 |F_o|)^2}$$

$$15 < I \leq 500 \quad \omega = \frac{1}{(\% \text{ error } \times |F_o|)^2}$$

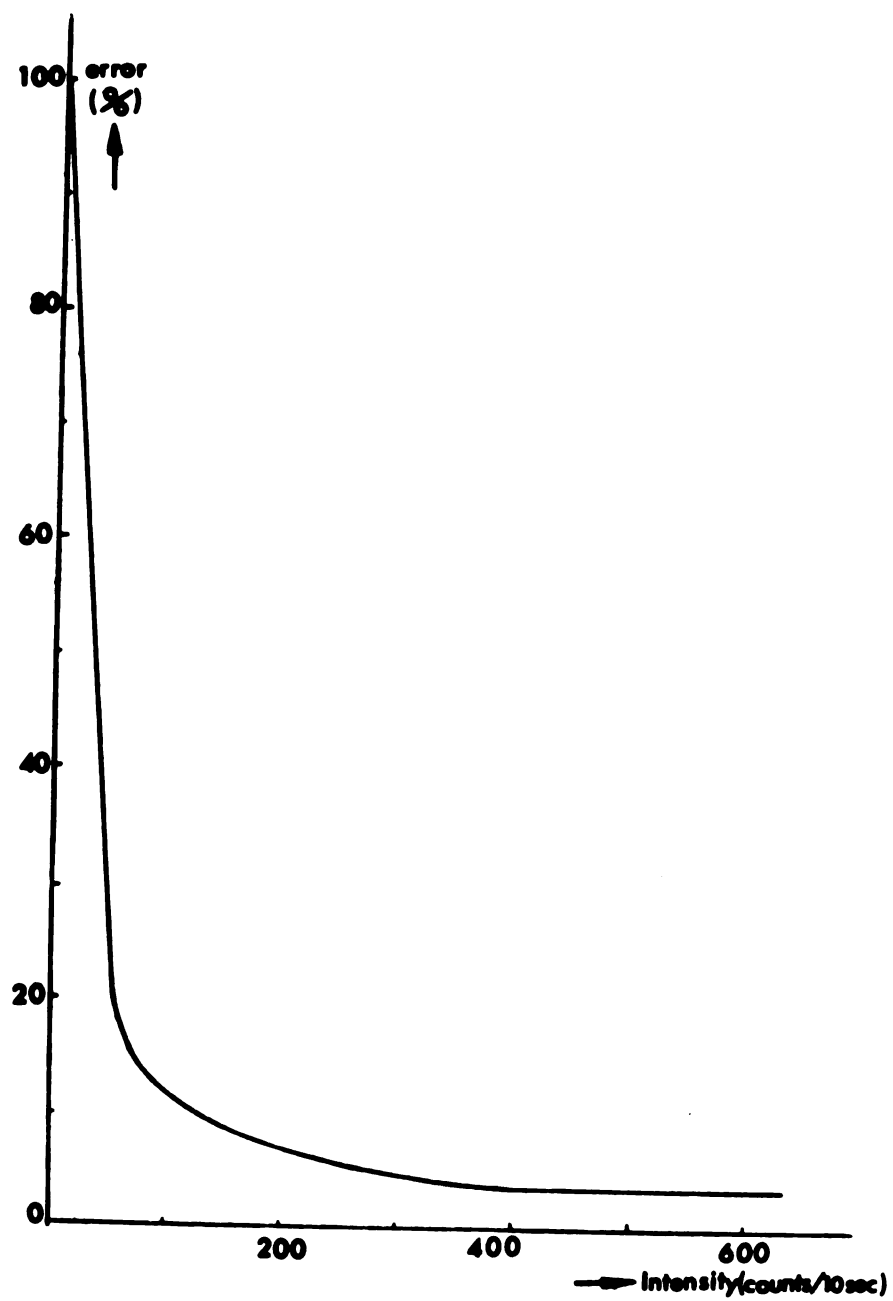


Figure 7. The percent error curve for porphine from Chen (15).

constant one time reduced R to 0.135. The phases obtained from this refinement were used to locate the hydrogen atoms.

#### 4. Location of the Hydrogen Atoms

The difference Fourier synthesis was used to determine the positions of the hydrogen atoms. This difference electron density function for a centrosymmetric crystal is defined as

$$\Delta\rho(x,y,z) = \frac{1}{V} \sum_{hkl} \sum_{l=-\infty}^{+\infty} (|F_o| - |F_c|) S_c \cos 2\pi(hx+ky+lz) \quad (14)$$

where  $S_c$  is the sign of the calculated structure factor. If the signs,  $S_c$ , are nearly correct this function provides a measure of the differences between the model structure used to calculate  $F_c$  and the true structure. The  $\Delta\rho(x,y,z)$  synthesis was used to locate the hydrogen atom positions which should be positive peaks since they were not included in the model.

Three difference Fourier syntheses and structure factor calculations (12 hydrogen atoms were located in the first map and R was reduced to 12.7; the six missing hydrogen atoms of one propyl group were located in the next two calculations, R = 12.2 and 11.7) revealed all of the hydrogen atoms' positions except that of the imino hydrogen atom. The position of the central hydrogen atom was obscured by the presence of residual density in the positions of the nitrogen atoms and by the presence of a relatively large peak at the center of symmetry of the molecule.

This residual peak is in a special position (00,1/2); also only reflections of the type  $k + l = 2n$  contribute to the density at this position. The electron density function for this position is of the form

$$\rho_o(00,1/2) = \sum_{hkl} \sum_{k+l=2n} F(hkl) \cos \pi l. \quad (15)$$

Since the average value of the squared geometric part of the electron density,  $\overline{\cos^2 \theta}$ , is not 1/2 for this special position, the standard error of  $\rho(00, 1/2)$  is larger than the  $\sigma(\rho)$  for a general position by a factor of  $\sqrt{2}$  (33).

The hydrogen atoms were assigned isotropic thermal parameters which were 25% larger than the isotropic B's of the atom to which they were bonded. The positions of the hydrogen atoms were taken from the difference Fourier calculations and were included in the refinement. The refinement was carried out by alternating the carbon and nitrogen atom refinement with cycles refining the coordinates of the hydrogen atoms. After several cycles, R reduced to 0.089. An examination of the values of  $|F_o|$  and  $|F_c|$  indicated that several low order reflections were affected by extinction.

##### 5. Secondary Extinction Correction

The secondary extinction effect is a decrease in the observed intensity of a reflection which is due to the incident beam being partially reflected by outer mosaic blocks before the beam penetrates to the deeper blocks. These blocks receive less power and thus reflect less intensity than they would have otherwise. The secondary extinction effect is largest for high intensity reflections at low values of  $2\theta$ . The secondary extinction coefficient,  $g$ , is defined as  $(I_c - I_o)/2I_c I_o$  and the values of the corrected structure amplitudes are obtained by

$$F_{\text{corr}}^2 = F_o^2 (1 + 2g(Lp)F_c^2) \quad (16)$$

where  $Lp$  is the Lorentz-polarization factor and  $(Lp)F_c^2$  is taken as  $I_c$  (34).

Thirteen low order reflections showed marked discrepancies between  $|F_o|$  and  $|F_c|$  and were corrected for extinction. The value of  $g$  was determined to be  $4.58 \times 10^{-7} e^2$  from the slope of a plot of  $F_c^2/F_o^2$  vs.  $(Lp)F_c^2$ . The corrected values of the extinction affected reflections along with the final calculated values are given in Table IV.

A structure factor calculation including the corrected values for the thirteen extinction affected reflections produced an R-factor of 0.080. A difference Fourier synthesis based on this calculation revealed a single position for the imino hydrogen atom which was included in the refinement in a manner similar to the other hydrogen atoms. The structure was refined as previously described for several cycles and the resulting R-factor was 0.065.

## 6. Evaluation of the Refinement

The calculated bond distances based on the atomic coordinates of the 0.065 structure showed some anomalies: the C-C distances in the propyl groups ranged from 1.492-1.526 Å, considerably shorter than the expected single bond distance of 1.54 Å; the bond distances of the pyrrole rings were within 0.01-0.02 Å of the expected values but the exact nature of the geometry of the rings was not evident. The apparent shortening of the propyl group distances suggested that the introduction of the hydrogen atoms early in the refinement procedure and the subsequent refinement of the hydrogen parameters might have prejudiced the least squares analysis. Such an effect has been observed in the structural analysis of  $\alpha, \beta, \gamma, \delta$ -tetra(4-pyridyl)porphinatomonopyridinezinc(II) (PyZnTPyP) (35). Two schemes of least squares refinement were used for PyZnTPyP; Refinement I excluded hydrogen atoms from the structure, and Refinement II included hydrogen atoms in the structure first in fixed positions and

Table IV. Reflections Corrected for Extinction

\* $|F_c|$  taken from final structure factor calculation (at  $R = 0.054$ )

hkl	$ F_o $	$ F_{corr} $	$ F_c *$
020	284	334	333
01 $\bar{1}$	192	218	215
002	273	382	377
02 $\bar{2}$	180	206	202
01 $\bar{5}$	281	318	330
11 $\bar{1}$	196	209	220
10 $\bar{2}$	215	234	246
112	449	609	599
113	292	323	326
104	423	527	527
114	372	434	437
11 $\bar{4}$	189	198	202
116	312	329	338



then allowed the positions to be refined. A comparison of the two structures obtained by these schemes showed that Refinement II had produced systematic shifts in the positions of the carbon atoms bonded to the hydrogen atoms with the cumulative effect being a "shrinkage" of the rings containing these carbon atoms.

To check the possibility of a similar effect in the TPrP structure, a refinement similar to Refinement I of the PyZnTPyP analysis was carried out on the  $R = 0.065$  structure excluding the propyl hydrogen atoms. After one full cycle of refinement on all of the carbon and nitrogen atom parameters,  $R$  decreased from 0.108 to 0.105. However, the propyl C-C distances increased somewhat and there were slight changes in the pyrrole distances.

Another aspect of the final structure that seemed a little inconsistent was the final  $R$ -factor of 0.065 which seemed large in terms of the reproducibility of the data determined from the redundant reflections ( $R_{OBS} = 0.011$ ). An examination of the weighted differences between  $|F_O|$  and  $|F_C|$  indicated that for the smaller structure amplitudes the estimated values of  $\sigma(|F_O|)$  were too large. A comparison of the weighting scheme with the observed differences,  $\Delta F = ||F_C| - |F_O||$ , was obtained from plots of  $\langle \Delta F \rangle / \langle |F_O| \rangle$  vs.  $\langle |F_O| \rangle$  and  $\sigma(|F_O|)$  vs.  $\langle |F_O| \rangle$  (Figure 8). These plots indicated that the weighting scheme was too severe for the small intensities; therefore, unit weights were used for all further refinement for all reflections.

## 7. The Gaussian Ellipsoid Approximation

The observed electron density based on phases calculated from the structure at  $R = 0.105$  was used to obtain a new set of atomic positions. At this level of refinement the phases should be correct so that an



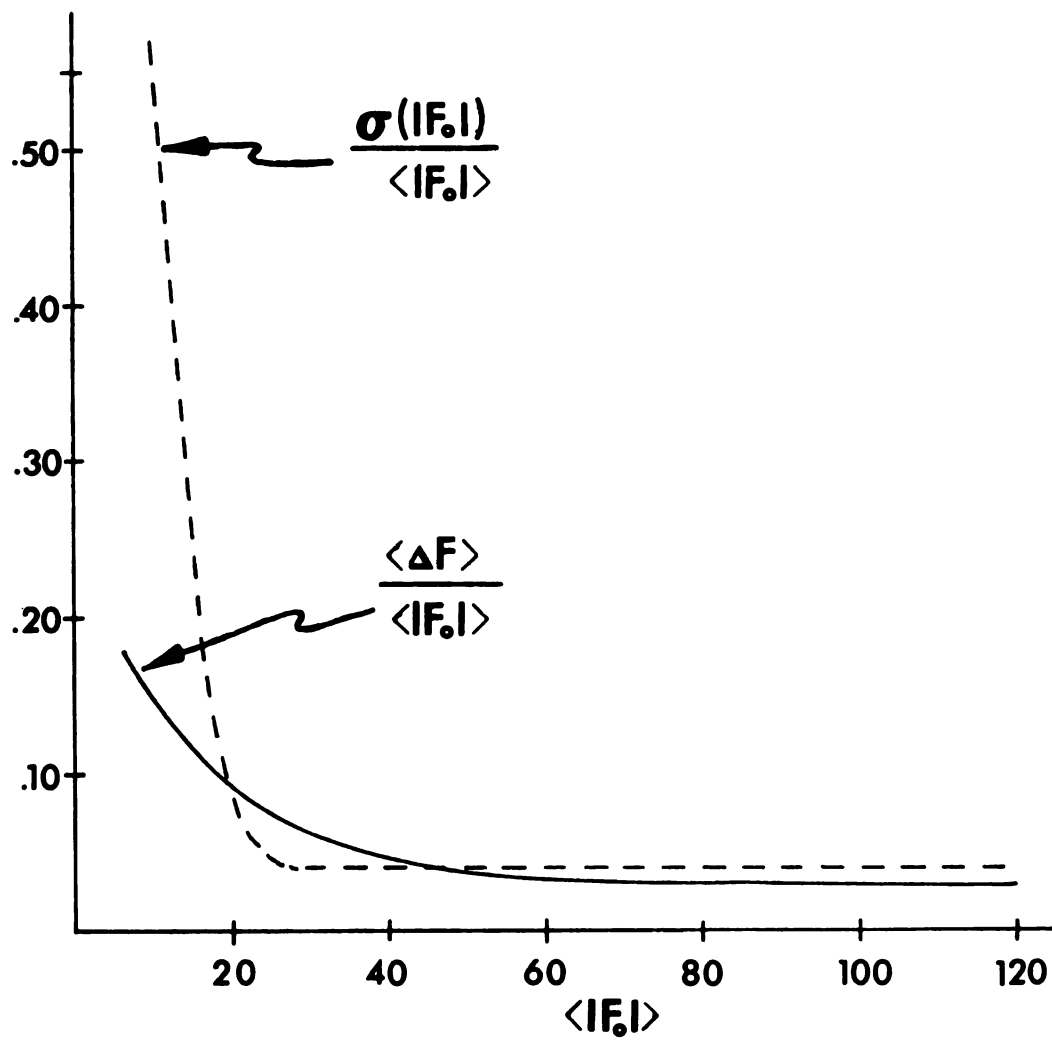


Figure 8. Plots of the observed and calculated weighting schemes.

electron density based on the observed structure amplitudes and these phases should be free of previous complications encountered in the refinement.

The atomic positions for the carbon and nitrogen atoms were obtained by a Gaussian analytical method first employed by Shoemaker, et al., (36). In this method the 27 points of the electron density in a 3x3x3 grid surrounding the peak maximum representing an atomic position are fitted by least squares to the three dimensional function

$$\rho = \exp\left(p - \frac{r}{2}x^2 - \frac{s}{2}y^2 - \frac{t}{2}z^2 + ux + vy + wz + \ell yz + mxz + nxy\right). \quad (17)$$

The values of the parameters  $p$ ,  $r$ ,  $s$ ,  $t$ , etc. for each peak are determined by least squares from the calculated values of the electron density  $\rho_i$ , at the 27 grid points. These parameters are determined by equations of the form:

$$\begin{aligned} p &= \frac{1}{k_p} \sum_{i=1}^{27} C_{i,p} \ell_n \rho_i, \\ r &= \frac{1}{k_r} \sum_{i=1}^{27} C_{i,r} \ell_n \rho_i, \text{ etc.}, \end{aligned} \quad (18)$$

where each parameter is determined by a set of constants and the values of  $\rho_i$ . These constants have been tabulated by Dawson (37) for the case of equally weighted grid points. The location of the peak center is obtained by setting the derivatives of  $\rho$  with respect to  $x$ ,  $y$ , and  $z$  equal to zero which results in the three simultaneous equations,

$$\begin{aligned} rx - \ell y - mx &= u \\ -\ell x + sy - nz &= v \\ -mx - ny + tz &= w, \end{aligned} \quad (19)$$

for the coordinates  $x$ ,  $y$ ,  $z$  of the maximum.

The Gaussian coordinates obtained from the observed electron density were corrected for non-convergence of the Fourier series which arises because the Fourier sum is not infinite. The effect of series termination errors is to enlarge the volume of density due to one atom, causing the atomic densities to overlap, which slightly displaces the point of the maximum. The back shift correction for this effect is obtained from the difference between the atomic coordinates used for a structure factor calculation and the coordinates obtained from the resulting calculated electron density. This difference between the input coordinates and the calculated map coordinates is subtracted from the coordinates obtained from the observed electron density.

The Gaussian method was used to obtain the atomic coordinates from the calculated electron density for the  $R = 0.105$  structure of TPrP. These coordinates were then used to make the back shift corrections. The coordinates of the hydrogen atoms were obtained by the Gaussian method from the corresponding difference Fourier method based on a structure factor calculation with all hydrogen atoms excluded.

The structure obtained from the electron densities was refined for three cycles, one on all the parameters of the carbon and nitrogen atoms, a second on the coordinates of the hydrogen atoms, and a third on the positional parameters of the carbon and nitrogen atoms. This refinement reduced the R-factor from 0.062 to 0.054. The change in R-factor between the last two cycles was only 0.001 and the shifts in the parameters were insignificant with respect to the standard deviations of the coordinates; therefore, the refinement was terminated at this point.

## V. Results

The final coordinates, anisotropic temperature factors and peak heights for all the carbon and nitrogen atoms are listed in Table V with the atoms labeled according to the scheme shown in Figure 9. The estimated standard deviations of the atomic parameters are given in the last row of Table V. The mean square displacements along the principal axes (ORTEP (38)) are given in Table VI for the carbon and nitrogen atoms. The final coordinates, isotropic thermal parameters, and peak heights for the hydrogen atoms are listed in Table VII. The hydrogen atoms have the label of the carbon or nitrogen atoms to which they are bonded.

The atomic positions of the nucleus of the TPrP molecule were fitted to least squares plane according to the method described by Shoemaker, et al., (39). The plane is described by the equation  $m_1x + m_2y + m_3z = d$ , where  $d$  is the origin-to-plane distance and the best values of  $m_1$  are obtained by minimizing the sum of the squares of the residuals,  $(m_1x_1 - d)^2$ . The equation obtained for the least squares plane (NLS) of the 24 atoms of the TPrP nucleus is  $3.37x + 2.02y + 13.03z = 9.89$  Å and the deviations of the atoms from this plane are given in Figure 9. The standard deviation of this plane is 0.015 Å. The individual pyrrole rings were also fit to least squares planes and the atomic deviations from these planes are given in Table VIII.

The interatomic distances and bond angles for TPrP are shown in Figure 10. The standard deviations in bond lengths vary from

Table V. Final Atomic Parameters, Carbon and Nitrogen\*

Atom	x	y	z	$\beta_{11}$	$\beta_{22}$	$\beta_{33}$	$\beta_{12}$	$\beta_{13}$	$\beta_{23}$	Peak height ( $\text{\AA}^{-3}$ )
NA	0.7655	-0.0206	0.5669	0.0545	0.0098	0.0022	-0.0004	0.0026	-0.00010	7.8
CA1	0.7425	-0.1215	0.5982	0.0496	0.0101	0.0020	-0.0029	0.0021	0.00005	6.2
CA2	0.5571	-0.0980	0.6388	0.0613	0.0113	0.0024	-0.0033	0.0034	0.00004	5.4
CA3	0.4728	0.0124	0.6308	0.0545	0.0122	0.0025	-0.0013	0.0034	-0.00046	5.8
CA4	0.6048	0.0628	0.5852	0.0439	0.0103	0.0021	0.0001	0.0018	-0.00041	6.4
CM5	0.5785	0.1767	0.5649	0.0436	0.0104	0.0021	0.0012	0.0011	-0.00047	6.0
NB	0.8820	0.1683	0.4892	0.0473	0.0093	0.0020	-0.0003	0.0016	0.00004	7.3
CB1	0.7127	0.2248	0.5211	0.0489	0.0090	0.0022	0.0017	0.0004	0.00002	5.8
CB2	0.6944	0.3443	0.5033	0.0666	0.0107	0.0027	0.0051	0.0024	0.00016	5.3
CB3	0.8479	0.3587	0.4605	0.0673	0.0093	0.0028	0.0025	0.0021	0.00052	5.0
CB4	0.9637	0.2478	0.4509	0.0491	0.0093	0.0022	0.0000	0.0007	0.00003	5.6
CM6	1.1338	0.2264	0.4090	0.0521	0.0092	0.0022	-0.0014	0.0014	0.00026	5.4
CP51	0.3943	0.2539	0.5941	0.0474	0.0114	0.0025	0.0036	0.0013	-0.00080	5.4
CP52	0.5346	0.3133	0.6513	0.0559	0.0123	0.0028	0.0024	0.0017	-0.00134	5.3
CP53	0.3437	0.3870	0.6803	0.0741	0.0143	0.0036	0.0056	0.0036	-0.00160	4.7
CP61	1.1840	0.3223	0.3657	0.0627	0.0096	0.0026	-0.0029	0.0024	0.00059	5.5
CP62	0.9992	0.3151	0.3047	0.0623	0.0114	0.0025	-0.0020	0.0023	0.00086	5.2
CP63	1.0848	0.3988	0.2586	0.0745	0.0133	0.0029	-0.0019	0.0021	0.00191	4.6
$\sigma \times 10^5$	46-77	21-35	10-17	137-265	27-52	6-12	51-95	24-44	11-20	

\*Anisotropic temperature factor =  $\exp[-(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)]$

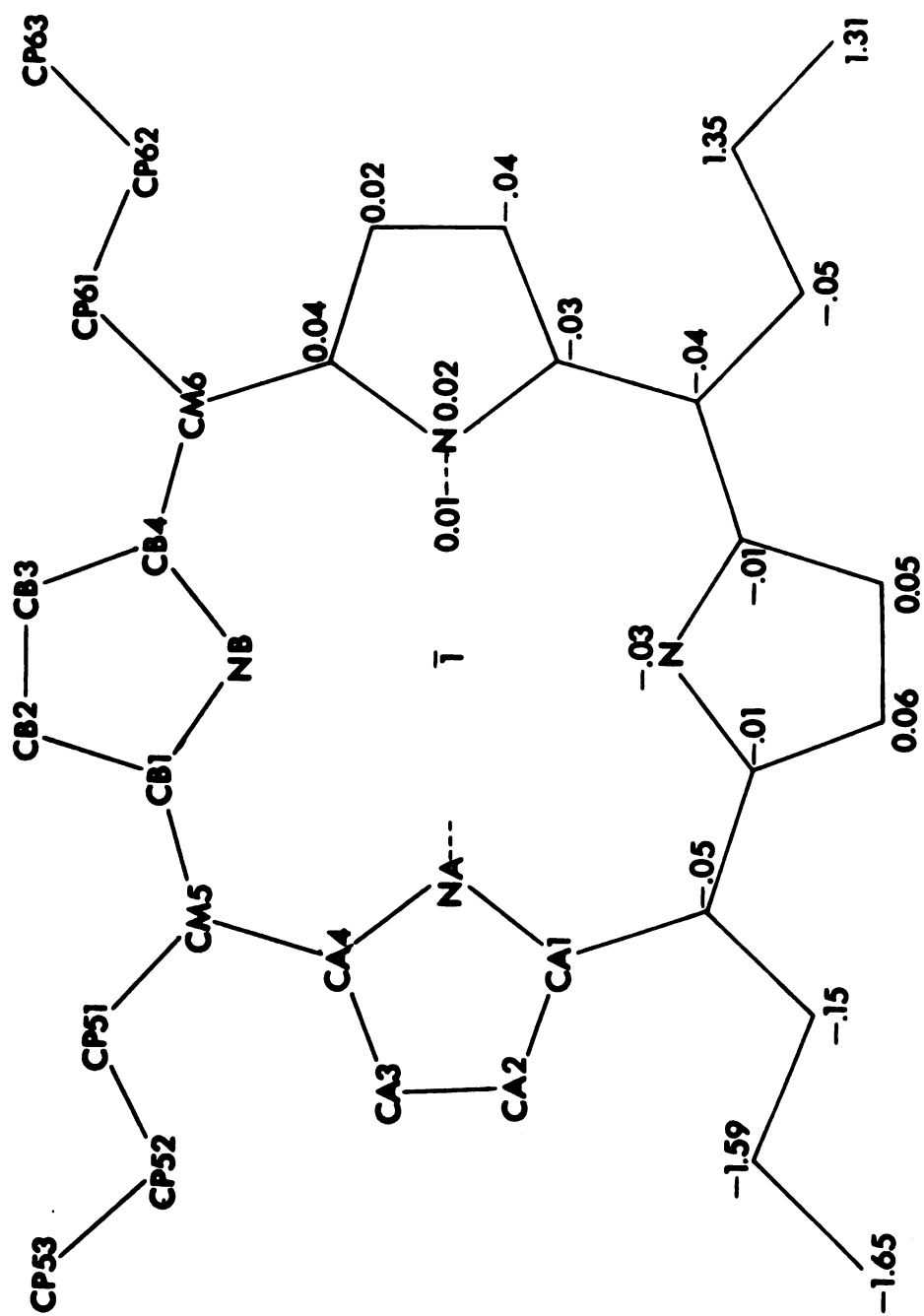


Figure 9. Labeling scheme for TPrP and the deviations from the NLS plane (in Å).



Table VI. The Mean Square Displacements (in  $\text{\AA}^2$ ) for the Carbon and Nitrogen Atoms.

