PART I: THE CRYSTAL AND MOLECULAR STRUCTURE OF DIMETHOXYPORPHINATO Ge(IV)

PART II: THREE DIMENSIONAL STUDY OF α-CHYMOTRYPSIN AT pH 8.7 and 2.7 WITH DIFFERENCE FOURIER METHOD.

By

Aristides Mavridis

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1975

ABSTRACT

PART I: THE CRYSTAL AND MOLECULAR STRUCTURE OF DIMETHOXYPORPHINATO Ge(IV)

PART II: THREE DIMENSIONAL STUDY OF α-CHYMOTRYPSIN AT pH 8.7 and 2.7 WITH DIFFERENCE FOURIER METHOD

By

Aristides Mavridis

The structure of dimethoxyporphinato Ge(IV) ($PGe(OMe)_2$) was determined by three dimensional crystallographic techniques. The molecule crystallizes in space group $P2_1/c$ with four molecules in the unit cell and cell dimensions $|\vec{a}|$ = 15.015 A, $|\vec{b}|$ = 14.441 A, $|\vec{c}|$ = 8.414 A and β = 91.85°. The structure was solved using Savre's equation and it was refined to an R-factor of 0.043.

There are two centrosymmetric and symmetry independent pairs of molecules in the unit cell. Small differences between the independent molecules are probably due to packing effects. The porphine molecules are stacked along the $\vec{c}$ axis composing a very efficient packing. The porphine ring is essentially planar with the methoxy groups slightly deviating from being perpendicular to this plane.

The structure of the hydrolytic enzyme α-CHT was studied in the pH's 8.7 and 2.7 with the difference Fourier method. Difference electron density maps between the electron density of the enzyme at pH 3.6 (native) and 8.7, as well as at pH 3.6 and 2.7 (using the phases of the native only) revealed a number of molecular changes. The main feature of the high pH difference map is a general lack of local two fold symmetry. The

largest changes in the high pH difference map correspond to movements of three separate segments of the α-CHT molecule.

The low 3.6 - 2.7 changes are essentially confined on the surface of the enzyme and they are generally small as compared with the 3.6 - 8.7 changes.

To the Greek students who died in Greece fighting for freedom and democracy during the period of April 21, 1967 to July 23, 1974.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Alexander Tulinsky for his assistance during this study.

The author is indebted to Dr. Richard L. Vandlen for his assistance with the computing aspects of this work, as well as for his general help.

To Mr. Michael Liebman, the author expresses his sincere appreciation for many helpful discussions.

Support of the Molecular Biology Section of the National Science Foundation during the course of this study is acknowledged.

The author also likes to thank his wife, Irene Moustakali Mavridis, not only for her assistance during this work, but also for her patience, courage, and understanding; things that so much one needs, particularly in a foreign country.

TABLE OF CONTENTS

PART I: THE CRYSTAL AND MOLECULAR STRUCTURE OF DIMETHOXYPORPHINATO Ge(IV)

I. INTRODUCTION 1

1. Historical Note
2. The Geometrical Theory of X-ray Diffraction. 2
   (a) The Atomic Scattering Factor. 2
   (b) Periodic Lattices 6
   (c) Phase Problem and Electron Density Representation of a Crystal. 10
   (d) Bragg's Law and Intensity Relations 15
   (e) The Patterson Function. 21
   (f) The Temperature Factor. 23

II. PORPHYRINS 26

III. EXPERIMENTAL 32

1. Crystals 32
2. Intensity Data Collections 32
3. Absorption 36

IV. STRUCTURE ANALYSIS 40

V. RESULTS. 58

VI. DISCUSSION 72

PART II: THREE DIMENSIONAL STUDY OF α-CHYMOTRYPSIN AT pH 8.7 AND 2.7 WITH DIFFERENCE FOURIER METHOD

PROLEGOMENA. 82

VII. INTRODUCTION. 84

1. Enzymes: Historical. 84
2. Alpha Chymotrypsin: A General Discussion 86
3. The Three Dimensional Structure of α-CHT. 92
4. Effect of pH on α-CHT 99

VIII. EXPERIMENTAL. 114

1. Crystal Preparation 114
2. Intensity Data Collection and Reduction 122
3. Data Processing 127

IX. DIFFERENCE FOURIER METHOD 131

X. RESULTS AND DISCUSSION. 135

GENERAL OBSERVATIONS. 135

A. High pH (8.3) Structural Changes of Crystalline α-CHT . . . 136

1. Dimeric Interactions. 136
2. Histidine 40. 144
3. The Active Site Region. 149
   3-1. Isoleucine 16 Valine 17 Dipeptide 150
   3-2. Aspartic 194. 156
   3-3. Tyrosine 146 Chain. 159
4. Surface Interactions. 164