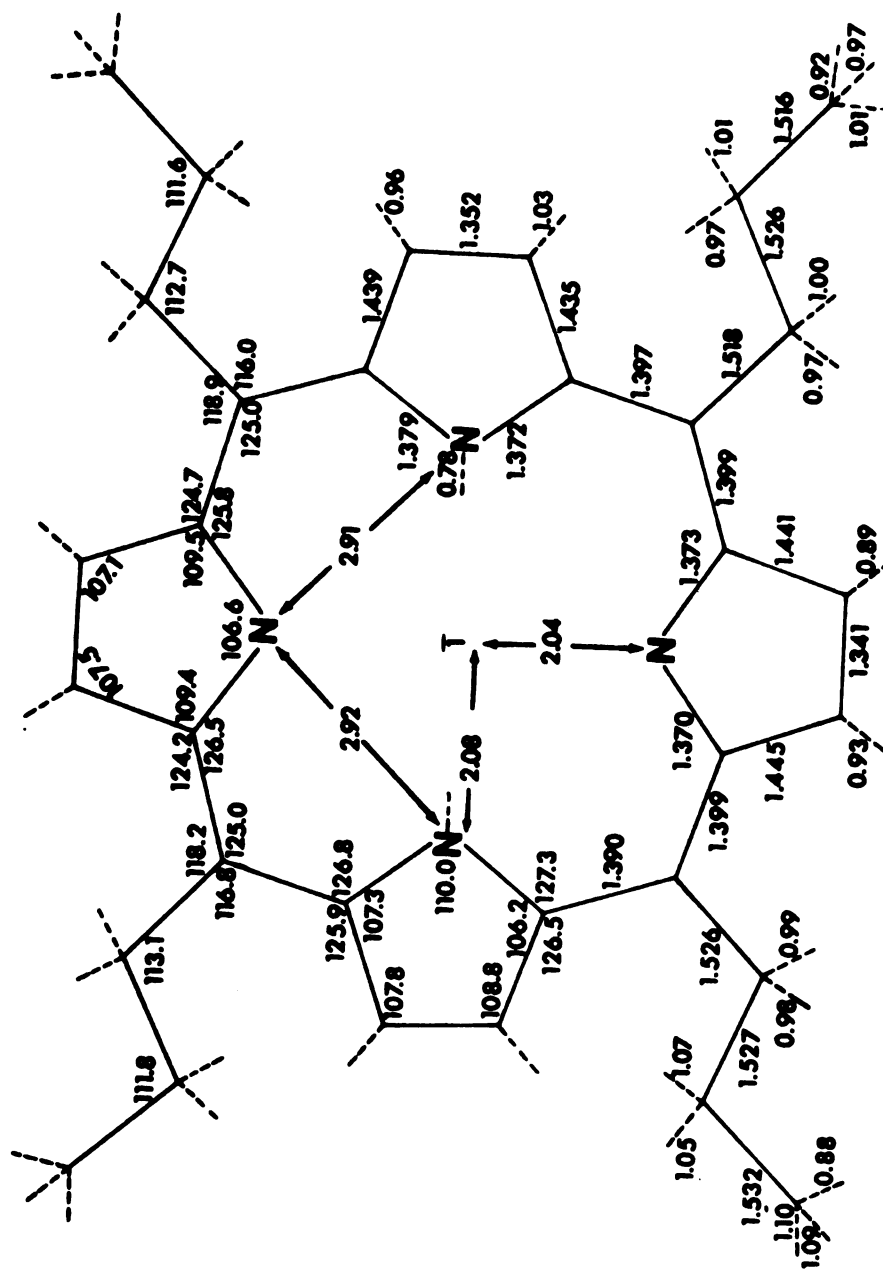
	$8\pi^2 \overline{u_1^2}$	$8\pi^2 \overline{u_2^2}$	$8\pi^2 \overline{u_3^2}$
NA	4.13	5.19	5.40
CA1	3.88	4.55	5.91
CA2	4.37	5.52	6.92
CA3	4.23	5.59	6.81
CA4	4.01	4.42	5.65
CM5	3.92	4.33	5.84
NB	3.90	4.73	5.01
CB1	3.92	4.68	5.53
CB2	4.95	5.23	7.48
CB3	4.59	5.79	6.94
CB4	4.12	5.01	5.26
CM6	4.19	4.75	5.57
CP51	4.25	4.67	6.99
CP52	4.52	5.49	7.83
CP53	4.89	7.62	9.55
CP61	4.23	5.55	6.65
CP62	4.24	5.99	6.92
CP63	4.26	7.26	8.83

Table VII. Final Hydrogen Atom Parameters

Atom	x	y	z	Isotropic Peak Height	
				B	eA <sup>-3</sup>
HNA	0.8548	-0.0146	0.5420	4.5	0.4
HA2	0.5026	-0.1533	0.6668	5.6	0.4
HA3	0.3390	0.0577	0.6515	5.5	0.5
HB2	0.5911	0.3906	0.5209	5.8	0.5
HB3	0.8845	0.4274	0.4423	5.6	0.5
H51A	0.3167	0.3136	0.5643	5.3	0.4
H51B	0.2438	0.2094	0.6034	5.3	0.4
H52A	0.6767	0.3593	0.6398	6.3	0.4
H52B	0.6005	0.2476	0.6796	6.3	0.4
H53A	0.4174	0.4276	0.7169	7.1	0.4
H53B	0.2552	0.4538	0.6534	7.1	0.3
H53C	0.1838	0.3541	0.6855	7.1	0.4
H61A	1.3736	0.3186	0.3598	5.6	0.4
H61B	1.1596	0.3968	0.3847	5.6	0.5
H62A	1.0051	0.2279	0.2888	5.7	0.4
H62B	0.8055	0.3353	0.3116	5.7	0.4
H63A	1.2899	0.3845	0.2524	6.5	0.4
H63B	0.9732	0.4016	0.2244	6.5	0.4
H63C	1.0901	0.4888	0.2735	6.5	0.4
$\sigma \times 10^4$	64-78	29-35	14-18		

Table VIII. Atomic Deviations from the Least Squares Plane  
of Individual Pyrrole Rings.

<u>Pyrrole A</u>		<u>Pyrrole B</u>	
Na	+ 0.002	NB	+ 0.010
CA1	- 0.004	CB1	- 0.007
CA2	+ 0.004	CB2	+ 0.002
CA3	- 0.002	CB3	+ 0.004
CA4	+ 0.000	CB4	- 0.008
<hr/>		<hr/>	
$\sigma = \pm 0.003 \text{ \AA}$		$\sigma = \pm 0.007 \text{ \AA}$	



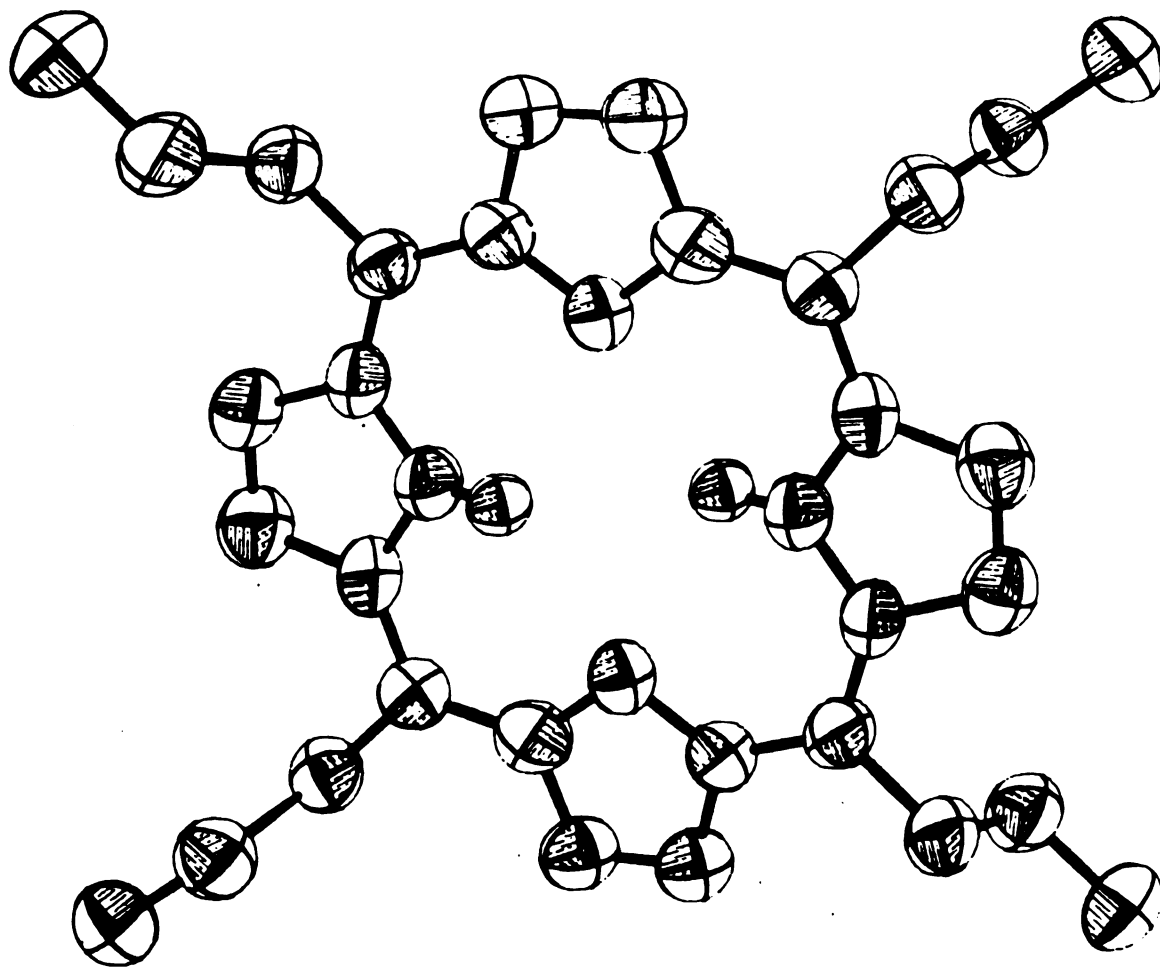


Figure 11. Computer plot of the TPrP molecule.

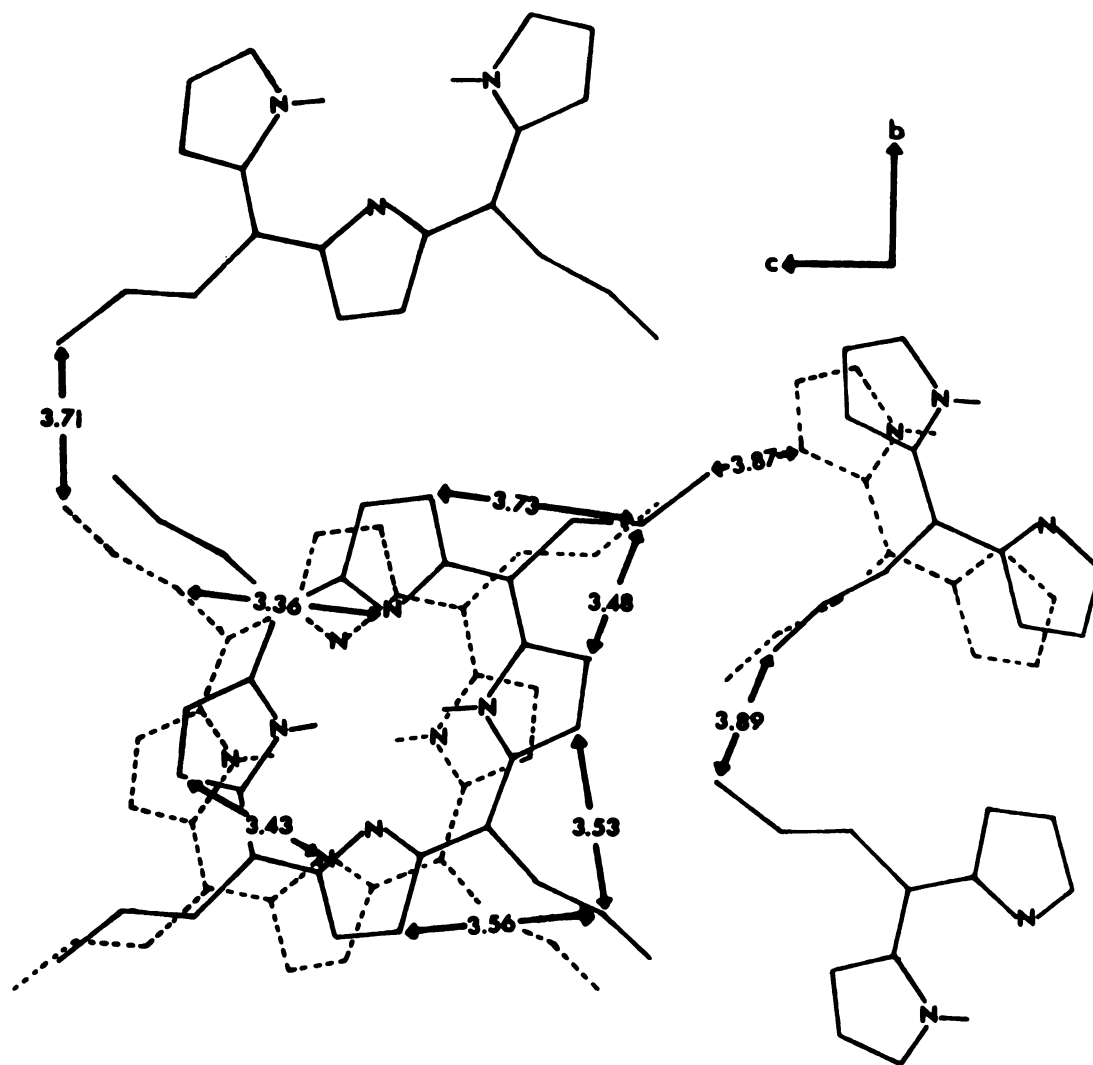


Figure 13. The packing along  $a^*$ .

$\pm 0.003$  Å for the interior of the molecule to  $\pm 0.005$  Å for the propyl group distances. The standard deviations in the bond angles range from  $\pm 0.2$ – $0.4^\circ$ .

A computer drawing in perspective of the TPrP molecule indicating the relative thermal ellipsoids is given in Figure 11 (ORTEP (38)). A composite final electron density is shown in Figure 12 projected onto the bc plane. In addition, Figure 12 shows the composite difference electron density obtained from a structure factor calculation excluding hydrogen atoms. The contours for the carbon and nitrogen atoms are drawn at  $1.0 \text{ e } \text{\AA}^{-3}$  levels. The hydrogen difference densities are drawn at  $0.1 \text{ e } \text{\AA}^{-3}$  contours, beginning with  $0.1 \text{ e } \text{\AA}^{-3}$ .

A packing diagram which indicates the arrangement of the molecules in the unit cell is shown in Figure 13. The packing is viewed in the a\* direction and the molecules on the right are related by the c glide plane. The molecules represented by dashed lines are translated by -a .

## VI. Discussion

The short value for the average CM-CP1 distance of 1.522 Å, as compared to the C-C single bond distance of 1.54 Å (18), could result from the  $sp^2$  hybridization of the bridge carbon atom. It is unlikely that the aromatic character of the macrocycle would affect the second and third bonds on the side chains so these short distances are probably due to some motion of the side chains.

The average value of the C-H distances is  $1.05 \pm 0.08$  Å which corresponds closely to the tabulated value of 1.08 Å for aromatic compounds (18). However, the N-H distance is approximately 0.2 Å shorter than the corresponding value of 0.99 Å expected for a N-H bond adjacent to a C-N linkage. This short distance may be due to a combination of the effect of the double bond character in the C-N bond and the close proximity of the hydrogen atoms in the interior of the ring. The latter might be especially important since the inner hydrogens show no deviation from the overall planarity of the macrocyclic ring.

The average value of the  $C_\alpha$ - $C_m$  distance, where  $C_\alpha$  is the pyrrole atom bonded to nitrogen and  $C_m$  is the bridge atom, is  $1.397 \pm 0.007$  Å which compares favorably with the aromatic C-C distance of  $1.395 \pm 0.003$  Å. In contrast, the average distances within the two pyrrole rings do not agree with the tabulated aromatic distances, e.g. C-N for an aromatic heterocyclic system is  $1.352 \pm 0.005$  Å and C-N in



TPrP is  $1.374 \pm 0.007$  Å. These differences arise because the aromatic pathway passes through either the inside C-N bonds or the outside C-C bonds in any resonance structure while the bridge carbon pathway is always aromatic (see Figure 14(c)). Thus, the bridge carbon atoms should be the most sensitive to changes in the  $\pi$  electron current.

The TPrP structure shows systematic differences in the distances and angles of the two crystallographically independent pyrrole rings. These differences can be interpreted in terms of a hybrid of the two major resonance forms of the porphyrinic nucleus (shown in Figure 14(a) and (b)). The average distances and angles for the two independent pyrrole rings are given in Figure 14(c). The largest differences between the two pyrrole rings are 0.011 Å in the  $C_{\beta}-C_{\beta}$  bonds and  $3.4^\circ$  in the  $C_{\alpha}-N-C_{\alpha}'$  angles. The  $C_{\beta}-C_{\beta}$  distance in pyrrole ring B of 1.341 Å is very similar to the tabulated double bond distance of  $1.337 \pm 0.06$  Å as would be expected from the hybrid structure. The major difference in angles at the nitrogen atom is accompanied by similar changes in the  $N-C_{\alpha}-C_{\beta}$  angle. The larger angle for the imino pyrrole ring (pyrrole A) may be due to the repulsion between the two hydrogen atoms which would tend to force these rings outward to reduce the H-H interaction. The center to NA distance is 0.04 Å larger than the corresponding distance for NB (see Figure 9).

Indications of the same free base macrocyclic structure are found when the results of this analysis are compared with those of porphine (15) and TPP (13). These three studies suggest the same hybrid structure discussed above. Since these structure determinations were all carried out to the same level of refinement,  $R = 0.059, 0.056, 0.054$  for porphine, TPP, and TPrP, respectively, their results were averaged to give an average general structure of the macrocycle. The three structures give four examples of each type of pyrrole ring whose parameters were averaged to give the distances

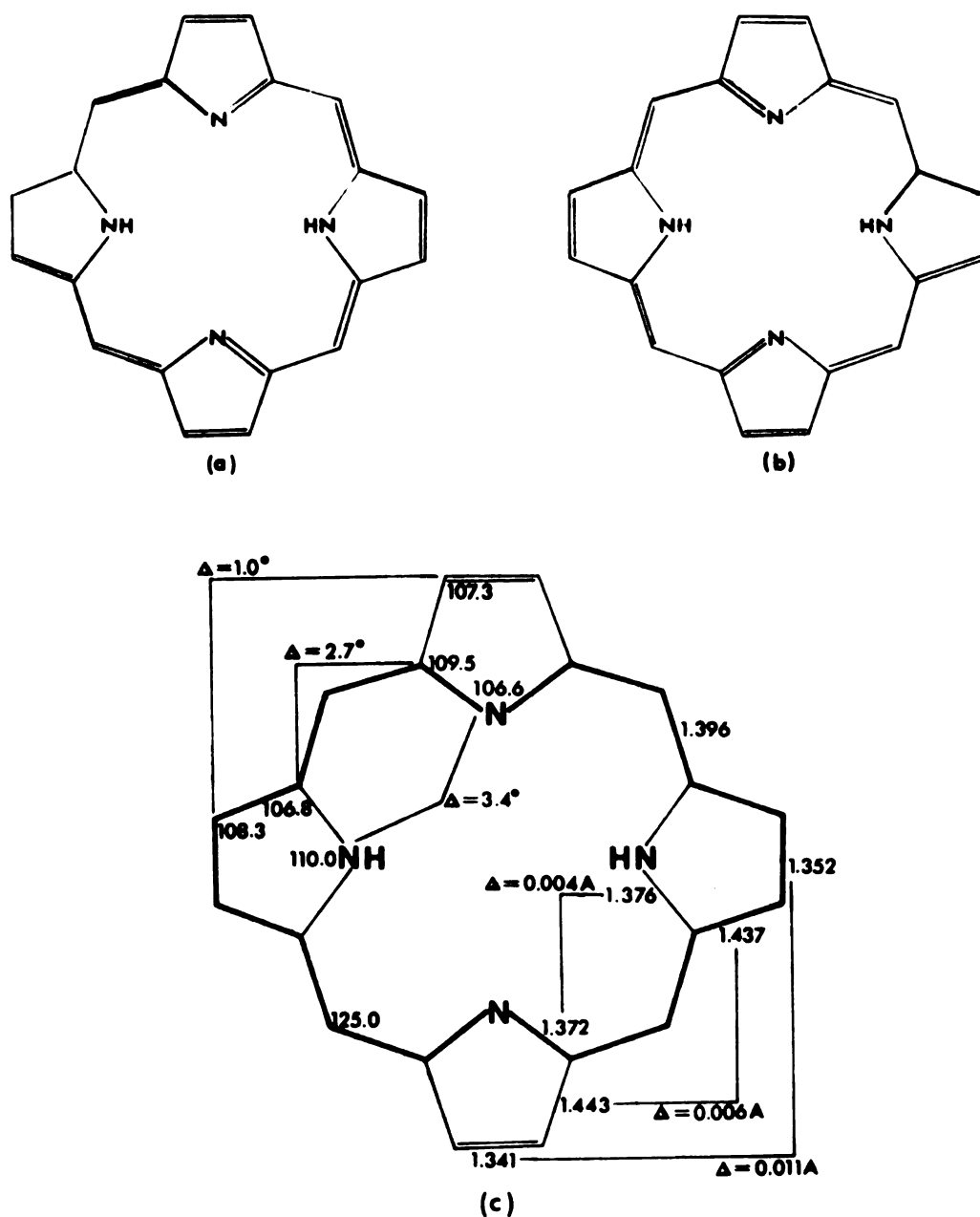


Figure 14. Dominant resonance structures of the porphine macrocycle (a,b). Average observed distances and angles of the independent pyrroles of TPrP(c). Differences ( $\Delta$ ) between corresponding bonds and angles also shown; expected nature of hybrid of dominant resonance structures shown with heavy line and double bond lines.

and angles shown in Figure 15. The standard deviations from the mean fluctuate from 0.003-0.010 Å for the bond lengths and from 0.3-0.7° for the bond angles and all of the observed values are within  $2\sigma$  of the mean value. A scale representation (2 inches/Å) of the two types of pyrrole rings is shown in Figure 16 which is intended to indicate the structural differences between the two rings. The imino pyrrole is wider than the aza pyrrole and is closer to a regular pentagon. The distance across pyrrole A is 2.25 Å and for pyrrole B, CB1 to CB4 is 2.20 Å.

The largest deviations from the average free base structure are associated with the methine carbon atoms where the average  $C_{\alpha}-C_m$  distance in the two substituted compounds is 0.016 Å longer than the mean distance in porphine and the  $C_{\alpha}-C_m-C_{\alpha}'$  angle is 1.8° smaller for the substituted porphyrins. The overall effect of these differences is to enlarge the interior of the ring by increasing the center to bridge carbon atom distance by about 0.02 Å in the substituted compounds. These small changes suggest that substitution at the bridge carbon atom produces some small localized changes in the  $\pi$  electron density which affect the sensitive methine atoms but that the basic macrocyclic structure remains essentially the same.

Another effect of the enlargement of the interior of the ring is to increase the separation of the imino nitrogens in the substituted compounds. The N-N distances are 4.11, 4.15, 4.20 Å for porphine, TPrP, and TPP, respectively. The 0.04 Å difference between the porphine and TPrP structure probably represents the effect of substitution because the TPP structure is nonplanar with the imino nitrogen atoms above and below the plane thus increasing their separation. The aza nitrogen separation is nearly the same in TPrP and porphine.

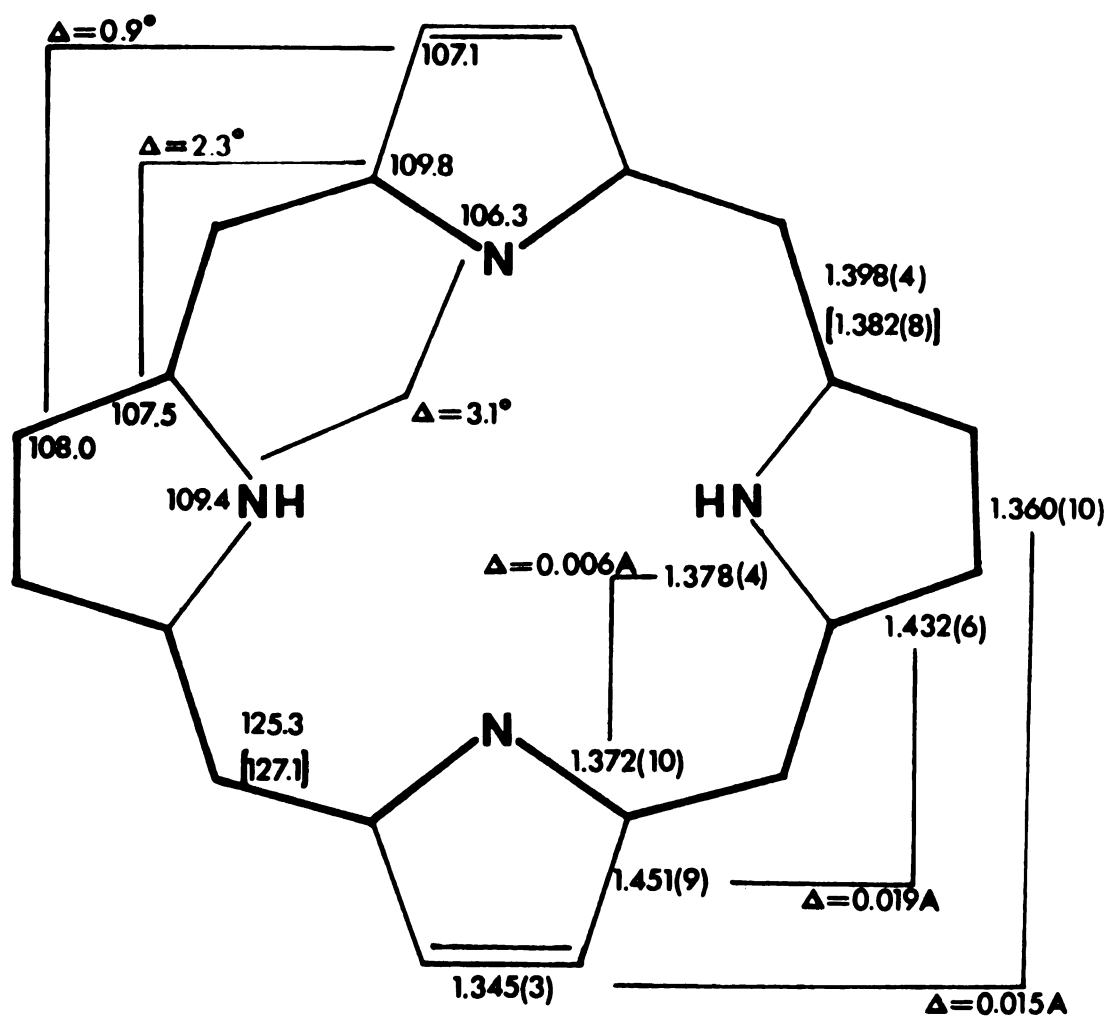


Figure 15. Average structure for free base porphine. Average obtained from TPP, TPrP, and porphine. Standard deviations of distances in parentheses, standard deviations of angles  $0.3\text{--}0.7^\circ$ . The average bridge distance and angle of porphine in brackets.

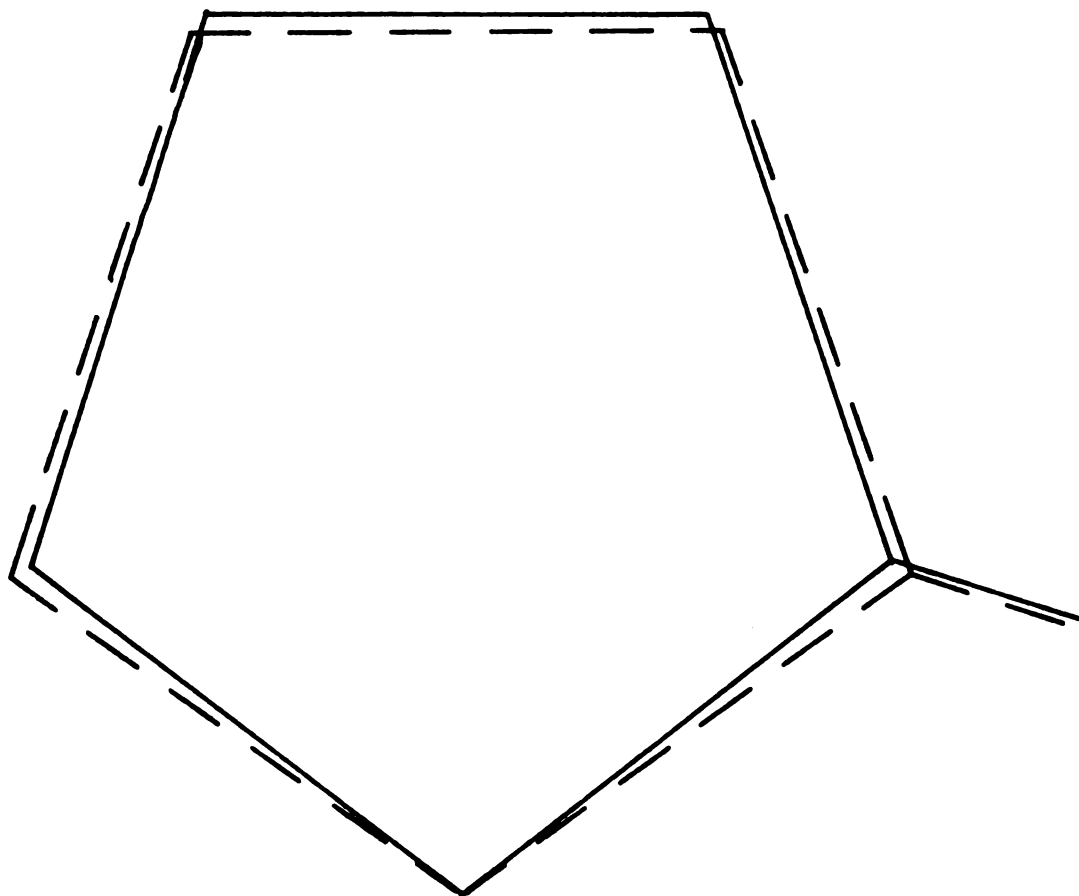


Figure 16. Scale representation of the two types of pyrrole rings.

2 inches/A.

Even though these three compounds have nearly the same structure for the free base macrocycle, some small but significant differences between the average TPrP structure (Figure 14) and the average structure of porphine and TPP are evident when the two types of pyrrole rings are compared. This is consistent with the changes in the visible absorption spectrum for the TPrP molecule (Figure 2). The aliphatic substitution appears to have reduced the differences in the bond distances between the two different pyrroles of the macrocyclic structure as indicated in Figure 17. The  $C_{\alpha}-C_{\beta}$  distances show the largest change in the difference which is a decrease of 0.017 Å and the difference between the  $C_{\beta}-C_{\beta}$  distances decreases by 0.006 Å. The locality of the former change is nearer the substitution than the latter which suggests that the effect decreases with distance from the side chain position. The C-N distances remain essentially unchanged and this has been observed in all three structures. The effect of reducing the asymmetry of the molecule can explain the unusual visible absorption spectra of TPrP. According to Platt (40), bands I and III (see Figure 2) are forbidden (0-0) vibrational bands and are therefore sensitive to any loss of symmetry in the porphyrin molecule. Thus, the relative increase in the intensities of bands I and III of the TPrP spectrum could arise from the increased symmetry of the macrocycle. In fact, the TPrP spectrum seems to be the only example where band I is more intense than band II.

The effect on the bond angles is different from that on the distances. The TPrP bond angles in comparison with those of the average of TPP and porphine suggest that the aza pyrrole (B) remains the same while the angles in pyrrole A change such that the differences between A and B are increased, especially for the  $C_{\alpha}-N-C_{\alpha}$  angle. The full significance of this change is not clear at present. However, the

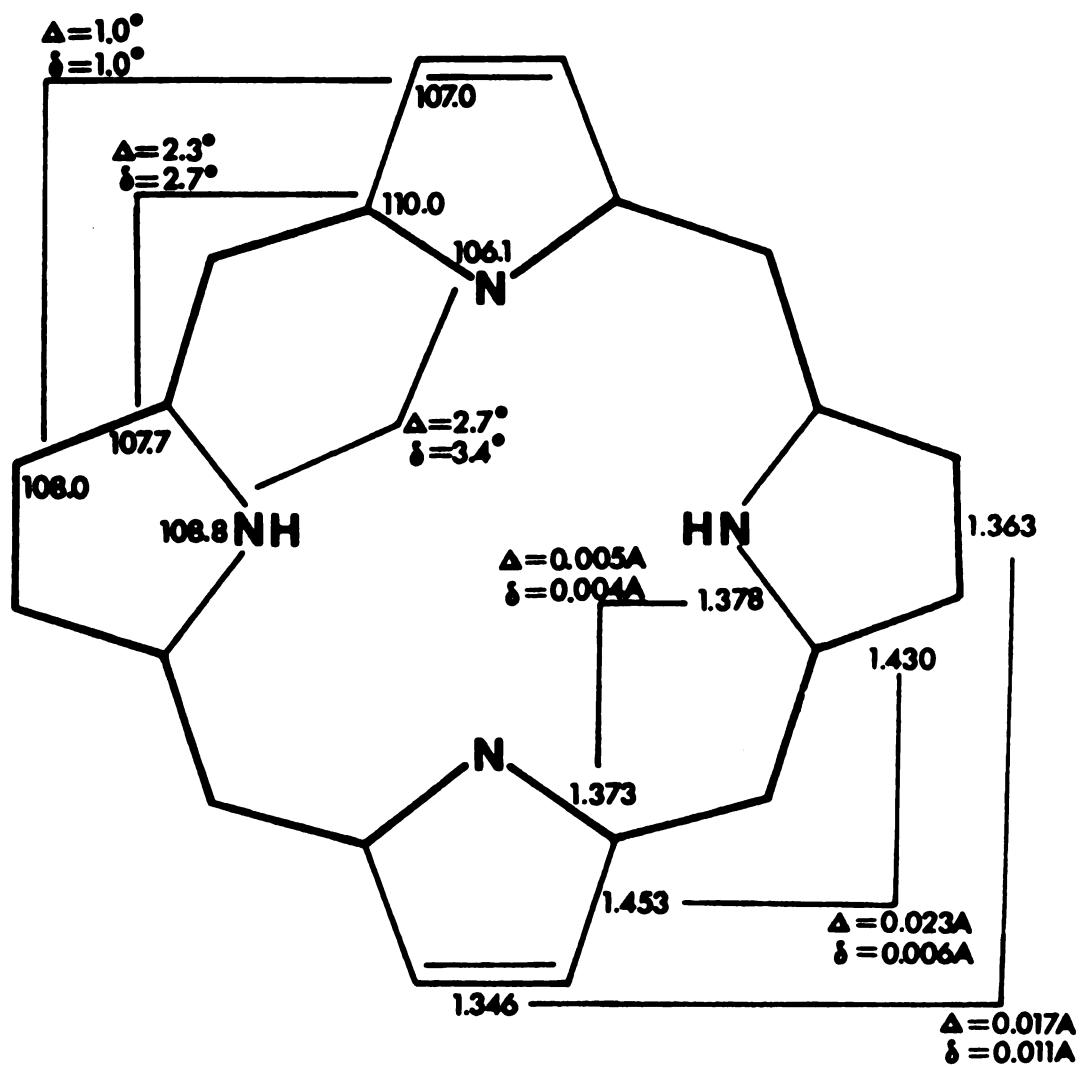


Figure 17. Average structure for porphine and TPP with the differences ( $\Delta$ ) between the pyrrole rings.  $\delta$  represents the corresponding differences for TPrP.

overall behavior which is observed is suggestive of  $\pi$  electron delocalization onto the side chains.

The individual pyrrole rings are planar to within  $\pm 0.007$  Å and although the TPrP nucleus shows deviations up to  $\pm 0.04$  Å it is essentially planar. The pyrrole rings form small but definite angles of approximately  $2^\circ$  to the NLS plane. The imino pyrrole ring (A) is tilted about the line from NA to the center of the CA2-CA3 bond while pyrrole ring B is tilted about the line joining CB1 and CB4. The small tilts are probably related to the packing arrangement. The bonding about the  $C_\alpha$  pyrrole atoms and the bridge carbon atoms is planar since the sum of the bond angles is  $360^\circ$  and is indicative of  $sp^2$  hybridization.

The molecules pack in layers along the  $a^*$  axis as shown in Figure 13. If the arrangement is viewed perpendicular to the molecular planes, one molecule is rotated about the normal to the NLS plane by approximately  $45^\circ$  with respect to the molecule below. This rotation in the stacking arrangement probably results in more space for the side chains; a similar but larger rotation of approximately  $90^\circ$  was found in the packing of TPP (3). The stacking is compact, producing a short NB-CP51 intermolecular distance of 3.36 Å which apparently causes the small tilt observed with pyrrole B. The intermolecular distance between NB and CA2 is also short (3.43 Å) and could effect contributions to the tilts of both pyrrole rings. The propyl side chains are in an extended configuration efficiently filling the space between the molecules related by the glide plane. One propyl group is rotated about the  $C_m$ -CP1 bond producing a short intramolecular contact of 3.48 Å to the adjacent  $\beta$  carbon atom of the pyrrole. This configuration probably arises from a close contact between the methyl carbon atoms, CP53-CP63,



of the molecules related by the screw axis along b. This separation of 3.71 Å is short in comparison to the sum of the van der Waals radii for two methyl groups of about 4.0 Å. The methyl group on the twisted propyl side chain is also fairly close (3.87 Å) to the  $\beta$  pyrrole carbon atom of the adjacent molecule along c. This twist of the propyl side chain is evident in both the composite electron density (Figure 11) and the computer drawing of the molecule (Figure 10).

The presence of the imino hydrogen atom at a peak height of  $0.4 \text{ e } \text{\AA}^{-3}$  bonded to the nitrogen atoms of opposite pyrroles presents a third example, along with porphine (2) and TPP (3), of localized hydrogen atoms. Along with the hydrogen atoms, the pyrrole rings are observed to be structurally different which indicates that the ring current follows different paths through the two rings. These results suggest that tautomerism of the N-H protons and the resulting magnetic equivalence of the pyrrole rings does not occur in the solid state. Thus, the example of a free base crystal structure with equivalent pyrrole rings and four half-hydrogen atoms (14) is the result of packing disorders.

The possibility of N-H  $\cdots$  N hydrogen bonding in the porphyrins is precluded by the consistently short N-H bond lengths observed in the free base. This distance for TPrP is the shortest, 0.78 Å, while porphine and TPP have similar distances of 0.86 and 0.96 Å, respectively. For hydrogen bonding to occur the N-H distance should be longer than the single bond distance. The average value of the N-H distance in hydrogen bonded systems has been determined from neutron diffraction studies to be 1.04 Å (41) which is longer than the observed N-H lengths in the free bases. Probably the porphyrin ring is too large for hydrogen

bonding to occur since the average N-N distance is 2.90 Å rather than the value for phthalocyanines of 2.65 Å. The lower energy for the N-H stretch which was observed in the porphyrins (9) may be due to complicated effects from the aromatic nature of the macrocycle.

Finally, in contrast to the other two structures, in the TPrP structure the imino hydrogen atom is within  $\pm 0.01$  Å of the NLS plane and the two CNH angles are essentially equal. However, the shorter N-H bond gives the longest H-H distance across the ring: 2.60 Å in TPrP, 2.41 Å in porphine and 2.36 Å in TPP.

This analysis has revealed a common structure for the free base macrocycle which undergoes small changes upon substitution. The aliphatic substituents appear to interact with the  $\pi$  electrons of the aromatic nucleus although the basic structural features of localized imino hydrogen atoms bonded to opposite pyrrole rings and structural differences between these rings and the other two pyrrole rings remain the same.

## PART II

N-FORMYL-L-PHENYLALANINE:  $\alpha$ -CHYMOTRYPSIN COMPLEX, THE EFFECT OF  
BINDING OF A COMPETITIVE INHIBITOR SUBSTRATE-LIKE MOLECULE

## VII. Introduction

### 1. Chymotrypsin

The proteolytic enzyme chymotrypsin catalyzes the hydrolysis of proteins. Chymotrypsin (CHT) is obtained from its inactive precursor chymotrypsinogen by tryptic cleavage of a specific peptide bond between arginine 15 and isoleucine 16. (A table of amino acids and the sequence of CHT are given in Appendix III). This cleavage yields a fully active enzyme which can then undergo further tryptic or chymotryptic hydrolysis liberating the dipeptides Ser 14-Arg 15 and Thr 147-Asn 148 to give  $\alpha$ -CHT.  $\alpha$ -CHT,  $C_{1113}S_{12}O_{349}N_{300}H_{1752}$ , has a molecular weight of approximately 25,300 amu. The enzyme is composed of 241 amino acids in three peptide chains: the A chain with 13 residues, the B chain with 131 residues, and the C chain with 97 residues.

The hydrolysis of peptide bonds occurs faster, by a factor of  $10^5$ , in the presence of proteolytic enzymes such as pepsin, trypsin, and chymotrypsin. CHT is specific for the hydrolysis of peptide bonds which are adjacent to the carboxyl function of the aromatic amino acids, phenylalanine, tyrosine, and tryptophan. CHT also can catalyze the hydrolysis of peptide bonds adjacent to large hydrophobic residues such as leucine (42) and methionine (43). In addition to peptide bonds, CHT catalyzes the hydrolysis of esters and amides of aromatic amino acids, e.g. N-acetyl-L-tyrosine ethyl ester or N-acetyl-L-tyrosine amide, and some other hydrophobic compounds.

The enzyme is specific for the L isomers of the substrate molecules which bind reversibly to the enzyme and are hydrolyzed. The D isomers of substrate molecules are competitive inhibitors in that they compete with the substrate for binding positions in the active center and their reversible binding inactivates the enzyme.

Niemann and Cohen and their co-workers (44,45) developed the concept of a tetrahedral arrangement of sites on the enzyme which would complement the arrangement of substituents around the  $\alpha$ -carbon of the substrate. These four sites on the enzyme would accommodate the aryl group, the acylamido group, the bond to be hydrolyzed and the hydrogen atom. In this concept, the structural arrangement of amino acids near the active site would determine the specificity of the enzyme.

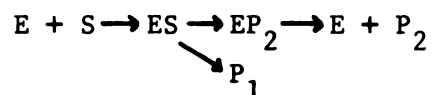
A serine residue and a histidine residue have been implicated in the catalytic reaction of  $\alpha$ -CHT. The importance of the serine residue was indicated by the reaction of diisopropylphosphofluoridate with CHT which inactivates the enzyme and phosphorylates a single serine residue. The result of this reaction suggests that this residue, Ser 195, plays a key role in the catalytic mechanism. Studies of the catalysis of p-nitrophenyl-acetate (46) indicated that the reaction proceeds through an acyl enzyme intermediate (the serine residue is acylated), so that normal peptide catalysis might also proceed through such an acyl enzyme intermediate.

A histidine residue was first implicated by the pH dependence of the catalytic activity. The rate of the reaction increases as a group with a  $pK_a$  of approximately 7 ionizes. Ionizing groups with a  $pK_a$  of about 7 in proteins are usually histidine residues. Schoellman and Shaw (47) showed that tosylamino-2-phenylethylchloromethylketone

(TPCK) alkylates His 57 irreversibly with accompanying loss of enzyme activity.

Also implicated in the activity of CHT is a carboxyl group, Asp 194, which appears to be involved in the activation of the zymogen, chymotrypsinogen (48).

The kinetics of  $\alpha$ -CHT catalysis follow a general mechanism of the type (49)



where ES is a complex,  $EP_2$  is the acyl enzyme, and  $P_1$  and  $P_2$  are products.  $P_2$  is an acid; and,  $P_1$  is an alcohol if S, the substrate, is an ester and is an amine if S is an amide. ES represents at least one intermediate formed prior to the acyl enzyme; the complex may result from conformational changes or from formation of a tetrahedral intermediate prior to acylation.

For esters the hydrolysis of the acyl enzyme is the rate determining step while for amides the rate is determined by the formation of the acyl enzyme. The kinetics of the hydrolysis of normal peptide substrates, esters, and amides are too fast to allow the acyl enzyme intermediate to be studied in the crystal state; therefore, pseudo-substrates have to be studied.

## 2. N-Formyl-Phenylalanine (NFP)

Many of the carboxylic acids which are products of  $\alpha$ -CHT catalyzed hydrolysis reactions are competitive inhibitors since they compete for the active site of the enzyme. Bender and Kemp (50) in studying these acids observed that  $\alpha$ -CHT catalyzed the exchange of oxygen atoms between the solvent and the carboxyl group; thus they can be considered to be substrates or virtual substrates of the enzyme.

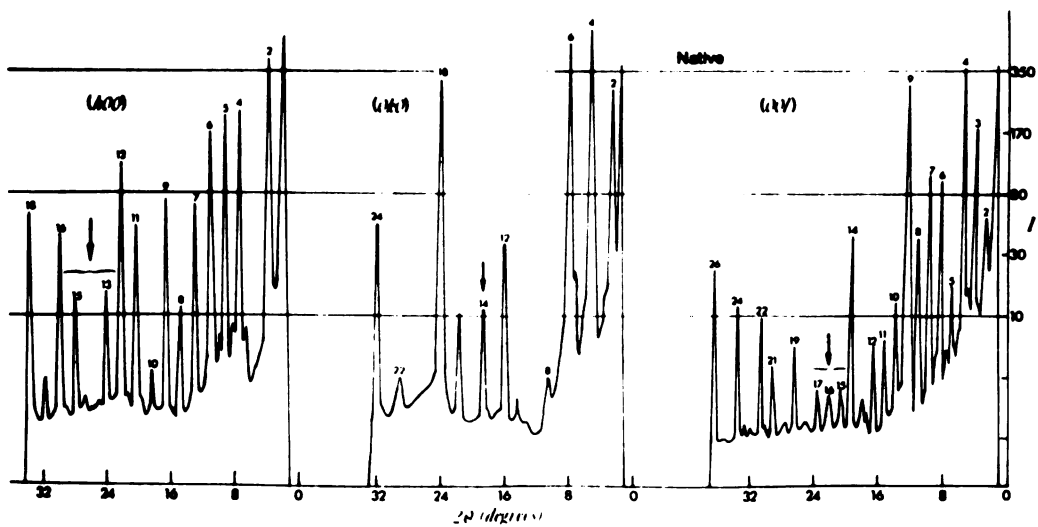
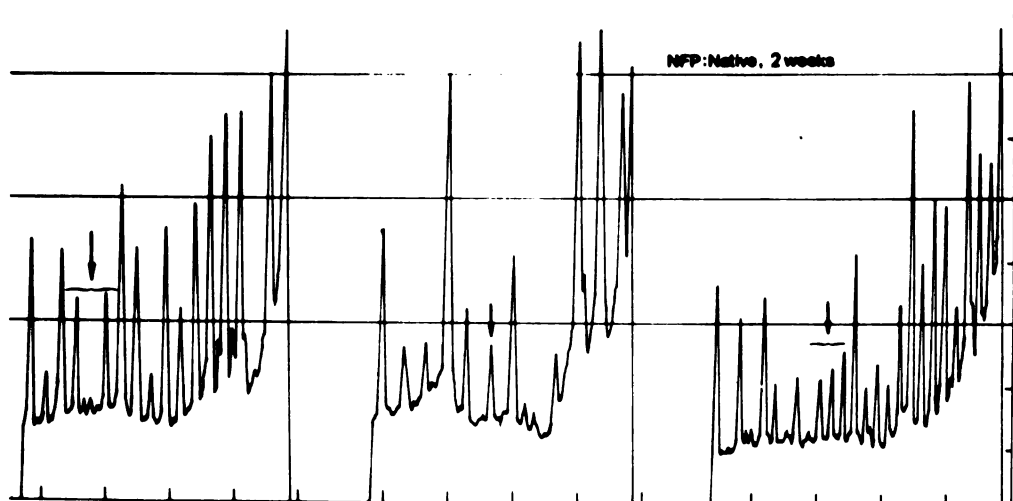
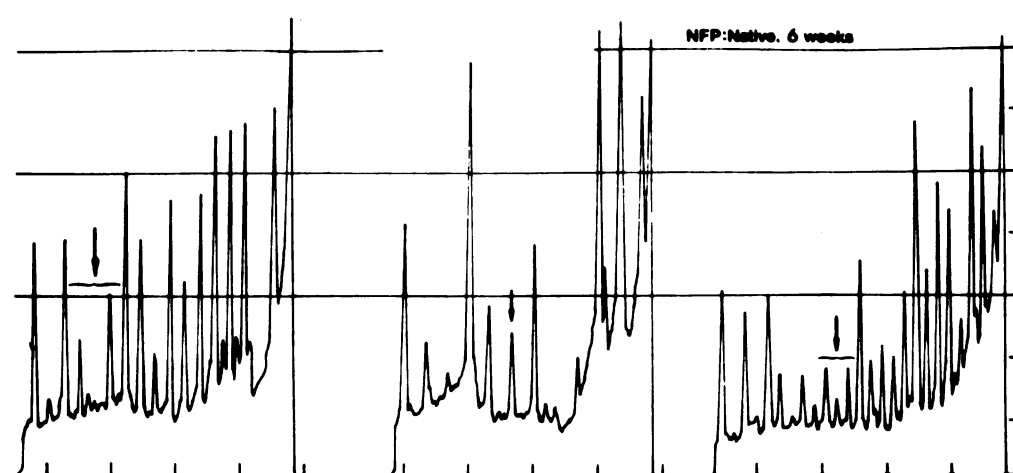
Such a virtual substrate should bind to  $\alpha$ -CHT in a similar mode as a true substrate so that a structural study of the enzyme:hydrolysis product complex might reveal the structural arrangement which helps determine the functionality of the native enzyme molecule. Upon examination of the two dimensional, centrosymmetric, difference electron densities (between the native enzyme and the enzyme complex) for several competitive inhibitors (51), NFP was found to show representative changes from the native structure so this complex was chosen to be studied in 3-dimensions and at high resolution.

The structure of NFP: $\alpha$ -CHT was briefly commented on by Steitz, Henderson, and Blow (52) in 1969. However, the crystals that were used for data collection were soaked in the inhibitor solution for only 3 to 24 hours before mounting. Crystals which were left soaking for longer periods of time were badly cracked. Furthermore, Steitz, et al., also reported that a dioxane molecule of crystallization was identified in the active site of the native enzyme and was subsequently diffused out of the crystals before intensity data collection. Finally, the difference Fourier map was an average of the difference electron density of the two crystallographically independent molecules. For such reasons, it was considered desirable to reinvestigate the structure of this complex.

In contrast to these results, the experiments carried out here yielded crystals which showed no cracking with prolonged soaking (about one month) and the substitution of NFP seemed to be a relatively slow equilibrium process (more than a month). This was indicated by an examination of the diffraction patterns of crystals soaked for two and six weeks. The axial intensity data (Figure 18) show greater differences from the native enzyme pattern for the six-weeks-soak than for the two-weeks-soak.

Figure 18. Axial diffraction patterns for Native, NFP (2 weeks), and NFP (6 weeks). Major changes are indicated by arrows.





Furthermore, the crystallization does not involve dioxane thus eliminating any complications arising from the inclusion of this molecule of solvation (53). Finally, the two molecules of  $\alpha$ -CHT in the asymmetric unit have observable differences (54); therefore, the NFP: $\alpha$ -CHT complex might also exhibit similar departures from local two-fold symmetry. Therefore, a structural study of this complex using the native enzyme structure determined at MSU was undertaken to investigate the geometry of the substrate interaction in the active site in an effort to gain some understanding of the functionality of the enzyme.

## VIII. Protein Crystallography

### 1. General

Since enzyme molecules are composed of L-amino acids they cannot be arranged in a unit cell which contains a center of symmetry. The phase problem for such non-centrosymmetric crystals is more difficult to solve because the phase angle,  $\alpha$ , can assume any value from 0 to  $2\pi$  radians.

For proteins the phase problem is usually solved using the method of multiple isomorphous replacement (55), wherein the data from several heavy atom complexes with the enzyme are used along with the positions of the heavy atoms determined by Patterson methods, to solve the phase problem. In principle, two heavy atom derivatives are sufficient to solve the phase problem uniquely when there are no errors present in the data. In practice, several derivatives are required to determine the phase of the native enzyme. The analysis of each heavy atom derivative yields a probability distribution for the value of the phase of the native enzyme. Two types of phases can be obtained from the product of these individual probability distributions, the most probable phase angle,  $\alpha_{\text{MAX}}$ , and best phase angle,  $\alpha_{\text{BEST}}$ , which is the phase of the center of gravity of the probability distribution. A measure of the quality of a phase determination is the figure of merit,  $m$ , of a reflection which is defined as the mean of the cosine of the error in the phase angle and is equal to 1.0 for a perfect phase determination.