B. Low pH (2.5) Structural Changes of Crystalline α-CHT. . . . 173

XI. FINAL COMMENTS. 183

REFERENCES. 194

APPENDIX. 205

LIST OF TABLES

Table		Page
I.	Crystal and Unit Cell Data of $PGe(OMe)_2$	34
II.	Statistical Distribution of $\|E\|$	53
III.	Final Atomic Coordinates, Thermal Parameters and Peak Heights. Molecule 1	60
IV.	Final Atomic Coordinates, Thermal Parameters and Peak Heights. Molecule 2	61
V.	Final Atomic Coordinates, Isotropic Temperature Factors and Peak Heights of the Hydrogen Atoms	62
VI.	Deviations of Porphine Skeleton From Least Squares Plane	63
VII.	Deviations of Pyrroles From Least Squares Planes	64
VIII.	Principal Mean-Square Displacements (A^2) in "Isotropic B" Notation	65
IX(a,b)	Comparison of Intrinsic Packing-Density of Porphyrins	74-75
X.	Some Molecular Interactions About the Local Two Fold Axis.	98
XI.	Ionizable Residues in α-CHT	106
XII.	Amino Acids that Commonly Occur in Proteins	107
XIII.	Charge Distribution of α-CHT as a Function of pH	111
XIV.	Scaling of High and Low pH Crystals of α-CHT With Respect to Native (h0l) Projection	128
XV.	Lattice Parameters of α-CHT Crystals at Four Different pH Values	134
XVI.	Negative Difference Densities Associated With the Movement of the Terminal Carboxylate Group (TYR146') of α-CHT as a Function of pH	138

Table		Page

XVII. Coordinates of Difference Density in ILE16 Region. . . . 153

XVIII. Coordinates of Difference Peaks in ASP194-SER190 Region. 157

XIX. Coordinates of Negative Difference Peaks in TYR146 Tetrapeptide Region. 162

XX. Coordinates of Positive Difference Peaks in TYR146 Tetrapeptide Region. 163

XXI. Surface Interactions 166

XXII. List of Residues for Which α-Carbon Atoms Differ in Position by More than 3.6A Between the CHTG and α-CHT Structures (From Reference 200). 184

TABLE OF FIGURES

Figure		Page
1.	Scattering by a charge distribution described by the vector $\vec{r}_n$. The unit vectors $\hat{s}_0$ and $\hat{s}$ give the directions of primary and scattered beams respectively; P is the observation point.	3
2.	Ewald's construction.	18
3.	Free base porphine.	27
4.	Two different "electron paths" for the porphyrin system .	29
5.	Absorption curves for $PGe(OMe)_2$ crystal; the numbers 006, 004, 002 indicate reciprocal lattice levels	39
6.	Harker section at v=1/2	42
7.	Labelling of $PGe(OMe)_2$.	56
8.	Bond distances (in A) and angles (in degrees); broken lines indicate C-H distances; molecule 1.	66
9.	Bond distances (in A) and angles (in degrees); broken lines indicate C-H distances; molecule 2.	67
10.	Intramolecular distances and angles of the central core region; molecule 1.	68
11.	Intramolecular distances and angles of the central core region; molecule 2.	69
12.	Computer plot of the $PGe(OMe)_2$ molecule (ORTEP, reference 100).	70
13.	Molecular packing of $PGe(OMe)_2$ (ORTEP, reference 100) . .	71
14.	Geometry of the Ge-OMe group with respect to the pyrollic nitrogens of the porphine macrocycle	77
15.	Average bond distances and angles of $PGe(OMe)_2$.	78

Figure | Page

16. Average bond distances and angles of the free base porphine (from reference 39). 79

17. The sequence of α-CHT 88

18. Formation of α-(or γ-) CHT from its inactive precursor chymotrypsinogen. 90

19. Schematic representation of the dimeric unit of α-CHT. View parallel to the local two fold axis. 93

20. Schematic packing diagram of α-CHT viewed along the $\vec{a}^*$ direction. Molecules I and I' form an asymmetric unit and are related by non crystallographic two fold rotation axes A and B; active site regions near center of dimer designated with asterisks; uranyl binding region designated by U (reference 154). 96

21. The α-CHT-catalyzed hydrolysis of N-acetyl-L-tryptophan amide (a), and of N-acetyl-L-phenyl-alanine (b) in water at 25°C as a function of pH; (reference 165). . . . 101