One limiting factor of heavy atom derivatives is that their phase determining power decreases with increasing resolution since lack of isomorphism becomes more serious with increasing resolution. The resolution of an electron density image depends upon the minimum interplanar spacing,  $d_{\min} = \lambda / 2 \sin \theta_{\max}$ , of the F's included in the calculation. Detail on a smaller scale than  $r = 0.61 d_{\min}$  cannot be resolved. Usually protein structure analyses cannot be extended to atomic resolution due to difficulties associated with phase determination and inherent lack of sufficient order to produce observations in this scattering range (high Bragg angle).

The native enzyme structure amplitudes along with the phases determined from the isomorphous replacement method can be used to calculate the electron density of the enzyme crystal. This native enzyme data may then be used along with the structure amplitudes of the inhibited enzyme crystal to compute the difference electron density which will reveal the changes (substitutions and enzyme movements) effected by the inhibitor molecule. This difference calculation assumes that the inhibitor:enzyme complex is sufficiently isomorphous with the native enzyme that the phases of the two structures are similar. The maximum probability difference Fourier has coefficients of the form

$$(|F_{P+I}| - |F_P|) e^{i\alpha_{\max}}, \quad (20)$$

where  $|F_{P+I}|$  and  $|F_P|$  are the structure amplitudes of the inhibitor:enzyme and native enzyme, respectively. The reflections included in the maximum probability synthesis are usually limited to those with well determined phases, e.g.  $m \geq 0.7$ . The best difference synthesis

employs nearly all of the observed reflections, but weights the terms so that contributions from reflections with poorly determined phases are decreased. The coefficients of the best difference electron density,  $\Delta\rho_{\text{BEST}}$ , are of the form

$$m(|F_{P+I}| - |F_P|)e^{i\alpha_{\text{BEST}}}, \quad (21)$$

where the mean square error in  $\Delta\rho_{\text{BEST}}$ , as given by Henderson and Moffat (56), is

$$\langle \Delta\rho^2 \rangle = \frac{2}{V^2} \sum \Delta F^2 (1-m^2) + \Delta F^2 + \delta^2 \quad (22)$$

and  $\Delta F^2 = (|F_{P+I}| - |F_P|)^2$ . The term  $\Delta F^2(1-m^2)$  is due to the experimental errors in the phases, the  $\Delta F^2$  term is due to the error in the use of  $\Delta F$  as the scattering factor of the inhibitor, and the  $\delta^2$  term is due to the experimental errors in the  $\Delta F$  values. The difference Fourier technique is a valuable method of determining the structures of isomorphous chemical modifications of the native enzyme. The method allows the investigation of many similar structures readily after the phases are determined for the native enzyme structure.

## 2. Structure Determination of $\alpha$ -Chymotrypsin

Since the investigation of the structure of NFP: $\alpha$ -CHT depends upon the determination of the phases and thus the structure of the native enzyme a brief summary of this work (53) will be presented here.

The crystals of  $\alpha$ -CHT are monoclinic, space group  $P2_1$  with four molecules in the unit cell and thus two molecules per asymmetric unit. The unit cell dimensions, from an average of the parameters of seven crystals, are  $a = 49.24(7)$ ,  $b = 67.20(10)$ ,  $c = 65.94(9)$  Å and  $\beta = 101.79(8)^\circ$ . The standard deviations of the last digit from

the average are in parenthesis. The two molecules in the asymmetric unit are related by a non-crystallographic local two-fold axis (57). The equation for this axis, as determined from the MSU 3.5 Å native enzyme electron density map, is

$$\begin{aligned} z &= 0.158 x + 0.367 \\ y &= 0.708 \end{aligned} \tag{23}$$

where (x, y, z) are fractional coordinates.

The phases were determined by multiple isomorphous replacement with six heavy-atom derivatives to 2.8 Å resolution. The phase analysis produced an average figure of merit,  $\langle m \rangle$ , of 0.76 with two-thirds of the data having a figure of merit greater than the average (see Figure 19). This gives a root mean square error in the best native electron density at 2.8 Å resolution of  $0.20 \text{ e}/\text{\AA}^3$ .

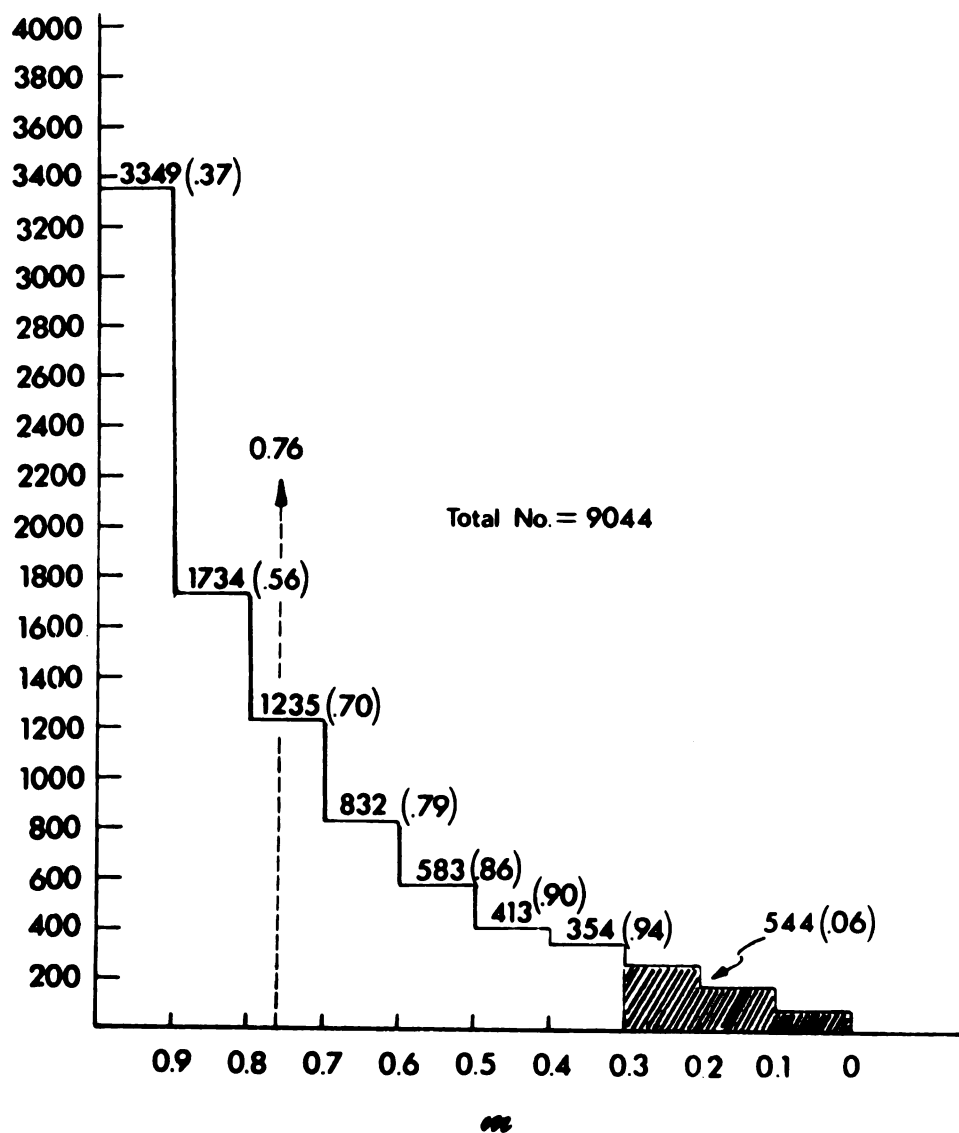


Figure 19. Histogram of the number of reflections in each figure of merit range for  $\alpha$ -CHT, from Tulinsky, et al., (53).

## IX. Experimental

### 1. Crystal and Derivative Preparation

Proteins crystallize with a considerable quantity of the solvent in a fairly mobile state within each unit cell. The presence of the mother liquor is necessary to keep the crystals well ordered but the composition of the solvent can be changed. Fairly large molecules can diffuse into protein crystals and enter each unit cell without disturbing the structure of the protein molecules. This is the basis of the preparation of heavy atom derivatives as well as the means for studying inhibitor:enzyme complexes.

$\alpha$ -CHT crystals contain approximately 40% mother liquor so that diffusion of inhibitor molecules into the crystals is readily accomplished.  $\alpha$ -CHT crystals are grown from about half-saturated ammonium sulphate solutions at pH  $\sim$  4.0 and then they are stored over 75% saturated ammonium sulphate at the same pH. NFP, obtained from International Chemical and Nuclear Corp., Irvine, California, and used without further purification, was dissolved in a few drops of acetonitrile and the solution was diluted to a volume of 1-2 ml with a 75% saturated ammonium sulphate solution. Then approximately one-half of the volume of storing solution in a tube of  $\alpha$ -CHT crystals was replaced with the inhibitor solution producing an approximate molar ratio of NFP: $\alpha$ -CHT of 25:1 (58).



After two weeks of soaking, a crystal was mounted and its diffraction pattern showed differences from that of the native enzyme (see Figure 18). The changes in the diffraction pattern became stationary after about six weeks of soaking; the diffraction data of the NFP: $\alpha$ -CHT complex were obtained at this time.

## 2. Intensity Data Collection

The NFP: $\alpha$ -CHT data were obtained to 2.8 Å resolution which includes  $\sim 11,000$  reflections to  $2\theta = 32.0^\circ$ . The data were collected from five crystals because the crystals suffer radiation damage and usually have to be abandoned after 35-40 hours of X-ray exposure. The data were collected in concentric shells of  $2\theta$  with small regions of overlap (see Table IX). The reflections in the overlap regions are used to verify the scale of the data from the different crystals.

The NFP: $\alpha$ -CHT data were collected with the same instrument, counting technique, and automatic realignment program as were used for TPrP (see Section II). The data for the five crystals were collected and converted to structure amplitudes separately and then merged into one data set. The total data collection included 11,885 intensity measurements over a 16 day period. The data treatment was much the same as that used for the TPrP data. The differences in these procedures caused by the special properties of the protein crystals are discussed below.

The crystals were always mounted with the unique axis,  $\underline{b}$ , parallel to the  $\phi$  axis of the four circle diffractometer (Figure 5). The quadrant for data collection was defined by  $-\underline{a}^*$  and  $+\underline{a}^*$  including  $+\underline{c}^*$ .

Table IX. Data Ranges for the Enzyme

Crystal	2 $\theta$ -Range (degrees)	d <sub>min</sub> (Å)	Number of Reflections
1	2.5-20.0	4.4	2800
2	19.7-25.0	3.6	2650
3	24.75-28.0	3.2	2200
4	27.80-30.15	3.0	1900
5	29.90-32.0	2.8	1900

Before data collection,  $2\theta$  scans of the axial data were made to ensure crystal quality and reproducibility of the derivative diffraction pattern. These axial runouts were repeated following the data collection to check for systematic changes in the diffraction which might arise from radiation damage or loss of humidity. In addition to this check, two sets of centric ( $h0l$ ) projection data to 6 Å resolution ( $2\theta < 15^\circ$ ) were collected before and after data collection.

These ( $h0l$ ) data sets served several purposes. A comparison of the before and after data provided a check on the stability of the substitution during data collection and provided some measure of the decay of the intensities. The ( $h0l$ ) data sets were used to calculate electron density projections from which an approximate scale factor to the native data could be obtained. Finally a comparison of the difference Fourier projections provided a further check on the reproducibility of the amount and position of substitution in the several crystals.

The absorption, lack of balance, and twin size of the crystals were examined before data collection. The absorption measurements were made as described for the TPrP crystal on four  $b$  axis reflections at  $\chi = 90^\circ$ ; these were the (020), (040), (060), and (0,18,0) with  $2\theta = 2.6, 5.3, 7.9$  and  $23.8^\circ$ , respectively. Thus, any  $2\theta$  dependence of the relative absorption could be detected. Lack of balance measurements were made as previously for the TPrP analysis.

Crystals of  $\alpha$ -CHT are twinned along the  $c^*$  direction (53). A twinned crystal appears to be a single crystal but is in fact two different interpenetrating orientations of a lattice. In the  $\alpha$ -CHT crystals the two lattices practically superimpose for the ( $0kl$ )

reflections. The size of the twin varied among the crystals and by careful choice of the crystals for data collection the effect of the twin lattice can be minimized. Before data collection the intensities of the (600) reflection for the real and the twin lattice are compared to determine the magnitude of the twin.

The cell dimensions for each inhibitor:enzyme complex crystal and the orientation of the lattice were determined from the angular coordinates of 12 reflections distributed throughout reciprocal space. The cell constants for the five data crystals are given in Table X. The average cell parameters for the NFP: $\alpha$ -chymotrypsin complex are  $\underline{a} = 49.27(2)$ ,  $\underline{b} = 67.10(6)$ ,  $\underline{c} = 65.84(7)$  Å and  $\beta = 101.75(2)^\circ$ . A comparison of these values with the average cell dimensions of the native enzyme crystals indicates that the two kinds of crystal are the same within the error of the determination.

The intensities of three reflections were monitored every 100 reflections ( $\sim 1$  hour) throughout the data collection as a check on the crystal alignment and on radiation damage. These reflections were the (0,18,0) at  $\chi = 90^\circ$  over both the  $+\underline{a}^*$  and  $+\underline{c}^*$  axes and the general reflection (4,14,5) at a  $\chi$  value of approximately  $59^\circ$ .

## 2. Crystal Motion

The enzyme crystals are mounted in glass capillaries and are held to the wall of the capillary by the surface tension of the mother liquor. Crystals mounted in this manner usually move continuously to new orientations during data collection. The computer program, discussed in Section II, compensates for the reorientation of the crystal. A plot of the intensity of the (0,18,0) monitor reflection at  $+\underline{c}^*$  from the crystal 2 data is shown in Figure 20. This plot

Table X. The Cell Parameters for NFP: $\alpha$ -Chymotrypsin

Crystal	<u>a</u> (Å)	<u>b</u> (Å)	<u>c</u> (Å)	$\beta$ (degrees)
1	49.27	67.04	65.91	101.75
2	49.24	67.16	65.79	101.77
3	49.28	67.11	65.77	101.74
4	49.26	67.16	65.81	101.74
5	49.28	67.02	65.92	101.75
Average	49.27(2)	67.10(6)	65.84(7)	101.75(2)
$\alpha$ -CHT	49.24(7)	67.20(10)	65.94(9)	101.79(8)

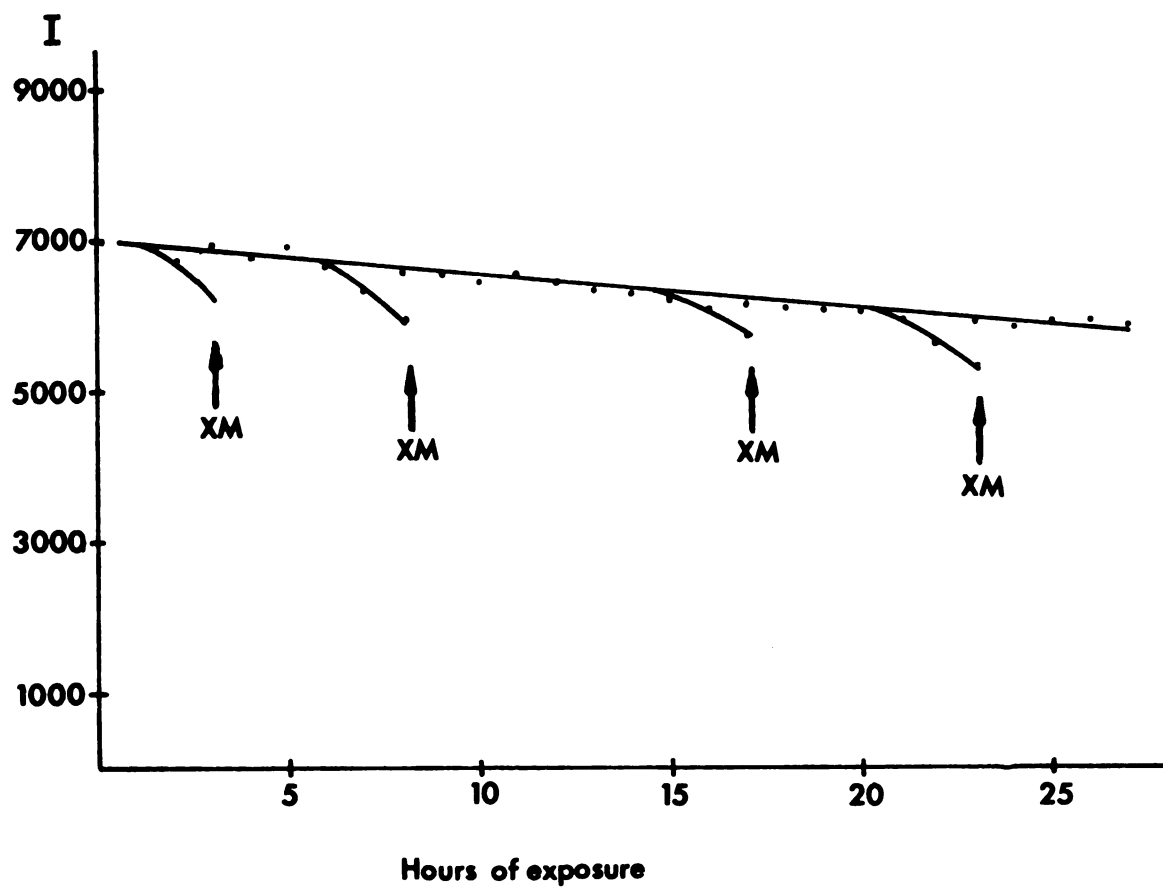


Figure 20. Plot of the intensity of the  $(0,18,0)$  vs. exposure time.

I in counts/16 seconds and XM means crystal was considered to have moved.

indicates the effective compensation for reorientation accomplished by the program. Table XI contains a record of the realignments necessary for each crystal along with the total hours of exposure for each crystal.

After the data for each crystal were collected the following corrections were applied: lack of balance, twin, absorption, decay, and Lorentz-polarization.

#### 4. Lack of Balance and Twin Corrections

The data were corrected for lack of balance by adding the measured correction, at the approximate  $2\theta$  angle, to the background intensity of the reflection. This correction was relatively large only for  $2\theta < 10^\circ$  and thus was important only for crystal 1.

The twin correction was applied to the  $(0kl)$  reflections by using a twin ratio calculated from the intensity of the  $(600)$  reflection for the real and the twin lattice. The twin ratios,  $k_t = I_{\text{real}}/I_{\text{twin}}$ , for the five data crystals are tabulated in Table XII. The observed intensities for the  $(0kl)$  reflections result from both the real and the twin lattice and therefore have to be reduced to only the real contribution. The correction is of the form

$$I_{\text{real}}(0kl) = \frac{k_t}{(k_t + 1)} I_{\text{OBS}}(0kl).$$

#### 5. Absorption Correction

The absorption measurements were corrected for lack of balance and then plotted as  $I_{\text{MAX}}/I(\phi)$  vs.  $\phi$ , where  $I_{\text{MAX}}$  is the maximum observed intensity of the reflection, to show the relative absorption of the crystal. These plots were compared for each of the four  $(0k0)$

Table XI. Record of Crystal Motion

Crystal	Number of Realignments	Total Exposure (hours)
1	4	38
2	4	36
3	3	30
4	2	27
5	1*	26

\*This crystal really showed no motion; it was made to be realigned after 20 hrs. as a precautionary measure.



Table XII. Ratios of the Intensities of Real and Twin Lattices

Crystal	$I_{\text{real}}/I_{\text{twin}}$
1	8.42
2	3.56
3	3.37
4	2.01
5	6.36

reflections and were averaged to give one absorption correction if no major differences in the curves were observed. These plots are shown in Figures 21 and 22. For crystals 1 and 3 all four curves were the same and only one absorption correction was used in each case. For crystal 2 the (020) reflection showed a different absorption so two corrections were applied: the (020) curve for  $k \leq 2$  and an average curve for  $k > 2$ . The (0,18,0) reflection showed a different absorption in crystals 4 and 5 and an average correction was applied for  $k \leq 6$  and the (0,18,0) curve for  $k > 6$ . These differences probably arise from the difference in the path length of the X-ray beam through the capillary, mother liquor, and crystal at the different  $2\theta$  values of the (0k0) reflections. Crystals 2 and 4 (Figure 22) showed the maximum relative absorption with  $(\text{Abs.})_{\text{MAX}} = 1.99$  and  $2.15$ , respectively, where  $(\text{Abs.})_{\text{MAX}} = [I_{\text{MAX}}/I(\phi)]_{\text{MAX}}$ .

The absorption correction was made using the semi-empirical method of North, Phillips, and Mathews (59). In this method the transmission of a reflection by a crystal,  $T(hk\ell)$ , is defined as the average of the transmission in the direction of the incident and reflected beams, thus

$$T(hk\ell) = [T(\phi_{\text{inc}}) + T(\phi_{\text{refl}})]/2 \quad (24)$$

where  $\text{Abs}(hk\ell) = 1/T(hk\ell)$ . This correction accounts for the change in the path length of the incident and reflected beams as  $\chi$  is changed from  $90^\circ$  to  $0^\circ$ . The values of  $\phi_{\text{inc}}$  and  $\phi_{\text{refl}}$  are calculated from  $\phi_{hk\ell}$  and for the diffractometer geometry are

$$\begin{aligned} \phi_{\text{inc}} &= \phi_{hk\ell} - \sin^{-1}(\sin\theta\cos\chi) \\ \phi_{\text{refl}} &= \phi_{hk\ell} + \sin^{-1}(\sin\theta\cos\chi) \end{aligned} \quad (25)$$

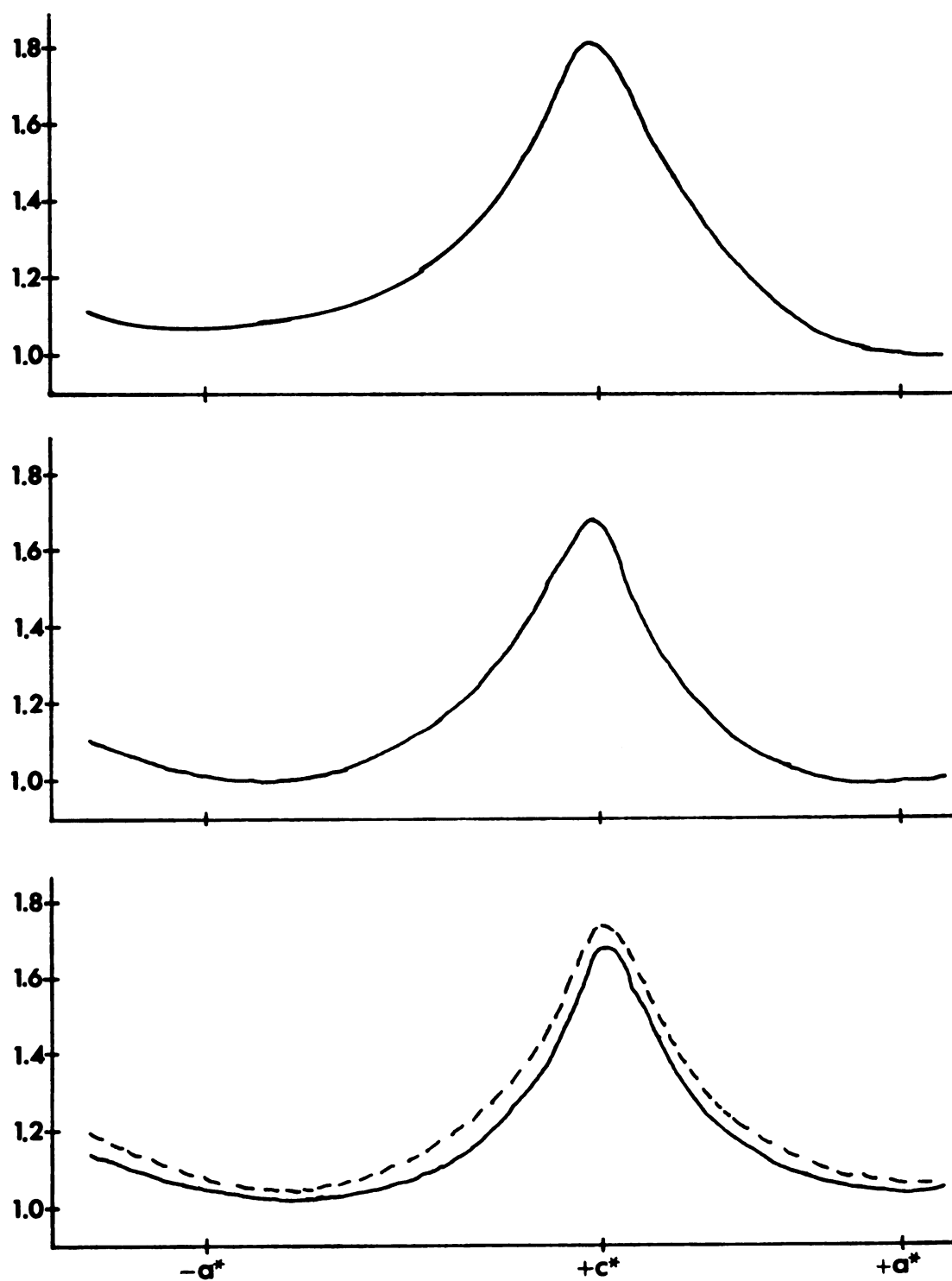


Figure 21. Plot of the relative absorption observed at  $\chi = 90^\circ$  for crystals 1,3, and 5. The average curves are solid lines and the individual curves are dashed lines.

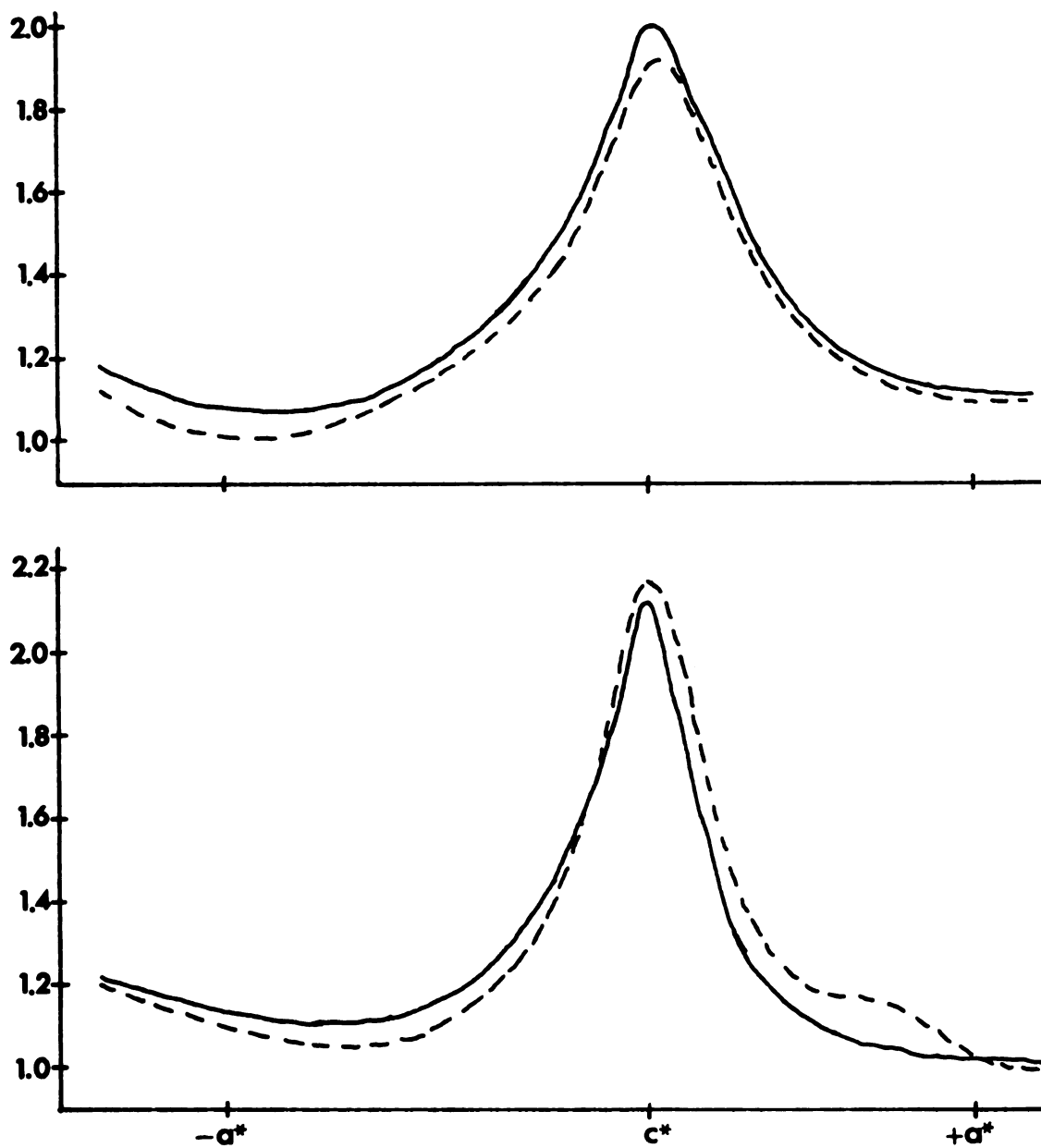


Figure 22. Plot of the relative absorption observed at  $\chi = 90^\circ$  for crystals 2 and 4. The average curves are solid lines and the individual curves are dashed lines.

With these  $\phi$  values the absorption corrections are obtained from the curves described above and used to calculate  $Abs(hkl)$  from

$$Abs(hkl) = \frac{2Abs(\phi_{inc})Abs(\phi_{refl})}{Abs(\phi_{inc}) + Abs(\phi_{refl})} \quad (26)$$

which is a rearrangement of Equation 24. This method gives a good estimate of the relative absorption correction.

## 6. Decay Correction

Another systematic error present in the data is the loss of intensity with exposure to X-rays. This effect is observed in the decrease in the intensities of the monitor reflection with increased exposure (see Figure 20) and is the major reason that several crystals have to be used for a data set. The decay correction is obtained from the monitor reflections by fitting a plot of  $I(1)/I(n)$  vs.  $n$  with a least square line, where  $n$  is the number of the reflection or the time at which it was measured. The reflections are numbered consecutively as they are measured and the reflection number,  $n$ , is used as an estimate of elapsed time. The slope of the least squares line (intercept = 1.0) is used to obtain the decay correction for any reflection from the equation

$$Decay(n) = 1.0 + (slope) \times n.$$

Most of the reflections collected for this data set have  $2\theta$  values between  $20^\circ$  and  $32^\circ$  thus the decay of the monitor reflections at  $2\theta = 21^\circ$  and  $23^\circ$  are good estimates for all reflections. However, the data set for crystal 1 contains the reflections with  $2\theta < 20^\circ$  and for this range, an additional decay correction may be obtained from

the before and after ( $h0l$ ) data. The decay correction for low  $2\theta (<15^\circ)$  is obtained from the scale factor between the ( $h0l$ ) data for the NFP complex before and after data collection. The maximum decay for the data with  $2\theta < 15^\circ$  is defined as the ratio between the scale factor for the after data and the scale factor for the before data.

Two decay corrections were applied to the crystal 1 data: for reflections with  $2\theta \leq 15^\circ$  the decay was obtained from the ( $h0l$ ) scale factors and for reflections with  $2\theta > 15^\circ$  the correction was obtained from the slope of the decay curve of the (4,14,5) reflection ( $2\theta = 21^\circ$ ). The rest of the data from crystals 2-5 were corrected for decay with the slope of the average decay curve of all three monitor reflections since these curves were approximately the same. Table XIII gives the decay correction data for all five crystals. The largest decrease in intensity was 21% for crystal 2.

With the above corrections the five data sets were converted to structure amplitudes which were then merged and averaged into one set of independent reflections.

## 7. Data Processing

The initial scale factor for each data set was obtained from a comparison of the peaks in the before ( $h0l$ ) projections, to 6 Å resolution, of the electron density of the native enzyme and the NFP complex. The scale factor for the NFP complex structure amplitudes was defined as the scale factor necessary for the peak heights in the two projections to be equal (excluding the positions where the NFP substitution produced changes).

Table XIII. Decay Correction Data

Crystal	2 $\theta$ Ranges	Number of Reflections	Slope( $\times 10^5$ )	Maximum Correction
1	0<2 $\theta$ <15	2904	1.72	1.05
	15<2 $\theta$ <20		5.08	1.15
2	all	2713	7.85	1.21
3	all	2301	5.33	1.12
4	all	1990	4.37	1.09
5	all	1977	4.91	1.10

The data sets were merged with the scale factor between each two sets of data being additionally verified by a comparison of the redundant reflections arising from the overlap of the data ranges. The value of  $|F|$  for the reflections in the overlap region was taken as the average of the values of  $|F|$  for the redundant reflections. The quality of each merge was assessed by evaluating the quantity  $R_r$ ,

$$R_r = \sum ||F_i| - \langle |F| \rangle| / \sum \langle |F| \rangle$$

where  $|F_i|$  and  $\langle |F| \rangle$  are the redundant structure amplitudes and their average value, respectively, and the sum includes all redundant reflections. The values of  $R_r$  for the four merges of the NFP complex data are given in Table XIV.

Another check on the correct scaling of the merged data was provided by the radial distributions of  $|F_p|^2$  and  $|F_{p+I}|^2$  which are plots of the  $\langle |F|^2 \rangle$  vs.  $\langle 2\theta \rangle$ . The relative distribution of the magnitudes of  $\langle |F|^2 \rangle$  over the range of the data should not change appreciably with the substitution of an inhibitor molecule. Therefore, the radial distribution curve of NFP: $\alpha$ -CHT should have nearly the same shape as the native curve and since the contribution of the inhibitor is small the magnitudes of the curves should also be nearly equal. The radial distribution curves for the native enzyme and the NFP complex to 2.8 Å resolution are shown in Figure 23. The shapes of these curves are similar and a final scale factor of 0.98 on the NFP: $\alpha$ -CHT data caused the magnitudes of the two curves to be approximately equal.

The final set of observations for the NFP complex contained 10473 independent reflections of which 9810 were taken as observed reflections with an average reproducibility of about 2.6%.



Table XIV.  $R_r$  Factors for the Redundant Reflections

Crystals Merged	Redundant Reflections	$R_r$	Total Observed Reflections
1 and 2	540	0.027	4932
1,2 and 3	381	0.026	6864
1,2,3 and 4	371	0.025	8321
1,2,3,4 and 5	413	0.027	9810

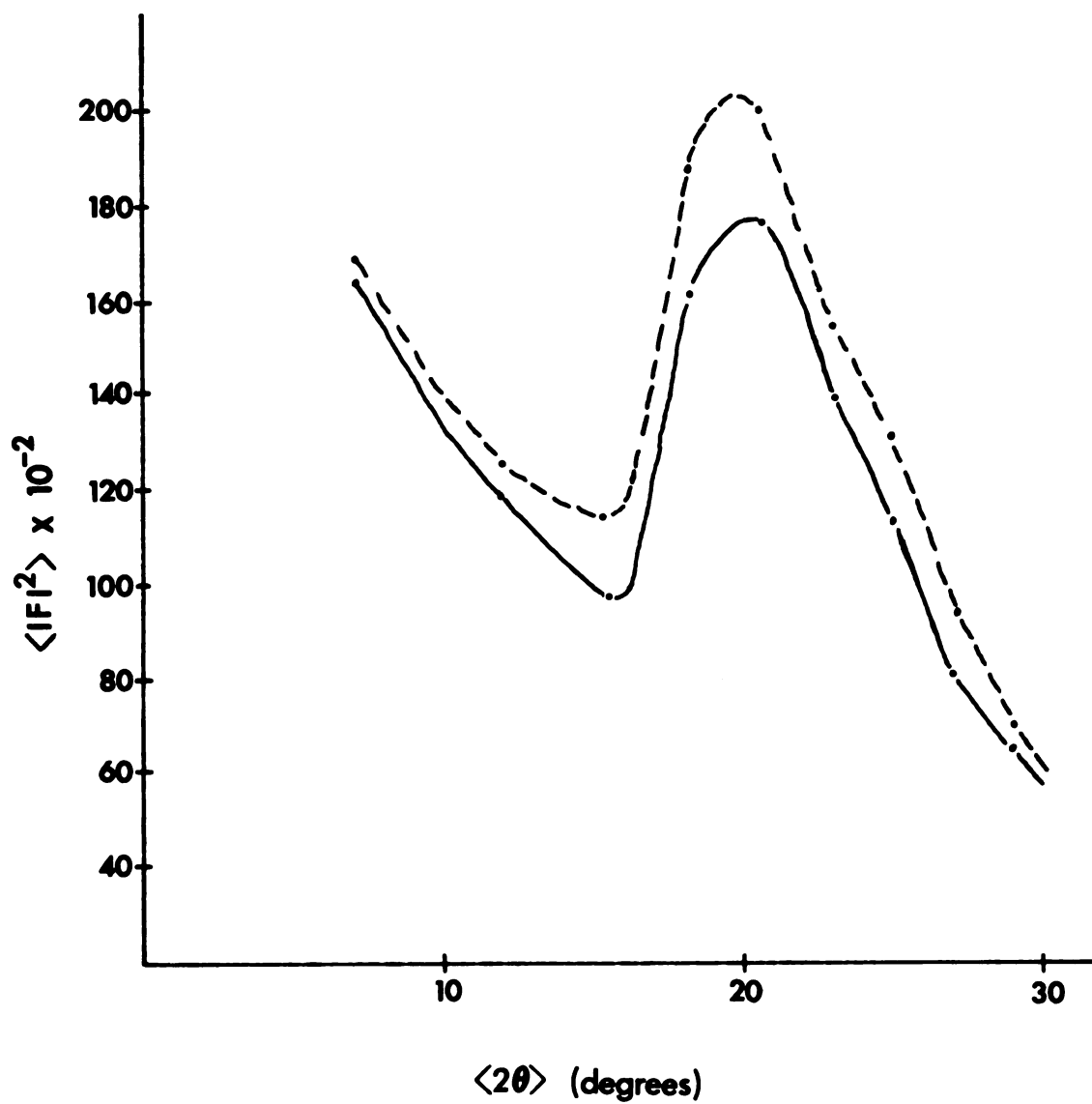


Figure 23. The radial distribution of Native (solid line) and NFP: Native (dashed line).

## X. Discussion of the Results

### 1. Substitution in the Active Site Region

The active site regions of the two molecules of  $\alpha$ -CHT in the asymmetric unit are spatially located close to each other across a local two-fold rotation axis. The Tyr 146 residue of one molecule of  $\alpha$ -CHT penetrates the molecular boundary of the two-fold related molecule and is located in close proximity to the active site region of this molecule.

The substitution of NFP is found to occur in a space in the active site region such that the inhibitor molecule is surrounded by portions of the peptide chains of  $\alpha$ -CHT. The composite 2.8 Å resolution electron density of the active site region along with the difference electron density of NFP is shown viewed parallel to the two-fold axis in Figure 24(a). The distances of the active site residues from the NFP substitution positions are given in Figure 24(b). The distances are taken from the midpoints of the contours except for the Ser residues where the distance was taken from the estimated position of the hydroxyl side chain. The NFP peaks in the unprimed molecule are attributed to the phenyl ring (P) with a peak height of  $0.18 \text{ eÅ}^{-3}$  and to either the carboxyl or acylamido group (PX) with a peak height of  $0.13 \text{ eÅ}^{-3}$ . The peaks P and PX are separated by 3.4 Å which agrees with the approximate geometry of the NFP molecule. The phenyl position partially overlaps the position of a localized water molecule which

1

Figure 24(a) Electron density in the active site region viewed parallel to the local two-fold axis.

The native electron density is drawn at  $0.25 \text{ eA}^{-3}$  intervals beginning with  $0.5 \text{ eA}^{-3}$ . The difference electron density is cross-hatched and drawn at  $0.04 \text{ eA}^{-3}$  intervals beginning with  $0.1 \text{ eA}^{-3}$ . The two-fold axis is marked with an asterisk.

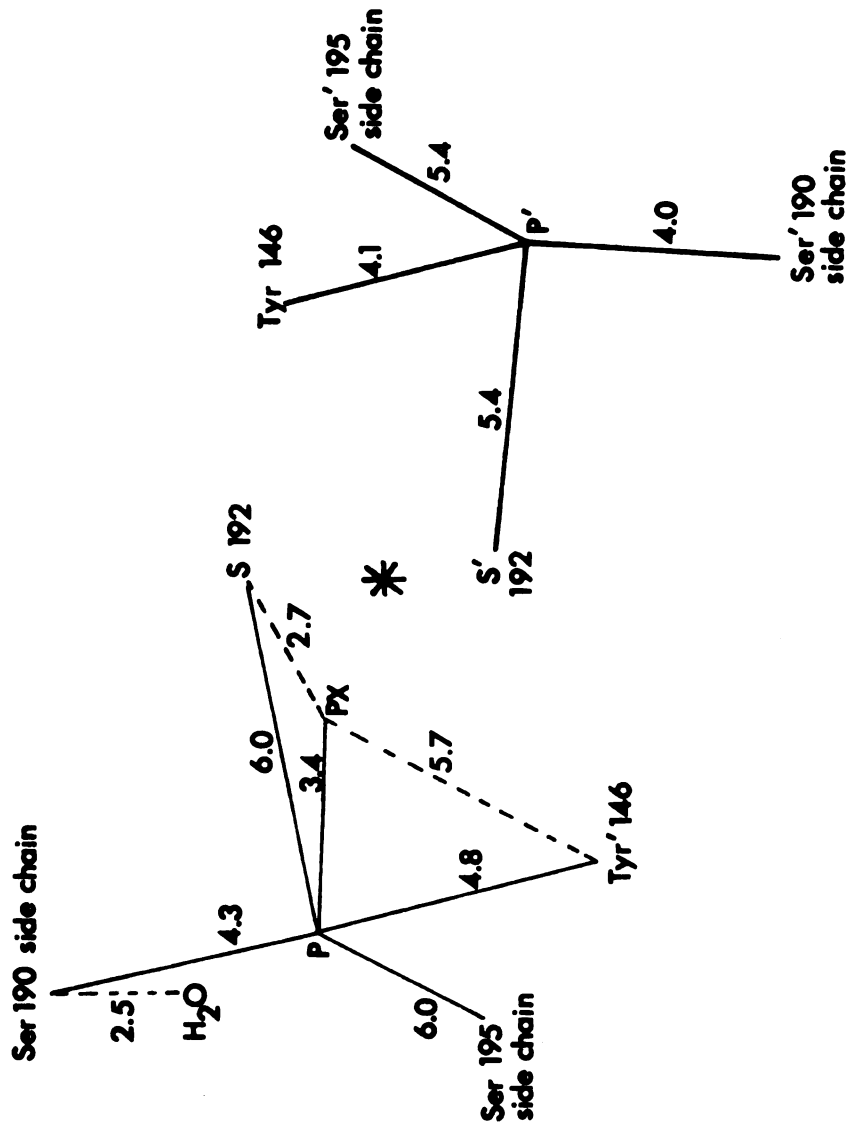


Figure 24(b) Distances (in A) from NFP in the electron density of Figure 24(a). Two-fold axis marked with asterisk. The Ser 190 and Ser 195 distances were taken from the estimated positions of their hydroxyl groups.

is hydrogen bonded to Ser 190 in the native enzyme structure and is displaced upon binding of the inhibitor molecule. The substitution position of the NFP molecule is such that the peak PX and the sulfur side chain of Met 192 have a close contact of about 2.7 Å. This interaction might be a hydrogen bond formation with the sulfur. The PX position is near the main chain of Met 192 which is above the substitution positions in the electron density presented in Figure 24(a). The interaction with the Met 192 residue and its side chain may help to position the substrate in the proper orientation for hydrolysis.

In the primed molecule only one peak (P') is observed (at a peak height of  $0.14 \text{ eÅ}^{-3}$ ) and is related by the local two-fold axis to the phenyl ring peak in the unprimed molecule. The smaller peak height for the NFP substitution in the primed molecule of  $\alpha$ -CHT suggests that this active site region has less affinity for the inhibitor molecule. The primed molecule of  $\alpha$ -CHT does not have a water molecule hydrogen bonded to Ser' 190; however, another water molecule ( $W_1'$ ) which is associated with the Tyr 146 residue occupies part of the active site region and may interfere with the substitution. Even though some of the residues have close contacts with the substitution positions (Tyr 146, Ser 190, see Figure 24(b)), there are no major changes in the orientation of these residues in the active site region with the substitution of NFP.

Some small shifts in the primed molecule are evident when the other residues in the active site region are examined. These residues and shifts are shown in Figure 25(a) where the electron density of the active site region is viewed perpendicular to the local two-fold axis.

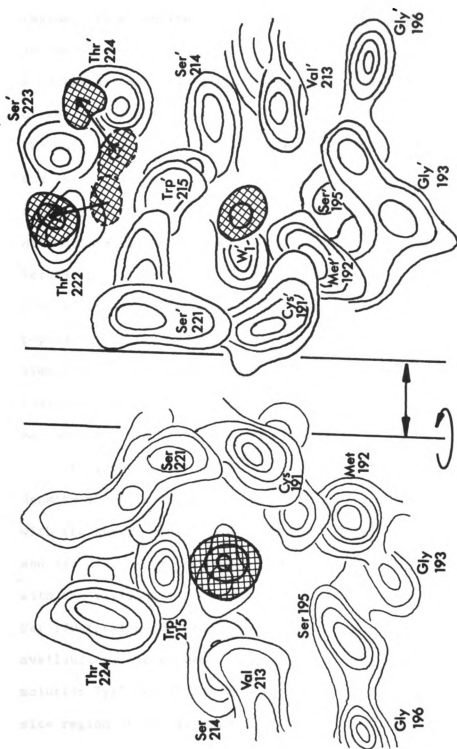


Figure 25(a) Electron density in the active site region perpendicular to two-fold axis. Contours are same as in Figure 24(a). The primed molecule is translated for clarity by the distance indicated by double arrows between the two-fold axes.



The shifts are in the positions of the main chain peptides of Thr' 222 and Ser' 223 by 0.25 Å and 0.35 Å, respectively, resulting in the movement of this part of the peptide chain away from the active site region. These shifts may be due to the more compact active site region in the primed molecule which is suggested by the distances in Figure 25(b). A comparison of these distances indicates that the two-fold symmetry of the active site region of the substituted molecules is not exact. The largest differences are those of the Thr 224 and Met 192 positions and these residues also show the largest deviations from two-fold symmetry in the native enzyme (see Figure 26). Thr 224 and Thr' 224 differ in position across the two-fold axis by about 3.8 Å and the Met 192 positions differ by approximately 2.0 Å. The resulting close Thr'-P' distance is probably the cause of the shift of the main chain peptides Thr' 222 and Ser' 223. The deviation of Met 192 from two-fold symmetry is particularly important because of the interaction of this residue with the inhibitor molecule. The short Met' 192-P' distance may be the cause of the decreased binding of NFP in the primed molecule.

The proposed PX position is beneath Cys 191 in the electron density shown in Figure 25(a). The other residues with close contacts with the phenyl position are the main chain peptides of Trp 215, Ser 214, and Val 213. Thus, the active site forms a cage around the NFP molecule with the residues in Figure 25(a) forming the walls and Tyr' 146 and Ser 190 forming the top and bottom of the cage. The NFP molecule is available to the solvent through an opening underneath Ser 221. In solution Tyr' 146 of one molecule does not interact with the active site region of the adjacent molecule so that the NFP molecule would be available to the solvent through the top of the cage also.

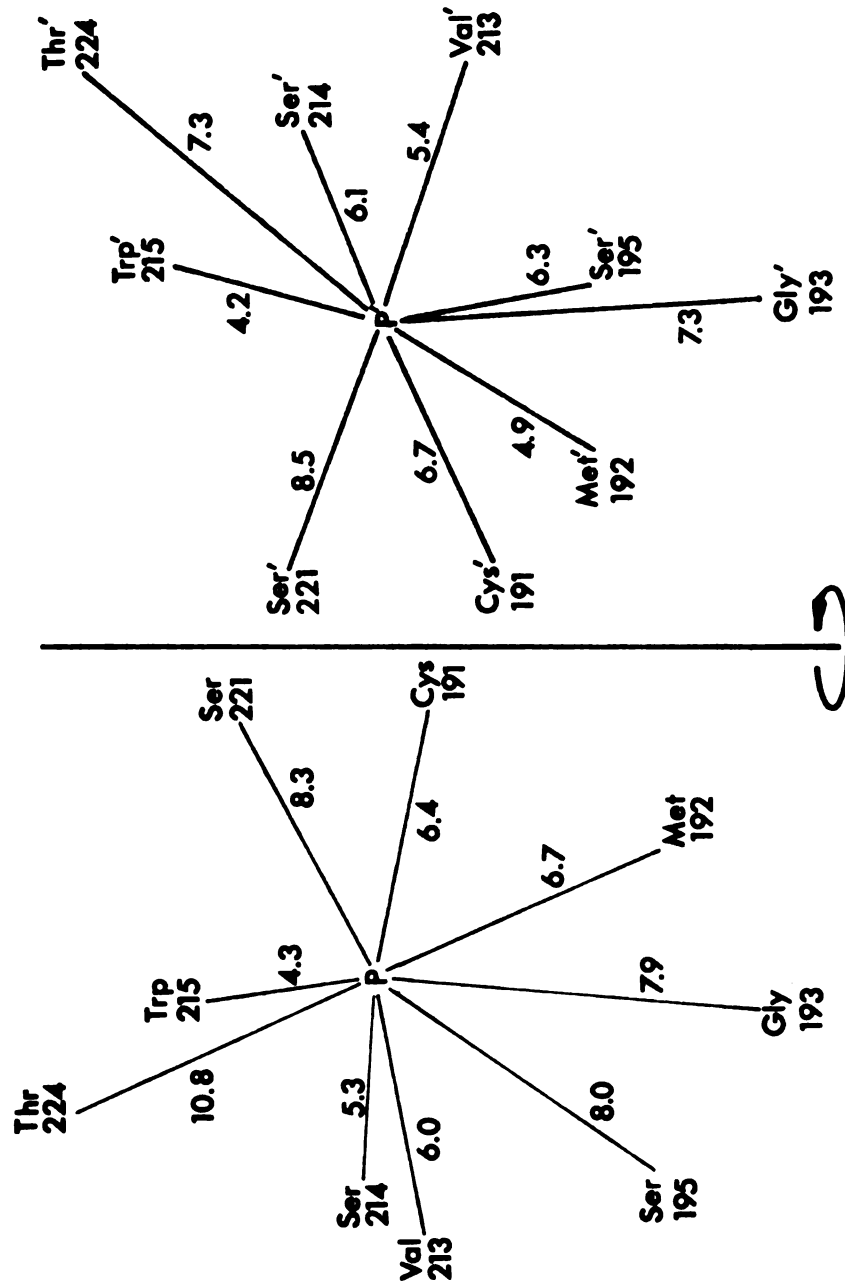


Figure 25(b) Distances (in Å) from NFP in the electron density shown in Figure 25(a). The distances were taken from the midpoint of the contours. Two-fold axis is heavy line.

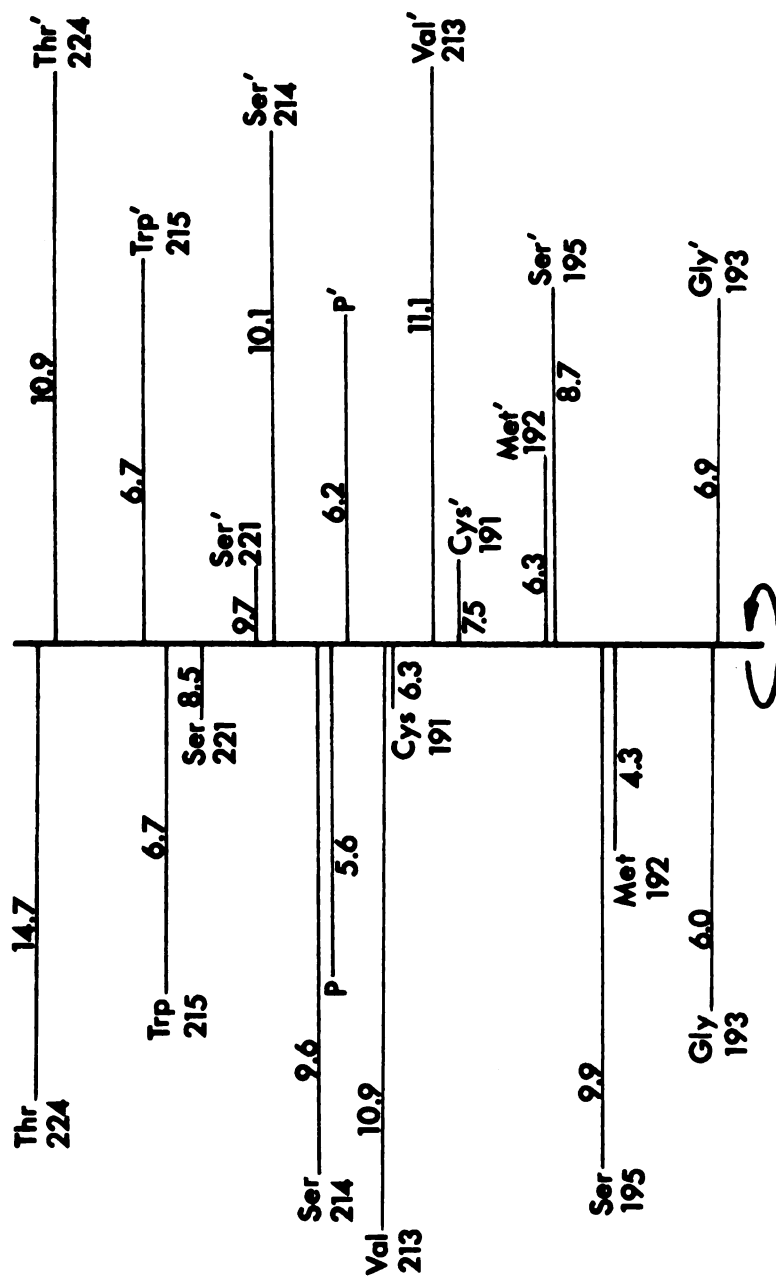


Figure 26 Distances (in Å) from the local two-fold axis (heavy line).

The relatively minor changes in the enzyme structure in the active site region are not typical of the effect of the NFP substitution on the rest of the molecule.

## 2. Histidine 40

One of the largest changes observed in the NFP: $\alpha$ -CHT complex occurs in the region of His 40 in both molecules and is shown in Figure 27. In the native enzyme structure the side chain of His' 40 appears to be hydrogen bonded to Gly' 43 while a similar interaction does not occur in the unprimed molecule. In addition, the positions of the main chain peptides of His 40 and His' 40 differ across the two-fold by 1.5-2.0 Å. Upon formation of the NFP: $\alpha$ -CHT complex the imidazole of His 40 moves by about 0.5 Å in the direction of Gly 43 possibly forming a hydrogen bond. The effect of NFP substitution on the primed molecule is more complicated and not fully understood because the shift of the His' 40 side chain is accompanied by a shift of 0.2 Å of the main chain peptide of Trp' 141 away from Gly' 43 and a shift of 0.2 Å by Cys' 42 toward the His' 40 side chain. The shift of Cys' 42 is not shown in Figure 27 since the residue lies under the Trp' 141 density.

A similar change in the region of His 40 and its two-fold related residue is observed when the pH of the native enzyme crystal is changed from 3.9 to 6.7 (60). The fact that this change occurs when a pseudo-substrate binds in the active site and when the pH is increased to a value near the pH of maximum activity suggest that this change is related to the activity of the enzyme. His 40 is found to be common in the sequences of chymotrypsinogen A, chymotrypsinogen B, trypsinogen, and porcine elastase (61) as is His 57 and the sequence Gly-Asp-Ser-Gly-Gly-Pro near Ser 195. These homologies in the sequences of the

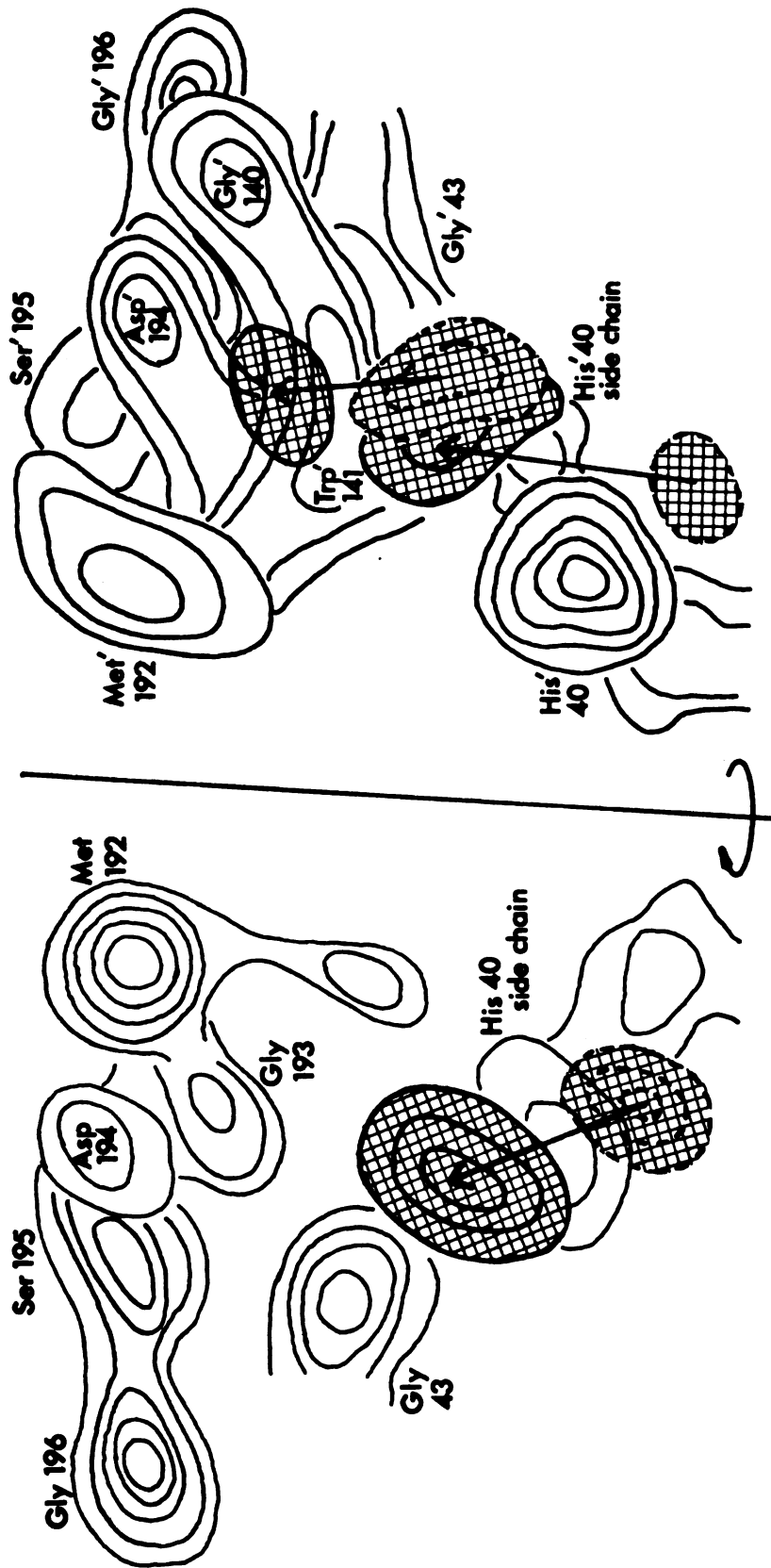


Figure 27. The region near Histidine 40. Two-fold axis is heavy line. The difference electron density is cross-hatched and contours are drawn as in Figure 24(a).

serine proteases along with the proximity of His 40 to the active site region, as shown in Figure 27, suggest that this residue may play a role in maintaining the conformation of the active site.

### 3. Other Similarities with pH Change

An intermolecular feature which is common to both pH change and NFP substitution occurs in the region of favored uranyl binding and is shown in Figure 28. This interaction involves the Arg' 154 side chain which shifts about 0.40 Å toward the side chain of Glu 21. This change is accompanied by a shift of the Val 23 side chain of 0.32 Å away from Glu 21 and a shift of the main chain peptide of Val' 23 and Ala' 22 of about 0.21 Å. The two-fold equivalent position does not exist since this region is formed by the molecular boundaries of three  $\alpha$ -CHT molecules.

Table XV gives the positions of changes in  $\alpha$ -CHT which result from the increase in pH. Those residues in boxes are also observed to change upon binding of NFP to the enzyme. The regions of similarity include: (1) those already discussed; (2) the change at Ser 159, which has a peak height of  $0.16 \text{ eA}^{-3}$ ; and (3) the change near Asp' 102, with a height of  $0.15 \text{ eA}^{-3}$ . The change in the region near Trp' 29- Trp' 207 is weak in the NFP density.

The lack of movement near Tyr' 146-His 57 is significant since these residues have been implicated in the dimerization of  $\alpha$ -CHT (62). The dimerization is a maximum at pH = 4.0 and the dimer dissociates at higher pH values. The movement of these residues in the pH study (60) was probably related to the dissociation and not particularly involved with the activity of  $\alpha$ -CHT. The substitution of NFP produces substrate-induced changes in the enzyme which have also been observed when the

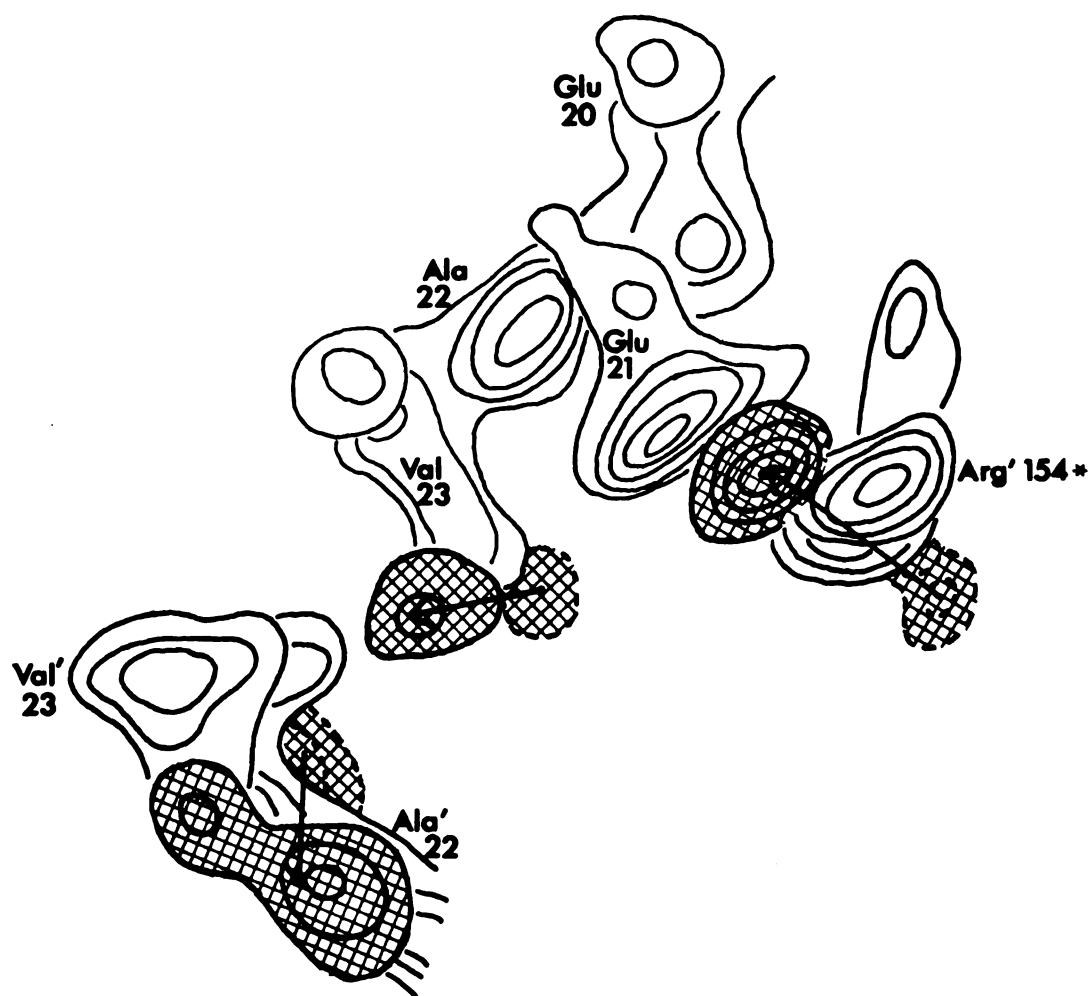


Figure 28. The uranyl substitution position. The Arg' 154\* residue is from the primed molecule related by the screw axis along b. The difference electron density is cross-hatched and the contours are drawn as in Figure 24(a).

Table XV. Positions of Structural Changes of  $\alpha$ -CHT  
with pH (from Vandlen and Tulinsky (60))

MAJOR CHANGES

His 40

Glu 21 - Arg' 154

Tyr' 146 - His 57

Trp' 29 - Trp' 207

-Phe 71 - Asp 72 - Gln 73 -

-Leu 155 - Arg 154 - Asp 153-

LESSER CHANGES

Ser 119

Ser 159

Ser 214

Trp' 141

Asp' 102 - Asn' 100



pH is changed to a value which increases the activity of the enzyme. The NFP substitution does not induce the same changes as increased pH when these changes are related to dimerization.

#### 4. Final comments

Several other changes with peak heights greater than  $0.15 \text{ eA}^{-3}$  are observed in the NFP difference electron density. The standard error in the difference electron density calculated using only the first term in Equation 22 is  $0.02 \text{ eA}^{-3}$ . These peaks, along with those already discussed, are identified in Table XVI. Substitution in an auxiliary binding site (51) was not observed for NFP and very few of the changes listed showed two-fold symmetry.

The results of this study are similar to those of the MRC group (52) in that the amino acid sequences near the aromatic residue are nearly the same; however, there are two major differences. The MRC study was confined to the active site region of the average map of the electron density of the two molecules. The MRC results are reported in detail for the N-formyl-L-tryptophan complex at pH = 5.7 (52). They report difficulty in the interpretation of the active site region due to a movement in the Tyr' 146 residue and the hydroxyl group of Ser 195. In contrast to these results, the NFP: $\alpha$ -CHT difference map showed no movement in either Tyr' 146 or Ser 195. Since the MRC work was done at a higher pH it is possible that the movements are related to the pH dependent dissociation of the dimer and are similar to the shifts found in the pH study of Vandlen and Tulinsky (60). It is also possible that the shifts are due to the inclusion of a molecule of dioxane in the active site region during crystallization. Neither of these complications, higher pH or dioxane, are present in the NFP study. The other major

Table XVI. The Major Peaks in the NFP Difference Electron Density.

x, y, z are fractional coordinates

peak	pk.ht.eA <sup>-3</sup>	x	y	z
P, phenyl ring of NFP	0.18	0.6119	0.2700	0.3780
PX, portion of NFP molecule	0.13	0.6269	0.2800	0.4268
P', two-fold of phenyl ring NFP	0.14	0.6269	0.3300	0.5488
Uranyl position, Arg' 154 - Glu 21	0.30	0.0448	0.5000	0.5976 (pH)
movement of Val' 23 main chain	0.17	0.1493	0.0500	0.2927
movement of Val 23 side chain	0.15	0.0149	0.0800	0.3049
shift of His 40	0.21	0.8806	0.2700	0.4024 (pH)
	-0.15	0.8358	0.3700	0.6220 (pH)
movement of His' 40	0.16	0.8507	0.3300	0.6098 (pH)
Trp' 141 movement, main chain	0.15	0.7910	0.4000	0.6098 (pH)
Cys' 42 movement	-0.18	0.8060	0.2400	0.6220
shift of Lys 170	0.22	0.3284	0.4300	0.5732
	-0.16	0.3284	0.4500	0.5366
shift of Ser' 223 main chain	-0.16	0.4925	0.3900	0.5732
	0.13	0.4478	0.3800	0.5732
movement of Thr' 222	0.16	0.4179	0.4200	0.5122
shift near Ser 217	0.16	0.4925	0.2600	0.4024
	-0.13	0.4328	0.2350	0.4329
movement of Ser 159	0.16	0.5075	0.1000	0.2317 (pH)
shift of Thr 224 main chain by 0.2 Å	-0.16	0.4478	0.1600	0.2927
movement of Asn' 150	0.16	0.8358	0.4400	0.5000
shift of Leu 162 side chain by 0.34 Å	-0.15	0.2657	0.2100	0.1463
	0.15	0.2239	0.2300	0.1585
movement near Cys' 182 - Ile' 181	-0.15	0.3134	0.3000	0.5732
movement of Leu 46 side chain	0.15	0.8955	0.3000	0.1829
movement of Asp' 102 - Asn' 100	0.15	0.6119	0.1900	0.7073 (pH)

difference in the two studies is that the NFP study showed a lack of two-fold symmetry in the active site region (which would have been obscured in the average map of the MRC group) and in the other major changes in the enzyme structure which were not discussed in the MRC report.

This study has determined the position of the substitution of NFP in the active site region of  $\alpha$ -CHT and has uncovered that the substitution produces substrate-induced changes in the interactions of other residues, which are also observed when the pH is increased to a value near the pH of maximum activity of the enzyme. The specific chemical interactions which produce these changes may be understood more fully when the molecular details of the present structure of  $\alpha$ -CHT have been determined.

## REFERENCES

## References

1. J. E. Falk, "Porphyrins and Metalloporphyrins", Elsevier Publishing Company, Amsterdam, 1964.
2. G. S. Marks, "Heme and Chlorophyll", D. van Nostrand Company Ltd., London, 1969.
3. F. R. Longo, J. D. Finarelli, E. Schmalzbach, and A. D. Adler, J. Phys. Chem., 74, 3296 (1970).
4. A. D. Adler, J. H. Green, M. Mautner, Organic Mass Spectrometry, 3, 955 (1970).
5. R. Lemberg and J. E. Falk, Biochem. J., 49, 674 (1951).
6. M. Gouterman, J. Chem. Phys., 30, 1139 (1959) and J. Mol. Spectry., 6, 138 (1961).
7. W. S. Caughey and W. S. Koski, Biochemistry, 1, 923 (1962).
8. R. J. Abraham, A. H. Jackson, G. W. Kenner, and D. Warburton, J. Chem. Soc., 853 (1963).
9. G. M. Badger, R. L. N. Harris, R. A. Jones, and Jenneth M. Sasse, J. Chem. Soc., 4329 (1962).
10. S. F. Mason, J. Chem. Soc., 976 (1958).
11. E. D. Becker, R. B. Bradley, C. J. Watson, J. Am. Chem. Soc., 83, 3743 (1961).
12. L. E. Webb and E. B. Fleischer, J. Chem. Phys., 43, 3100 (1965).
13. S. Silvers and A. Tulinsky, J. Am. Chem. Soc., 89, 3331 (1967).
14. M. J. Hamor, T. A. Hamor, and J. L. Hoard, J. Am. Chem. Soc., 86, 1938 (1964).
15. B. M. L. Chen, Ph.D. Thesis, Michigan State University, East Lansing, Michigan (1970).
16. G. H. Stout and L. H. Jensen, "X-ray Structure Determination", The MacMillan Company, New York, 1968.

17. N. F. M. Henry, H. Lipson, and W. A. Wooster, "The Interpretation of X-ray Diffraction Photographs", MacMillan and Company, Ltd., London, 1961.
18. "International Tables for X-ray Crystallography", Vols. I-III, The Kynoch Press, Birmingham, England, 1952.
19. R. L. Vandlen and A. Tulinsky, Acta. Cryst., B27, 437 (1971).
20. H. W. Wyckoff, M. Doscher, D. Tsernoglou, T. Inagami, L. N. Johnson, K. D. Hardman, N. M. Allewell, D. M. Kelly, and F. M. Richards, J. Mol. Biol., 27, 563 (1967).
21. T. C. Furnas, "Single Crystal Orienter Instruction Manual", General Electric Company, Milwaukee (1957).
22. I. L. Karle, H. Hauptman, J. Karle, and A. B. Wing, Acta. Cryst., 11, 257 (1958).
23. A. J. C. Wilson, Nature, 150, 151 (1942).
24. D. Sayre, Acta. Cryst., 5, 60 (1952).
25. H. Hauptman and J. Karle, "Solution of the Phase Problem I. The Centrosymmetric Crystal", A.C.A. Monograph No. 3, Brooklyn: Polycrystal Book Service.
26. I. L. Karle and J. Karle, Acta. Cryst., 17, 1356 (1964).
27. W. Cochran and M. M. Woolfson, Acta. Cryst., 8, 1 (1955).
28. R. E. Long, Ph.D. Thesis, University of California, Los Angeles, California, 1965.
29. D. J. Cromer and J. T. Waber, Acta. Cryst., 18, 104 (1965).
30. R. F. Stewart, E. R. Davidson, and W. T. Simpson, J. Chem. Phys., 42, 3175 (1965).
31. E. W. Hughes, J. Am. Chem. Soc., 63, 1737 (1941).
32. W. R. Bushing, K. O. Martin, and H. A. Levy, "A Fortran Crystallographic Least-Squares Program", Report ORNL-TM-305, Oak Ridge, Tennessee (1962).
33. H. Lipson and W. Cochran, "The Determination of Crystal Structures", Cornell University Press, New York (1966), pp. 319-323.
34. R. W. James, "The Optical Principles of the Diffraction of X-rays", G. Bell and Sons, Ltd., London, 1950, pp. 292-294.
35. D. M. Collins and J. L. Hoard, J. Am. Chem. Soc., 92, 3761 (1970).

36. D. P. Shoemaker, J. Donohue, V. Schoemaker, and R. B. Corey, J. Am. Chem. Soc., 72, 2328 (1950).
37. B. Dawson, Acta. Cryst., 14, 999 (1961).
38. C. K. Johnson, "ORTEP, a Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations", ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1965.
39. V. Schomaker, J. Waber, R. E. Marsh, and G. Bergman, Acta. Cryst., 12, 600 (1959).
40. J. R. Platt, in "Radiation Biology", Vol. 3, A. Hollaender, Ed., McGraw Hill, New York, 1956, p. 101.
41. W. C. Hamilton and J. A. Ibers, "Hydrogen Bonding in Solids", W. A. Benjamin, Inc., New York, 1968, pp. 262-263.
42. M. Röver, M. Poilroux, A. Yoshida, and P. Desnuelle, Biochem. Biophys. Acta., 23, 608 (1957).
43. N. Brenner, H. R. Muller, and R. W. Pfister, Helv. Chim. Acta., 33, 568 (1950).
44. G. E. Hein and C. Niemann, J. Am. Chem. Soc., 84, 4495 (1962).
45. S. G. Cohen, A. Milovanoic, R. M. Schultz, and S. Y. Weinstein, J. Biol. Chem., 244, 2664 (1969).
46. B. S. Hartley and B. A. Kilby, Biochem. J., 50, 672 (1952).
47. G. Schoellman and E. Shaw, Biochemistry, 2, 252 (1963).
48. G. P. Hess in "The Enzymes", Vol. III, P. D. Boyer, Ed., Academic Press, New York, 1971, pp. 236-244.
49. G. P. Hess, J. McConn, E. Ku, and G. McConkey, Phil. Trans. Roy. Soc. Lond., B257, 89 (1970).
50. M. L. Bender and K. C. Kemp, J. Am. Chem. Soc., 79, 116 (1957).
51. A. Tulinsky, P. W. Coddington, and R. L. Vandlen, Abstracts, Winter Meeting of the American Crystallographic Association, Columbia, South Carolina, Feb. 1971, No. K3.
52. T. A. Steitz, R. Henderson, and D. M. Blow, J. Mol. Biol., 46, 337 (1969).
53. A. Tulinsky, N. V. Mani, C. N. Morimoto, and R. L. Vandlen, unpublished results.
54. R. L. Vandlen, N. V. Mani, C. N. Morimoto, and A. Tulinsky, Abstracts, Winter Meeting of the American Crystallographic Association, Columbia, South Carolina, Feb. 1971, No. K2.

55. R. E. Dickerson, J. C. Kendrew, and B. E. Strandberg, Acta. Cryst., 14, 1188 (1961).
56. R. Henderson and J. K. Moffat, unpublished results.
57. P. B. Sigler, D. M. Blow, B. W. Matthews, and R. Henderson, J. Mol. Biol., 35, 143 (1968).
58. R. L. Vandlen, private communication.
59. A. C. T. North, D. C. Phillips, and F. S. Mathews, Acta. Cryst., A24, 351 (1968).
60. R. L. Vandlen and A. Tulinsky, Abstracts, 62nd Annual Meeting of American Society of Biological Chemists, San Francisco, California, June 1971, No. 444.
61. B. S. Hartley, Phil. Trans. Roy. Soc. Lond., B257, 77 (1970).
62. K. C. Aune and S. N. Timasheff, Biochemistry, 10, 1609 (1971).



## APPENDICES

## APPENDIX I

### The Reciprocal Lattice

The reciprocal lattice is a three dimensional array of points around the crystal in which each point represents in orientation and spacing a direct lattice plane (hkl). The vector from the origin to the reciprocal lattice point represents the normal to the direct lattice plane and the length of that vector is equal to  $1/d_{(hkl)}$  where  $d_{(hkl)}$  is the spacing of the lattice planes.  $d_{(hkl)}$  is related to  $\theta$ , the angle of the incident and diffracted beams to the direct lattice plane (hkl) by

$$2d_{(hkl)} \sin\theta = \lambda$$

where  $\lambda$  is the wavelength of the radiation. The unit lengths  $\underline{a^*}$ ,  $\underline{b^*}$ ,  $\underline{c^*}$  and angles  $\underline{\alpha^*}$ ,  $\underline{\beta^*}$ ,  $\underline{\gamma^*}$  of the reciprocal lattice are related to the unit cell parameters by

$$\begin{matrix} \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \underline{a^*} \cdot \underline{b} & = & \underline{a^*} \cdot \underline{c} & = & \underline{b^*} \cdot \underline{c} & = & \underline{b^*} \cdot \underline{a} & = & \underline{c^*} \cdot \underline{a} & = & \underline{c^*} \cdot \underline{b} & = & 0 \end{matrix}$$

and

$$\begin{matrix} \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \underline{a^*} \cdot \underline{a} & = & \underline{b^*} \cdot \underline{b} & = & \underline{c^*} \cdot \underline{c} & = & 1 \end{matrix}$$

**APPENDIX II**  
**The Observed and Calculated Structure**  
**Factors of Tetrapropylporphine**

H	K	L	F (ORS)	F (ALC)	H	K	L	F (ORS)	F (ALC)	H	K	L	F (ORS)	F (ALC)	H	K	L	F (ORS)	F (ALC)
0	2	0	334.0	-732.4	0	0	0	-1	32.6	3	0	-1	24.0	-23.6	1	9	-2	16.3	-33.4
0	0	0	51.4	-52.4	0	0	0	-1	42.9	3	0	-1	24.0	-43.7	1	10	-2	30.4	-33.1
0	6	0	44.8	63.0	0	10	-1	34.4	35.3	3	9	-1	10.8	-9.4	1	11	2	7.4	-7.3
0	8	0	36.8	32.0	1	1	-1	209.0	220.3	3	9	-1	6.2	8.3	1	11	2	30.6	-32.7
0	12	0	22.4	27.3	1	1	-1	18.4	21.6	3	10	-1	14.4	15.9	2	0	-2	148.1	-154.2
1	1	0	83.9	-49.1	1	2	-1	241.6	254.7	4	1	-1	25.9	-21.0	2	0	-2	138.4	-141.0
1	1	0	18.6	-21.4	1	2	-1	163.7	-164.3	4	1	-1	7.0	-4.6	2	1	-2	68.2	67.7
1	1	0	138.4	-140.4	1	3	-1	60.7	61.4	4	2	-1	9.7	-8.9	2	1	-2	7.7	-9.4
1	3	0	89.8	-95.8	1	3	-1	44.4	47.4	4	2	-1	15.4	15.0	2	2	2	34.0	-39.8
1	5	0	77.7	-76.7	1	4	-1	24.4	22.6	4	3	-1	32.6	-32.6	2	2	-2	35.4	29.3
1	6	0	55.4	53.3	1	4	-1	102.2	99.0	4	3	-1	27.2	-22.7	2	3	-2	22.7	17.9
1	7	0	16.6	14.7	1	5	-1	71.7	-67.4	4	4	-1	27.6	-21.0	2	3	-2	37.8	-36.2
1	8	0	9.8	-7.9	1	6	-1	15.4	-18.2	4	5	-1	27.1	28.2	2	4	-2	85.7	80.2
1	9	0	16.6	-17.5	1	6	-1	101.9	-97.2	4	5	-1	7.1	-4.3	2	4	-2	71.4	-71.8
1	10	0	41.7	-42.1	1	7	-1	6.2	6.9	4	7	-1	22.9	-23.6	2	5	-2	14.5	-13.7
2	0	0	26.3	23.4	1	8	-1	57.6	56.2	4	7	-1	11.9	-10.7	2	5	-2	55.5	-53.3
2	1	0	15.0	-15.4	1	9	-1	29.3	-26.2	4	8	-1	6.1	5.7	2	6	-2	46.9	-46.9
2	2	0	32.4	-39.1	1	10	-1	52.9	-54.2	5	1	-1	7.0	-6.2	2	6	-2	21.9	-18.7
2	3	0	20.4	18.1	1	10	-1	25.5	23.9	5	1	-1	6.0	-5.1	2	7	-2	55.9	-55.9
2	4	0	44.6	43.1	1	11	-1	19.4	23.2	5	2	-1	20.9	20.5	2	8	-2	34.2	-32.7
2	5	0	29.1	-29.4	1	11	-1	4.7	-8.5	5	3	-1	9.4	-9.4	2	8	-2	29.9	28.4
2	6	0	49.4	-44.1	1	12	-1	13.7	-16.7	5	3	-1	7.3	-7.1	2	9	-2	20.3	22.4
2	7	0	24.2	-31.1	2	1	-1	98.3	-95.7	5	4	-1	22.8	23.6	2	10	-2	12.1	-14.5
2	8	0	31.4	31.4	2	1	-1	42.7	67.3	5	4	-1	382.0	377.2	2	10	-2	6.7	-8.5
2	9	0	10.8	11.4	2	2	-1	36.7	39.2	0	1	-2	90.5	-91.1	2	11	2	16.4	-19.1
3	0	0	30.1	-22.4	2	2	-1	32.1	-31.7	0	2	-2	206.0	-202.4	3	0	2	26.8	-28.3
3	1	0	30.1	32.4	2	3	-1	74.6	74.7	0	3	-2	29.3	-28.3	3	1	-2	44.7	47.2
3	2	0	30.4	-34.1	2	4	-1	58.4	62.6	0	4	-2	104.8	109.5	3	2	-2	24.7	-24.9
3	3	0	32.9	31.2	2	5	-1	33.7	-24.5	0	4	-2	57.0	56.9	3	3	-2	35.9	35.0
3	4	0	7.4	7.0	2	5	-1	48.3	84.2	0	7	-2	24.5	24.8	3	4	-2	72.8	70.4
3	5	0	31.6	-39.1	2	6	-1	39.2	34.4	0	8	-2	26.5	24.2	3	5	-2	9.2	-11.7
3	6	0	45.5	63.4	2	6	-1	25.9	25.4	0	9	-2	13.4	-12.6	3	5	-2	17.5	-16.6
3	7	0	11.0	9.2	2	7	-1	34.6	-35.4	0	10	-2	55.5	-57.1	3	6	-2	30.5	-30.0
3	8	0	7.3	-8.4	2	7	-1	34.6	-41.7	0	11	-2	11.9	-13.2	3	6	-2	19.8	-21.0
3	9	0	24.6	21.4	2	8	-1	34.6	-41.7	0	12	-2	13.3	-14.8	3	7	-2	27.4	24.4
4	0	0	21.5	21.5	2	8	-1	11.7	-12.8	1	0	-2	200.7	208.0	3	7	-2	33.1	-31.2
4	1	0	31.7	-30.6	2	9	-1	28.1	32.3	1	0	-2	234.0	-246.3	3	8	-2	44.9	47.1
4	2	0	23.9	-22.2	2	9	-1	32.1	32.4	1	1	-2	609.0	598.4	3	9	-2	7.1	-6.3
4	3	0	24.4	23.4	2	10	-1	37.4	37.4	1	1	-2	109.2	-115.6	3	10	-2	27.4	-27.4
4	4	0	16.7	15.7	2	10	-1	37.4	12.6	1	2	-2	37.3	-41.2	4	0	-2	61.1	-61.9
4	5	0	6.2	-4.0	2	11	-1	40.4	-23.5	1	2	-2	13.7	-16.8	4	1	-2	14.4	10.8
4	6	0	9.6	11.2	3	1	-1	40.4	43.3	1	3	-2	51.3	-56.0	4	1	-2	45.4	45.9
4	7	0	7.0	6.1	3	1	-1	54.4	-52.5	1	3	-2	166.5	171.0	4	2	-2	34.5	34.0
4	8	0	11.5	13.2	3	2	-1	44.7	-54.4	1	4	-2	10.3	25.4	4	2	-2	13.1	-14.5
4	9	0	32.5	-33.6	3	3	-1	44.7	-54.4	1	5	-2	18.6	-20.0	4	3	-2	6.4	3.4
5	0	0	215.0	-215.2	3	4	-1	18.4	17.2	1	6	-2	26.9	-29.3	4	4	-2	27.3	26.0
5	1	0	22.7	-21.7	3	4	-1	15.4	-14.9	1	6	-2	31.5	-31.1	4	5	-2	25.5	-24.0
5	2	0	157.9	-162.2	3	5	-1	10.7	9.7	1	7	-2	49.4	-45.4	4	6	-2	14.8	-14.7
5	3	0	32.5	-33.4	3	5	-1	22.7	21.7	1	7	-2	44.7	-45.4	4	6	-2	14.8	-14.7
5	4	0	90.2	-92.2	3	6	-1	28.5	28.1	1	8	-2	17.0	-15.1	4	7	-2	6.2	5.7
5	5	0	44.6	-42.4	3	6	-1	34.2	-35.9	1	8	-2	52.2	51.1	4	8	-2	22.9	-21.4
5	6	0	64.5	64.7	3	7	-1	34.2	-35.9	1	9	-2	11.2	9.1	5	0	-2		

M	L	F (ORS)	F (CALC)	M	L	F (ORS)	F (CALC)	M	L	F (ORS)	F (CALC)	M	L	F (ORS)	F (CALC)	M	L	F (ORS)	F (CALC)
5	0	22.9	-21.6	2	11	-3	13.3	1	4	17.0	-11.4	4	6	31.0	-31.9	2	10	10.0	-10.0
5	0	12.3	11.9	1	3	17.1	-13.5	2	5	121.2	-11.4	4	6	28.0	-31.9	2	10	10.0	-10.0
5	0	5.0	5.4	3	1	19.8	-19.3	1	5	14.6	11.4	4	7	7.2	6.2	4	7	7.2	6.2
5	2	19.2	18.0	3	2	35.2	-37.3	1	7	38.6	39.6	5	0	31.1	-31.9	5	0	31.1	-31.9
5	2	11.0	10.5	2	3	51.6	-51.3	1	7	25.4	-25.1	5	0	18.4	-20.9	5	0	18.4	-20.9
5	3	15.1	-16.7	3	3	75.1	-69.4	1	8	24.4	20.5	5	1	14.0	-14.0	5	1	14.0	-14.0
5	3	13.2	-14.0	3	3	8.4	8.5	1	9	11.7	-10.9	5	1	21.9	-25.6	5	1	21.9	-25.6
5	4	24.6	-26.0	3	4	39.3	40.2	1	9	53.8	56.5	5	2	22.3	22.9	5	2	22.3	22.9
0	1	54.2	-57.9	3	4	10.3	13.2	1	10	24.3	22.2	5	3	9.1	-10.6	5	3	9.1	-10.6
0	2	46.4	47.2	3	5	32.0	31.7	1	11	17.4	18.1	5	4	12.6	13.4	5	4	12.6	13.4
0	3	204.5	211.4	3	5	6.1	6.5	1	11	27.7	-30.7	0	1	119.0	129.4	0	1	119.0	129.4
0	4	24.4	24.5	3	6	14.4	12.6	2	0	137.8	-143.9	0	2	5 218.5	-210.4	0	2	5 218.5	-210.4
0	5	159.4	-154.9	3	6	26.9	-25.9	2	0	65.1	59.9	0	3	5 42.6	-39.1	0	3	5 42.6	-39.1
0	6	26.4	-23.4	3	7	37.4	38.6	2	1	83.7	84.0	0	4	5 11.9	-13.2	0	4	5 11.9	-13.2
0	7	54.7	50.1	3	7	31.1	-30.8	2	2	36.9	31.1	0	5	5 15.4	-17.0	0	5	5 15.4	-17.0
0	8	40.4	-40.0	3	8	37.7	-34.3	2	2	16.0	-15.0	0	6	5 40.4	76.4	0	6	5 40.4	76.4
0	9	21.4	24.1	3	8	7.1	-6.4	2	3	115.1	-115.3	0	7	5 20.7	19.5	0	7	5 20.7	19.5
0	10	40.2	41.0	3	9	16.2	-15.4	2	3	76.8	75.3	0	8	5 53.9	54.6	0	8	5 53.9	54.6
0	11	24.4	-21.0	3	9	41.8	-42.5	2	4	130.4	-129.1	0	9	5 20.2	21.2	0	9	5 20.2	21.2
0	12	12.2	14.5	3	10	3.4	6.4	2	4	45.3	43.0	0	10	5 8.7	10.0	0	10	5 8.7	10.0
1	1	123.0	-126.0	4	1	14.0	20.6	2	5	79.7	-78.2	0	11	5 7.3	-4.7	0	11	5 7.3	-4.7
1	1	91.2	102.4	4	2	56.7	52.2	2	6	31.4	-29.0	1	1	5 25.1	-18.5	1	1	5 25.1	-18.5
1	2	34.9	-31.4	4	2	16.6	-15.4	2	7	90.3	89.0	1	1	5 211.4	-214.2	1	1	5 211.4	-214.2
1	2	142.4	141.5	4	3	22.6	22.1	2	7	24.4	24.6	1	2	5 155.2	-159.9	1	2	5 155.2	-159.9
1	3	11.1	9.1	4	3	4.1	-7.2	2	8	20.1	-21.4	1	3	5 41.7	42.4	1	3	5 41.7	42.4
1	4	25.4	21.4	4	3	18.0	-17.3	2	8	16.5	15.3	1	3	5 15.2	16.5	1	3	5 15.2	16.5
1	5	35.4	36.6	4	5	9.4	-7.9	2	9	4.2	-7.0	1	4	5 30.4	-31.7	1	4	5 30.4	-31.7
1	5	24.2	-27.5	4	6	3.4	-14.5	2	10	4.6	-7.0	1	5	5 100.2	-95.2	1	5	5 100.2	-95.2
1	6	101.2	-92.4	4	6	3.4	-4.9	2	11	21.9	-25.7	1	6	5 165.0	157.4	1	6	5 165.0	157.4
1	7	10.1	10.2	4	7	11.6	12.1	3	0	149.1	145.4	1	7	5 43.4	-40.7	1	7	5 43.4	-40.7
1	7	22.1	-19.9	4	8	14.0	15.5	3	1	44.5	-46.2	1	8	5 44.1	-45.7	1	8	5 44.1	-45.7
1	7	25.4	-25.5	5	1	7.5	4.2	3	1	45.3	43.4	1	9	5 32.7	36.4	1	9	5 32.7	36.4
1	8	34.0	34.4	5	1	17.1	-15.8	3	2	35.1	-34.2	1	9	5 17.4	-18.7	1	9	5 17.4	-18.7
1	8	34.0	32.1	5	2	35.6	-36.6	3	2	17.6	-17.1	1	10	5 12.2	-13.1	1	10	5 12.2	-13.1
1	9	47.1	-47.1	5	3	7.6	-8.2	3	3	51.4	-53.2	1	10	5 31.4	30.0	1	10	5 31.4	30.0
1	9	20.4	-17.9	5	4	136.1	-147.6	3	4	61.9	-58.6	1	11	5 17.4	16.5	1	11	5 17.4	16.5
2	1	11.4	34.5	6	1	42.8	-40.6	3	5	11.8	9.5	1	11	5 30.4	32.4	1	11	5 30.4	32.4
2	1	44.0	50.4	6	2	72.5	-72.8	3	5	4.0	4.2	1	12	5 12.2	-13.1	1	12	5 12.2	-13.1
2	2	5.4	5.9	6	3	48.6	-48.3	3	6	20.8	19.0	1	13	5 43.2	-44.3	1	13	5 43.2	-44.3
2	3	43.4	45.1	6	4	4.4	-8.3	3	7	31.7	-29.3	1	14	5 27.2	-28.4	1	14	5 27.2	-28.4
2	3	119.3	121.1	6	5	33.7	28.7	3	7	24.2	24.4	1	15	5 27.1	-27.7	1	15	5 27.1	-27.7
2	4	58.8	57.4	6	6	52.4	50.1	3	8	36.4	-39.0	2	1	5 27.1	-27.7	2	1	5 27.1	-27.7
2	4	34.7	-36.0	6	7	47.2	-46.5	3	8	37.4	35.0	2	2	5 9.5	-9.1	2	2	5 9.5	-9.1
2	5	21.3	-20.0	6	8	25.5	-26.1	3	9	8.0	8.2	2	3	5 19.3	19.4	2	3	5 19.3	19.4
2	6	13.3	13.4	6	9	17.4	-18.7	3	10	4.7	4.1	2	4	5 20.2	20.2	2	4	5 20.2	20.2
2	6	17.4	14.4	6	10	181.0	181.4	4	0	4.6	-8.3	2	5	5 51.5	50.2	2	5	5 51.5	50.2
2	7	9.7	12.3	6	11	527.0	526.6	4	0	21.2	-21.6	2	6	5 27.2	27.7	2	6	5 27.2	27.7
2	7	28.8	-25.5	6	12	198.0	201.7	4	2	19.5	-18.1	2	7	5 72.2	74.0	2	7	5 72.2	74.0
2	8	45.8	-45.1	6	13	44.0	437.1	4	3	25.1	-26.8	2	8	5 27.3	27.3	2	8	5 27.3	27.3
2	9	16.7	16.4	6	14	66.5	-71.6	4	4	30.4	-28.7	2	9	5 27.3	27.3	2	9	5 27.3	27.3
2	9	66.0	67.1	6	15	137.6	139.1	4	5	40.0	49.5	2	10	5 70.6	70.6	2	10	5 70.6	70.6
2	10	21.8	24.0	6	16	13.9	-16.5	4	5	17.5	-25.6	2	11	5 60.7	-59.3	2	11	5 60.7	-59.3
2	10	4.0	-9.3	6	17	41.3	-46.3	4	6	26.1	-25.6	2	12	5 24.2	-22.7	2	12	5 24.2	-22.7
3	11	8.5	9.4	6	18	45.4	-46.3	4	7	19.5	19.4	2	13	5 24.2	-22.7	2	13	5 24.2	-22.7

M	K	L	F (OHS)	F (CALC)	M	K	L	F (OHS)	F (CALC)	M	K	L	F (OHS)	F (CALC)	M	K	L	F (OHS)	F (CALC)
2	7	-5	24.2	-22.7	4	1	-6	50.7	-50.4	4	1	-6	15.2	15.1	2	7	-7	48.3	-45.4
2	8	-5	6.8	-5.9	4	2	-6	142.6	140.1	4	2	-6	19.8	-20.8	2	7	-7	28.7	-21.7
2	9	-5	28.4	-20.9	4	3	-6	143.7	-140.7	4	3	-6	30.5	-28.5	2	8	-7	11.4	-11.6
2	10	-5	19.9	-17.4	4	4	-6	9.2	11.6	4	4	-6	33.9	13.1	2	9	-7	41.7	-43.0
2	11	-5	33.6	13.8	4	5	-6	26.3	-25.0	4	5	-6	6.2	5.2	2	10	-7	24.2	-24.3
2	12	-5	5.7	7.3	4	6	-6	22.8	20.7	4	6	-6	54.0	53.7	2	11	-7	38.7	39.8
3	1	-5	19.7	19.4	4	7	-6	70.0	66.8	4	7	-6	25.8	-25.8	2	12	-7	8.2	-9.2
3	2	-5	15.9	14.9	4	8	-6	20.0	18.1	4	8	-6	23.3	-24.8	3	1	-7	7.6	7.7
3	3	-5	9.4	-29.1	4	9	-6	35.3	-33.9	4	9	-6	18.8	-20.8	3	2	-7	41.8	40.0
3	4	-5	15.7	-77.4	4	10	-6	42.1	-41.6	4	10	-6	46.4	-47.1	3	3	-7	72.8	72.2
3	5	-5	10.4	28.1	4	11	-6	43.6	44.1	4	11	-6	6.5	-7.8	3	4	-7	20.4	20.8
3	6	-5	9.9	12.5	4	12	-6	32.0	32.5	4	12	-6	24.8	26.9	3	5	-7	21.3	21.9
3	7	-5	26.5	-25.8	4	13	-6	14.3	-16.3	4	13	-6	11.0	-11.3	3	6	-7	58.4	53.4
3	8	-5	6.9	-8.0	4	14	-6	28.3	-31.7	4	14	-6	52.6	59.6	3	7	-7	76.0	-74.3
3	9	-5	28.6	23.2	4	15	-6	8.0	11.4	4	15	-6	43.3	41.4	3	8	-7	41.1	42.1
3	10	-5	10.4	11.9	4	16	-6	78.3	-82.7	4	16	-6	56.5	-51.2	3	9	-7	48.5	-4.0
3	11	-5	12.1	13.3	4	17	-6	138.4	132.6	4	17	-6	97.1	94.3	3	10	-7	5.8	-4.0
3	12	-5	5.6	-5.7	4	18	-6	60.1	-61.5	4	18	-6	63.0	64.2	3	11	-7	14.3	-14.0
3	13	-5	6.5	-6.8	4	19	-6	155.6	156.3	4	19	-6	17.5	17.2	3	12	-7	9.2	-10.0
3	14	-5	48.6	-49.2	4	20	-6	139.9	137.3	4	20	-6	30.4	-29.9	4	1	-7	8.4	9.4
4	1	-5	16.7	-17.4	4	21	-6	42.7	-46.5	4	21	-6	21.9	-24.3	4	2	-7	40.0	-42.3
4	2	-5	42.1	41.4	4	22	-6	8.7	12.6	4	22	-6	36.7	42.0	4	3	-7	9.4	8.4
4	3	-5	25.3	-25.7	4	23	-6	188.1	184.6	4	23	-6	78.7	-76.3	4	4	-7	10.6	-10.0
4	4	-5	5.5	7.5	4	24	-6	39.0	-33.8	4	24	-6	29.6	29.3	4	5	-7	5.8	-4.7
4	5	-5	31.7	-31.4	4	25	-6	4.4	1.7	4	25	-6	41.6	-40.6	4	6	-7	11.2	-8.0
4	6	-5	12.4	-11.4	4	26	-6	19.1	-14.0	4	26	-6	113.5	115.7	4	7	-7	12.0	-10.3
4	7	-5	29.5	29.4	4	27	-6	69.6	66.2	4	27	-6	149.9	152.1	4	8	-7	10.4	-12.0
4	8	-5	10.6	-11.9	4	28	-6	13.2	11.3	4	28	-6	21.8	23.4	4	9	-7	11.9	14.3
4	9	-5	27.2	27.5	4	29	-6	8.0	4.0	4	29	-6	59.1	57.1	4	10	-7	10.4	11.5
4	10	-5	15.6	-9.9	4	30	-6	16.0	-17.7	4	30	-6	47.3	-42.9	4	11	-7	5.9	-5.1
4	11	-5	17.0	16.1	4	31	-6	11.9	-11.7	4	31	-6	24.2	-21.9	4	12	-7	37.5	32.9
4	12	-5	27.5	32.3	4	32	-6	34.1	35.9	4	32	-6	65.5	-63.0	4	13	-7	11.7	13.8
4	13	-5	14.7	14.4	4	33	-6	5.5	4.8	4	33	-6	17.1	15.2	4	14	-7	19.5	9.7
4	14	-5	13.8	-14.8	4	34	-6	40.9	36.4	4	34	-6	11.5	-9.9	4	15	-7	26.3	-27.4
4	15	-5	21.6	-20.5	4	35	-6	21.9	-25.0	4	35	-6	8.2	3.6	4	16	-7	95.5	-92.2
4	16	-5	27.9	-29.4	4	36	-6	9.9	12.6	4	36	-6	18.4	-16.8	4	17	-7	4.2	4.9
4	17	-5	102.7	-109.0	4	37	-6	20.4	-16.7	4	37	-6	67.3	66.7	4	18	-7	22.3	21.9
4	18	-5	56.6	56.9	4	38	-6	74.5	-72.9	4	38	-6	42.8	45.0	4	19	-7	10.4	-6.2
4	19	-5	129.0	-128.4	4	39	-6	25.4	-24.1	4	39	-6	11.8	-13.7	4	20	-7	16.3	-15.4
4	20	-5	56.1	-52.1	4	40	-6	19.1	14.1	4	40	-6	22.1	-25.3	4	21	-7	10.2	10.7
4	21	-5	126.6	118.5	4	41	-6	9.1	-11.3	4	41	-6	12.3	-12.9	4	22	-7	31.8	16.0
4	22	-5	43.3	-39.4	4	42	-6	25.2	-25.0	4	42	-6	13.5	-14.3	4	23	-7	120.2	-124.4
4	23	-5	44.5	-44.2	4	43	-6	11.7	10.4	4	43	-6	90.6	-129.5	4	24	-7	52.6	49.1
4	24	-5	46.9	-42.3	4	44	-6	17.1	-17.1	4	44	-6	154.0	-159.2	4	25	-7	8.3	7.7
4	25	-5	10.0	7.9	4	45	-6	12.3	11.1	4	45	-6	149.2	-144.1	4	26	-7	73.4	75.5
4	26	-5	11.4	10.8	4	46	-6	6.3	-4.4	4	46	-6	37.0	-37.9	4	27	-7	11.8	14.3
4	27	-5	37.3	34.5	4	47	-6	13.3	-13.4	4	47	-6	40.5	-39.7	4	28	-7	24.8	-27.1
4	28	-5	129.7	-133.8	4	48	-6	14.4	19.4	4	48	-6	80.3	79.8	4	29	-7	8.7	-11.6
4	29	-5	329.0	317.8	4	49	-6	14.4	-14.6	4	49	-6	60.6	43.5	4	30	-7	146.1	-160.5
4	30	-5	34.0	33.3	4	50	-6	4.6	8.0	4	50	-6	54.2	-53.6	4	31	-7	74.6	71.4
4	31	-5	40.5	36.9	4	51	-6	23.5	22.7	4	51	-6	28.6	29.0	4	32	-7	62.0	-59.5
4	32	-5	92.3	-93.4	4	52	-6	4.5	11.7	4	52	-6	43.2	-38.5	4	33	-7	13.8	11.5
4	33	-5	89.0	-86.2	4	53	-6	21.1	20.7	4	53	-6	15.8	-16.4	4	34	-7	10.5	-10.7

M	K	L	F (OBS)	F (CALC)	H	K	L	F (OBS)	F (CALC)	H	K	L	F (OBS)	F (CALC)
1	7	-A	10.5	-10.7	3	6	-9	45.3	47.3	2	4	10	30.3	26.4
1	7	-A	44.6	41.5	3	7	-9	33.5	35.5	2	5	10	5.8	4.0
1	8	-A	41.6	-39.7	3	7	9	14.8	-15.1	2	5	-10	72.9	-69.5
1	8	-A	15.0	-16.4	3	8	-9	19.8	15.9	2	6	-10	14.1	12.1
1	9	-A	20.8	18.2	3	8	9	9.9	-11.7	2	6	10	30.1	28.1
1	9	-A	28.7	31.4	4	1	-9	14.3	15.7	2	7	10	7.5	7.0
2	0	-B	136.4	132.4	4	2	-9	11.4	10.1	2	7	10	24.6	27.5
2	0	-B	129.5	-130.8	4	2	9	14.3	15.7	2	8	10	47.2	-48.6
2	1	-A	30.4	28.4	4	4	-9	A.3	7.8	2	8	10	47.2	-48.6
2	1	-A	72.7	-67.3	4	4	9	24.5	-23.7	2	9	-10	13.0	5.4
2	2	-A	42.4	38.1	4	5	-9	14.9	14.7	2	9	-10	13.0	5.4
2	2	-A	226.1	227.0	4	5	9	51.4	50.5	2	9	10	8.0	-7.0
2	3	-A	38.1	-35.5	4	5	9	23.5	-24.0	3	0	10	17.4	14.2
2	3	-A	6.2	6.2	4	6	-9	10.8	-12.0	3	0	-10	98.1	102.8
2	4	-A	50.5	43.9	5	2	-9	16.8	19.1	3	1	10	44.4	-44.4
2	4	-A	45.0	41.3	5	3	9	17.2	-14.1	3	2	-10	42.9	-42.1
2	5	-A	23.8	-19.8	0	0	-10	17.8	-14.1	3	2	-10	39.9	37.2
2	5	-A	12.2	-10.6	0	1	-10	18.8	-102.7	3	3	-10	64.7	65.0
2	6	-A	16.2	15.5	0	2	-10	19.2	-80.5	3	3	10	46.8	46.1
2	7	-A	46.3	84.4	0	3	-10	79.9	-79.8	3	4	-10	23.1	-23.6
2	7	-A	13.6	-13.2	0	4	-10	7.7	-10.2	3	4	-10	26.7	-24.5
2	7	-A	61.3	61.0	0	5	-10	75.7	71.4	3	5	-10	59.9	-60.3
2	8	-A	6.2	-7.6	0	6	-10	12.3	-8.3	3	6	-10	47.2	47.1
2	9	-A	10.8	-12.4	0	7	-9	19.5	32.8	3	6	10	14.1	13.0
2	9	-A	13.1	9.2	0	8	-9	14.7	19.3	3	7	-10	13.4	-12.7
2	10	-A	20.8	23.1	0	10	-10	30.6	25.5	3	8	-10	25.2	-29.4
3	0	-A	51.2	-53.5	1	0	-10	39.5	35.5	4	9	-10	13.9	-20.5
3	0	-A	6.1	-64.7	1	1	-10	9.6	10.5	4	0	-10	6.8	4.4
3	1	-A	18.5	-37.7	1	1	-10	45.2	-43.5	4	0	-10	62.1	-63.5
3	1	-A	18.9	-37.4	1	2	-10	40.0	-41.5	4	1	10	11.2	-12.1
3	2	-A	20.2	16.9	1	2	-10	45.1	-42.6	4	1	-10	54.5	53.2
3	3	-A	43.2	-41.4	1	3	-10	55.0	52.0	4	2	10	12.0	11.2
3	3	-A	11.6	9.9	1	4	-10	46.0	-44.8	4	3	-10	10.0	10.4
3	4	-A	18.7	-18.5	1	4	-10	17.8	11.4	4	4	-10	49.7	-52.4
3	4	-A	18.3	-18.4	1	5	-10	16.7	16.7	4	4	-10	20.7	-22.4
3	5	-A	12.4	-11.7	1	5	-10	24.8	-31.7	4	5	-10	18.4	-20.2
3	6	-A	9.4	-11.4	1	6	-10	120.6	112.1	4	6	-10	20.7	23.9
3	7	-A	10.9	10.9	1	7	-10	41.4	39.2	5	0	-10	24.1	-24.1
3	8	-A	17.7	18.1	1	7	-10	40.9	59.4	5	1	-10	24.1	-24.1
3	9	-A	11.9	-14.7	1	8	-10	39.1	-41.9	5	2	-10	17.3	12.4
4	0	-A	16.2	-15.5	1	9	-10	37.7	-30.9	5	3	-10	13.4	13.0
4	0	-A	10.7	7.4	1	10	-10	32.1	-10.9	0	3	-11	26.2	-25.9
4	1	-A	72.4	75.5	1	10	-10	19.2	20.1	0	4	-11	81.4	74.4
4	2	-A	15.3	15.0	1	11	-10	6.1	-6.2	0	5	-11	20.6	15.7
4	3	-A	20.6	22.2	2	0	-10	48.2	-50.9	0	6	-11	11.7	-10.2
4	4	-A	11.5	-12.4	2	0	-10	103.9	-91.3	0	7	-11	58.7	-60.2
4	5	-A	16.0	-34.3	2	1	10	41.5	42.8	0	8	-11	28.0	21.7
4	6	-A	11.3	-11.9	2	2	-10	22.0	-19.6	0	9	-11	28.0	20.7
4	7	-A	21.7	-23.0	2	2	-10	11.9	-16.7	1	1	-11	140.4	-131.1
4	8	-A	22.0	25.0	2	3	-10	70.8	69.8	1	1	-11	25.5	-24.4
5	0	-A	50.3	-50.3	2	3	-10	15.2	-17.2	1	2	-11	10.9	4.9
5	1	-A	19.9	-42.2	2	3	-10	7.9	-7.1	1	3	-11	9.4	-7.5
5	2	-A	11.4	12.5	2	4	-10	27.1	25.7	1	3	-11	60.7	54.4

M	K	L	F (ORS)	F (CALC)	M	K	L	F (ORS)	F (CALC)	M	K	L	F (ORS)	F (CALC)	M	K	L	F (ORS)	F (CALC)
1	3	11	60.7	59.9	1	0	-12	12.3	13.7	4	4	-12	27.7	-29.0	3	8	-13	17.0	-19.6
1	4	11	49.4	-45.8	1	1	-12	38.3	-38.1	4	1	-13	12.2	8.9	4	1	-13	12.2	12.7
1	4	-11	45.8	-41.9	1	1	12	14.4	12.1	0	1	-13	64.0	-60.0	4	1	-13	21.6	-23.2
1	5	-11	36.4	-34.9	1	2	12	33.4	29.2	0	2	-13	64.7	59.1	4	4	-13	11.9	11.5
1	5	11	31.8	32.8	1	2	-12	106.2	97.9	0	3	-13	24.0	26.4	4	5	-13	15.3	-14.2
1	6	11	18.5	9.2	1	3	12	16.5	15.7	0	4	-13	95.9	-90.8	4	6	-13	16.9	-17.3
1	7	-11	39.2	-34.4	1	4	-12	17.4	-18.9	0	5	-13	8.9	-4.9	4	6	-13	71.2	66.3
1	8	-11	18.3	-19.0	1	5	-12	31.4	-29.2	0	6	-13	11.2	-8.2	0	1	-14	66.6	64.0
1	9	-11	40.6	-42.0	1	5	12	21.4	-24.1	0	7	-13	10.3	-11.8	0	2	-14	34.9	29.7
1	10	-11	6.5	9.7	1	6	-12	30.4	-28.8	0	8	-13	21.0	23.6	0	3	-14	42.7	-41.7
2	1	11	14.3	14.7	1	6	-12	66.2	62.5	0	9	-13	9.8	13.7	0	4	-14	22.4	20.4
2	1	-11	62.3	63.7	1	7	-12	73.3	74.1	0	10	-13	8.4	-11.0	0	5	-14	54.3	50.8
2	2	-11	33.3	28.9	1	7	12	25.8	25.5	1	1	13	11.6	-9.0	0	6	-14	73.6	70.6
2	2	11	19.7	19.0	1	8	12	21.3	23.1	1	2	13	35.6	-32.5	0	8	-14	44.5	-48.5
2	3	-11	11.8	-9.4	1	9	-12	20.6	-24.7	1	2	-13	11.5	10.4	1	0	14	25.9	-28.4
2	3	11	15.3	-14.8	1	9	12	13.4	-14.4	1	3	-13	20.9	21.2	1	0	-14	124.4	-123.9
2	4	-11	72.6	67.2	1	10	-12	10.5	12.6	1	3	13	47.6	-47.4	1	1	-14	16.9	16.1
2	4	11	22.6	20.0	1	10	12	25.2	25.3	1	3	13	80.7	81.9	1	1	-14	46.4	46.1
2	5	11	30.4	30.4	2	0	-12	75.8	-73.5	1	4	-13	9.4	10.5	1	2	14	15.2	-15.1
2	5	-11	17.4	-14.0	2	0	12	34.1	-33.6	1	5	-13	68.7	-62.7	1	2	-14	9.9	7.4
2	6	-11	41.2	41.4	2	1	12	5.6	7.8	1	5	13	47.6	-47.4	1	2	-14	31.7	28.4
2	6	11	38.8	-37.6	2	1	-12	80.7	-81.3	1	6	13	7.3	-5.8	1	3	14	19.7	18.1
2	7	11	29.3	-32.5	2	2	-12	17.0	16.7	1	6	-13	24.7	-21.2	1	4	-14	27.4	25.4
2	7	-11	63.4	-62.3	2	2	12	32.5	-33.3	1	7	-13	39.7	-42.3	1	5	-14	55.3	-55.4
2	8	11	10.0	12.7	2	3	12	51.4	-51.4	1	4	-13	12.1	-13.5	1	5	14	11.4	-10.5
2	9	-11	14.4	19.8	2	3	-12	8.7	7.5	1	8	-13	5.7	-4.1	1	6	14	38.4	-38.4
2	10	-11	16.0	17.0	2	4	-12	31.5	-31.0	1	9	-13	24.3	26.4	1	6	-14	31.7	30.1
3	1	-11	10.2	11.4	2	4	12	7.4	-8.4	1	9	13	10.5	-12.8	1	7	14	24.4	-24.5
3	2	-11	63.4	-62.3	2	5	-12	21.3	-21.3	2	1	13	31.3	30.0	1	8	-14	12.1	15.2
3	2	11	43.2	-42.8	2	5	12	10.7	-9.7	2	2	-13	21.9	-21.6	1	9	-14	5.8	5.0
3	3	-11	49.9	-47.0	2	6	-12	29.7	26.1	2	2	-13	87.7	-84.6	1	9	-14	10.4	-13.2
3	3	11	90.2	-87.4	2	6	12	19.4	-19.5	2	3	13	6.1	-5.7	1	9	14	26.9	24.2
3	4	-11	17.1	16.7	2	7	-12	53.5	54.9	2	3	13	41.3	39.3	2	0	-14	26.8	24.2
3	4	11	19.0	-18.4	2	7	12	51.9	-51.1	2	3	-13	24.0	21.7	2	0	14	16.8	14.5
3	5	-11	17.2	16.9	2	8	-12	35.9	-36.5	2	4	-13	14.8	17.0	2	1	-14	44.4	-44.4
3	5	11	19.0	20.1	2	8	12	9.6	-9.9	2	4	13	31.0	29.7	2	1	-14	12.5	-11.0
3	6	-11	6.1	5.4	2	9	-12	10.4	-11.6	2	5	-13	8.4	-6.2	2	2	-14	26.7	26.3
3	6	11	14.2	14.8	2	9	12	47.5	43.4	2	5	13	24.4	24.1	2	2	14	16.4	-35.0
3	7	-11	15.3	18.4	2	0	-12	31.5	-34.4	2	6	-13	76.8	79.2	2	3	14	7.3	8.7
4	1	-11	45.2	45.2	3	0	12	36.2	-36.4	2	6	13	9.8	10.4	2	3	-14	5.4	-5.4
4	2	-11	4.5	-4.2	3	1	-12	28.0	26.8	2	7	-13	7.9	-8.0	2	4	-14	56.7	-47.2
4	2	11	27.5	28.4	3	1	12	63.9	62.1	2	7	13	15.9	15.2	2	5	14	12.3	10.7
4	4	-11	29.3	31.0	3	3	-12	16.1	15.8	2	8	-13	9.4	-11.2	2	5	-14	15.2	12.2
4	4	11	20.3	23.2	3	3	12	48.7	48.8	3	1	-13	11.3	7.6	2	6	-14	28.5	29.4
4	6	-11	29.3	-29.6	3	4	-12	75.1	-76.9	3	1	13	8.8	10.4	2	6	14	31.8	-31.9
5	1	-11	11.7	-11.6	3	4	12	10.8	8.4	3	2	-13	41.7	-41.8	2	7	-14	28.2	-30.0
0	0	-12	79.8	78.7	3	5	-12	12.8	-12.6	3	2	13	15.9	-14.0	2	8	-14	21.1	-24.7
0	0	12	17.7	-15.6	3	6	12	13.1	12.7	3	3	-13	47.4	52.2	2	8	14	16.9	18.4
0	3	-12	8.4	9.7	3	7	-12	10.2	14.4	3	3	13	52.8	-51.6	2	9	-14	14.2	-12.7
0	4	-12	10.5	-29.3	3	7	12	24.3	28.4	3	4	-13	24.3	23.4	3	0	-14	37.8	-19.3
0	5	-12	103.9	96.4	4	0	-12	87.8	-88.1	3	5	-13	23.8	-23.7	3	0	14	21.7	-23.1
0	6	-12	20.8	-22.7	4	1	-12	14.3	-13.6	3	5	13	17.6	15.0	3	1	-14	62.5	63.0
0	7	-12	59.7	-59.0	4	1	12	9.0	-10.2	3	5	13	14.1	-13.7	3	1	14	14.0	-14.0
0	8	-12	11.5	-10.4	4	2	-12	39.4	40.9	3	6	-13	11.3	12.9	3	2	-14	37.1	-37.1
0	9	-12	23.3	21.5	4	2	12	6.8	-8.5	3	6	13	32.8	-36.0	3	2	14	27.6	27.7
1	0	12	101.0	-99.4	4	3	-12	24.9	-24.8	3	7	-13	13.2	12.4	3	3	-14	40.8	-41.4



M	K	L	F (ORS)	F (CALC)	M	K	L	F (ORS)	F (CALC)	M	K	L	F (ORS)	F (CALC)	M	K	L	F (ORS)	F (CALC)
3	3	-14	40.8	-41.4	3	4	-15	26.1	-28.4	0	2	-17	21.9	20.4	1	6	18	16.1	-17.2
3	3	-14	13.2	14.0	4	3	-15	10.7	-8.7	0	2	-17	15.8	-16.4	1	6	18	16.1	-17.2
3	4	-14	27.8	-31.6	4	3	-15	18.3	17.4	0	4	-17	26.2	25.8	1	7	-18	10.8	-28.8
3	5	-14	33.8	33.9	4	3	-15	45.5	-48.9	0	5	-17	54.2	-52.4	1	7	18	14.9	-12.9
3	5	-14	11.4	-14.5	4	4	-15	8.2	-8.6	0	6	-17	22.3	-23.6	1	8	-18	7.4	8.7
3	6	-14	9.3	9.0	0	0	-16	36.4	33.5	0	7	-17	21.4	-21.8	2	0	-18	15.1	-15.5
3	7	-14	21.4	23.4	0	1	-16	15.7	12.4	0	8	-17	15.3	18.6	2	0	-18	16.7	-20.1
4	0	-14	8.2	-10.7	0	2	-16	45.0	48.3	1	1	-17	25.7	-25.1	2	1	18	24.3	-27.1
4	1	-14	13.8	35.2	0	4	-16	23.6	-25.2	1	2	-17	21.0	-18.6	2	2	-18	7.5	-72.8
4	2	-14	18.5	38.7	0	5	-16	26.1	26.0	1	3	-17	20.8	-20.4	2	3	18	10.9	-12.0
4	4	-14	12.5	11.4	0	6	-16	9.0	69.8	1	3	-17	40.8	40.4	2	4	-18	39.9	38.3
4	5	-14	13.3	-15.9	0	7	-16	24.2	-24.5	1	3	-17	48.2	-45.8	2	4	-18	39.9	38.3
4	5	-14	58.1	-52.9	0	8	-16	8.9	-12.1	1	4	-17	16.8	14.6	2	5	-18	34.2	37.0
0	1	-15	51.0	58.7	0	9	-16	15.4	18.1	1	5	-17	31.3	32.1	2	6	-18	7.7	-8.0
0	2	-15	65.2	-62.2	1	0	-16	46.8	66.4	1	5	-17	6.8	7.1	2	7	-18	17.1	19.6
0	3	-15	57.0	-55.1	1	0	-16	6.8	-43.2	1	6	-17	10.5	10.8	3	0	-18	21.5	22.2
0	5	-15	6.2	7.1	1	1	-16	9.6	7.1	1	6	-17	35.4	34.9	3	1	-18	44.2	44.1
0	6	-15	7.6	-4.4	1	1	-16	32.3	33.3	1	7	-17	20.8	-22.5	3	2	-18	15.1	14.3
0	7	-15	20.6	-18.4	1	2	-16	32.0	-34.5	1	7	-17	11.3	13.9	3	3	-18	11.0	-14.5
0	8	-15	12.2	14.7	1	2	-16	6.3	-7.1	1	8	-17	25.6	28.8	3	4	-18	21.3	21.1
0	9	-15	7.2	-6.5	1	3	-16	14.3	-13.4	2	1	-17	20.6	21.6	3	6	-18	17.4	20.0
1	1	-15	25.9	26.5	1	3	-16	34.3	-32.9	2	1	-17	37.3	-33.3	4	0	-18	6.1	7.5
1	1	-15	19.9	21.3	1	4	-16	15.9	15.5	2	2	-17	15.4	-15.9	4	0	-18	4.9	10.1
1	2	-15	52.7	-50.5	1	4	-16	12.3	-16.7	2	2	-17	11.2	-10.2	0	2	-19	16.8	16.1
1	2	-15	10.2	7.4	1	5	-16	7.3	8.5	2	3	-17	7.3	-6.7	0	3	-19	24.7	24.4
1	3	-15	21.4	19.4	1	5	-16	10.1	-31.1	2	3	-17	101.0	103.3	0	4	-19	6.4	7.1
1	3	-15	10.8	-29.2	1	6	-16	16.4	-17.5	2	4	-17	5.9	-5.1	0	5	-19	17.4	17.4
1	4	-15	40.5	37.4	1	6	-16	7.3	-7.4	2	5	-17	7.2	-7.9	0	7	-19	8.5	-8.4
1	4	-15	56.9	54.3	1	7	-16	59.5	-62.5	2	5	-17	19.8	-21.8	1	1	-19	44.9	-45.1
1	5	-15	31.0	33.5	1	7	-16	28.7	32.4	2	6	-17	29.6	-30.2	1	1	-19	29.4	-27.9
1	5	-15	29.5	29.4	1	8	-16	7.0	-8.5	2	8	-17	12.8	-15.9	1	2	-19	86.0	85.3
1	6	-15	14.1	14.4	1	8	-16	14.4	-13.9	2	8	-17	12.8	17.7	1	3	-19	33.9	35.3
1	6	-15	11.7	9.5	1	9	-16	7.8	9.0	3	1	-17	16.8	14.1	1	3	-19	18.0	-16.1
1	7	-15	45.8	-44.5	2	0	-16	61.5	57.3	3	2	-17	12.5	-16.4	1	4	-19	19.0	20.4
1	7	-15	18.3	19.2	2	0	-16	20.2	-18.5	3	3	-17	34.8	-31.1	1	4	-19	47.2	-49.8
1	9	-15	20.7	26.4	2	1	-16	11.6	-12.5	3	4	-17	17.5	21.9	1	5	-19	7.4	9.0
2	1	-15	8.4	-6.4	2	2	-16	15.9	-17.3	3	5	-17	20.7	-19.7	1	5	-19	20.5	-21.4
2	1	-15	52.9	-50.4	2	2	-16	40.0	-39.7	4	1	-17	18.6	-18.0	1	6	-19	5.5	-5.2
2	2	-15	12.3	-11.0	2	4	-16	12.1	12.1	4	2	-17	9.7	-10.8	2	1	-19	12.6	14.7
2	2	-15	40.6	-41.7	2	5	-16	6.4	6.0	4	3	-17	7.3	-9.4	2	1	-19	16.9	17.1
2	3	-15	31.4	29.4	2	5	-16	32.7	36.8	4	3	-17	26.4	-56.5	2	2	-19	15.7	-14.0
2	3	-15	38.4	38.1	2	6	-16	18.5	-16.6	0	0	-18	55.8	23.5	2	3	-19	22.9	-24.4
2	4	-15	29.9	31.4	2	7	-16	9.0	9.4	0	1	-18	8.6	-7.0	2	3	-19	40.4	41.2
2	5	-15	11.0	11.3	3	0	-16	6.1	6.3	0	3	-18	8.6	-8.4	2	5	-19	22.1	-24.2
2	6	-15	18.0	14.7	3	1	-16	37.0	37.2	0	6	-18	14.9	14.5	2	6	-19	8.4	-10.9
2	6	-15	9.0	-10.4	3	2	-16	12.6	-14.1	0	6	-18	27.3	-27.7	3	1	-19	28.3	-28.8
2	6	-15	13.1	-14.7	3	2	-16	27.8	-29.3	0	7	-18	35.6	31.7	3	3	-19	28.4	-29.9
3	1	-15	13.3	14.8	3	3	-16	53.1	-55.6	1	0	-18	33.5	34.2	3	4	-19	11.9	13.3
3	1	-15	44.8	44.4	3	3	-16	17.8	-15.8	1	1	-18	15.4	14.3	3	5	-19	5.9	-8.5
3	2	-15	21.3	21.3	3	5	-16	30.1	30.3	1	2	-18	20.2	19.5	0	0	-20	34.5	34.9
3	2	-15	31.8	-29.8	3	6	-16	29.4	29.1	1	2	-18	27.7	26.2	0	1	-20	34.8	34.1
3	3	-15	16.3	14.4	3	7	-16	7.1	7.9	1	3	-18	49.8	-47.9	0	2	-20	44.5	-45.5
3	3	-15	21.2	-20.9	4	0	-16	31.9	32.2	1	3	-18	24.4	23.3	0	3	-20	9.5	-10.2
3	4	-15	9.6	11.9	4	1	-16	10.7	14.0	1	4	-18	37.7	37.3	0	4	-20	12.5	-15.0
3	4	-15	13.5	37.0	4	2	-16	18.9	18.0	1	4	-18	14.1	-14.6	0	5	-20	20.1	-19.4
3	5	-15	10.9	-8.4	4	4	-16	10.4	11.1	1	5	-18	40.3	-42.7	1	0	-20	24.4	23.1



**APPENDIX III**  
**The Amino Acids and**  
**The Sequence of  $\alpha$ -Chymotrypsin**

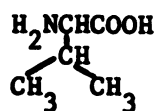
## APPENDIX III

### Aliphatic Acids

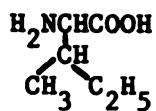
Glycine (Gly)



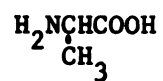
Valine (Val)



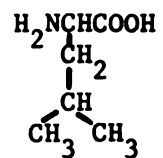
Isoleucine (Ile)



Alanine (Ala)

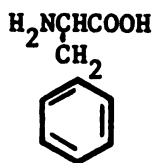


Leucine (Leu)



### Aromatic Amino Acids

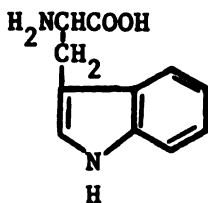
Phenylalanine (Phe)



Tyrosine (Tyr)

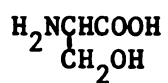


Tryptophan (Trp)

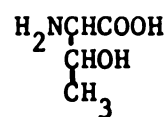


## Hydroxyamino Acids

Serine (ser)

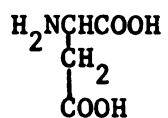


Threonine (Thr)

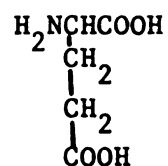


## Acidic Acids

Aspartic acid (Asp)

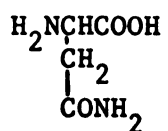


Glutamic acid (Glu)

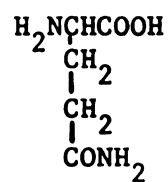


## Acid Amide Amino Acids

Asparagine (Asn)

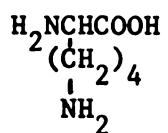


Glutamine (Gln)

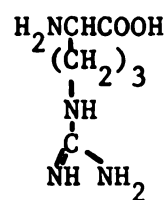


## Basic Amino Acids

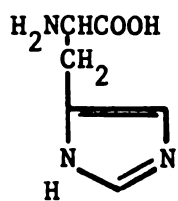
Lysine (Lys)



Arginine (Arg)

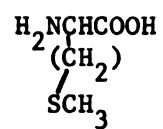


Histidine (His)

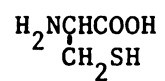


## Sulfur-Containing Amino Acids

Methionine (Met)

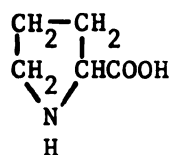


Cystine (Cys)



## Secondary Amino Acids

Proline (Pro)



Val·Leu·Ser·Gly·Leu————Ilu·Val·Asn·Gly·Glu·Glu·Alg·Val·Pro·Gly·Ser·Trp·Pro·Trp  
21 Gln  
Pro  
Gln Glu·Asn·Ilu·Leu·Ser·Gly·Gly·Cys·Phe·His·Phe·Gly·Thr·Lys·Asp·Gln·Leu·Ser·Val  
Ilu Asn S 40  
Ala Trp·Val·Val·Thr·Ala·Ala·His·Cys·Gly·Val·Thr·Thr·Ser·Asp·Val·Val·Val  
57  
Pro  
Val Ilu·Lys·Leu·Lys·Gln·Ilu·Lys·Glu·Ser·Ser·Ser·Gly·Gln·Asp·Phe·Gly·Gly  
Gly Ala  
Cys Lys·Val·Phe·Lys·Asn·Ser·Lys·Tyr·Asp·Ser·Leu·Thr·Ilu·Asn·Asn·Asn  
S Ilu  
S Thr  
Cys·Val·Ala·Ser·Val·Thr·Gln·Ser·Phe·Ser·Ala·Ala·Thr·Ser·Leu·Lys·Leu·Leu  
Leu  
Pro Gly·Leu·Thr·Arg·Tyr————Ala·Asn·Thr·Pro·Asp·Arg·Leu·Gln·Gln·Ala  
154 Ser  
Ser Trp  
Ala Gly Thr·Gly·Trp·Tyr·Lys·Lys·Cys·Asn·Thr·Asn·Ser·Leu·Leu·Pro·Leu  
Ser Lys S  
Asp Thr Ilu·Lys·Asp·Ala·Met·Ilu·Cys·Ala·Gly·Ala·Ser·Gly Val  
Val Ser  
Phe·Ala·Ala·Gly·Thr·Thr·Cys·S S Cys·Val·Leu·Pro·Gly·Gly·Ser·Asp·Gly·Met·Cys·Ser  
195 S 190  
Lys  
Lys S  
Asn Cys·Ser·Thr·Ser  
221 Thr  
Gly Ser  
Ala·Trp·Thr·Leu·Val·Gly·Ilu·Val·Ser·Trp·Gly·Ser  
215  
Asn Ala Ala Leu Thr Gln Gln Val Trp Asn Val Leu Ala Thr Val Arg Ala Tyr







MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03046 5342