22. Dependence of the natural logarithm of the dimerization equilibrium constant on pH. Conditions: 0.01M NaCl, 0.01M acetate buffer, 25°C; maximum dimer formation occurs at pH 4.4 (reference 166). 102

23. The pH dependence of the specific rotation of α-CHT at 340 nm and 12°C (reference 167). 103

24. Model of a globular protein molecule, based on the assumption that charge is uniformly distributed over its surface 105

25. Histogram of total algebraic charge (Q) on α-CHT molecule as a function of pH. 112

26. Axial scans along $\vec{a}^*$, $\vec{b}^*$ and $\vec{c}^*$ directions formative (a) pH 8.3 (b) and pH 2.5 (c) α-CHT crystals. The log of the diffracted intensity is plotted as a function of 2θ. The order of diffraction is indicated above each peak. 119-121

27. Absorption curves for the high pH crystals of α-CHT as a function of φ 125

28. Absorption curves for the low pH crystals of α-CHT as a function of φ 126

Figure Page

29. Radial distribution plots of native (——), pH 8.3 (---) and pH 2.5 (oooo) crystals of α-CHT. 130

30. Rotation of terminal carboxylate group of TYR146' around the NH-C_α bond. The local two fold axis shown is perpendicular to the rotation axis. 140

31. Difference electron density between pH 8.3 and 3.6 crystals of α-CHT in vicinity of active site viewed down the local two fold axis; contours at $0.05eA^{-3}$ beginning with $0.15eA^{-3}$; solid contours positive, broken negative; thickness shown ∿6A; 3-1, 3-2, 3-3 are section numbers where the difference densities are discussed; position of local two fold axis indicated by cross 142

32. The quenching of the fluorescence of δ-chymotrypsin as a function of pH, measured at 340 nm and 25°C. (reference 143). 146

33. Movement of the imidazole ring of HIS40 by rotation around the C_α-C_β bond; the rotation axis is approximately perpendicular to the local two fold axis. 148

34. Schematic representation of α-CHT folding around the internal ion pair ILE16⟷ASP194; view perpendicular to the local two fold axis 151

35. Negative difference electron density layers perpendicular to the plane of the paper (yz plane); they correspond to Table XVII. Their relative size is approximately proportional to peak heights 154

36. Schematic representation of the U shaped folding of the tetrapeptide TYR146-ARG145-THR144-LEU143 of α-CHT view perpendicular to the local two fold axis. 160

37. Geometrical arrangement of the residues THR135, SER159, SER186 and the corresponding residues on molecule II' of another dimer. The circled positive and negative signs indicate positive and negative difference densities respectively. 171

38. Possible schemes of interactions between the interdimer carboxyl groups of ASP153' and GLU21 residues. . . . 175

39. Schematic representation of the residues ASP64 and ALA149' of α-CHT; view perpendicular to the local two fold axis. 178

Figure | Page

40. Geometrical arrangement of a part of α-CHT molecule showing the manner of formation of an extended pleated sheet. 180

41. Possible mechanism of action for α-CHT; (a) acylation step; (b) deacylation step (reference 136) 186

PART I

THE CRYSTAL AND MOLECULAR STRUCTURE OF DIMETHOXYPORPHINATO Ge(IV)

I. INTRODUCTION

1. Historical Note

Few things in nature are more perfect than a crystal, in which immense numbers of atoms or molecules are stacked in almost perfect alignment. Crystallography, which is a small branch of the vast field of solid state physics, studies mainly the geometrical arrangement in space of atoms or molecules taking advantage of their ordering. Crystallography made its start as a science in 1669 when the Danish physician N. Steno discovered that although quartz crystals may differ in appearance from one another, the angles between corresponding faces are always the same. By 1891 the complete space group theory was published by A. Schoenflies[1] and E. von Fedorov[2], but yet nothing was known about the interior of crystals. The discovery of X-rays by W. C. Röntgen in 1895, and its subsequent use by von Laue in 1912[3-5] to prove that crystals are natural diffraction gratings for X-rays, was a great impetus to solid state physics and crystallography (at the time, centered mostly around mineralogy).

Immediately after Friedrich and Knipping's discovery of X-ray diffraction and von Laue's mathematical analysis[4], W. H. Bragg and his son W. L. Bragg approached the diffraction problem in a different way[6,7] namely, as the "reflection" of X-rays by planes in crystals. Still another approach to the diffraction problem was taken by P. P. Ewald[8] who developed a more sophisticated and complete theory, called the dynamical theory. Essentially the same mathematical theories and

experimental techniques were used later, after L. de Broglie's proposal of his famous equation of matter waves in 1924[9], and its subsequent verification via electron diffraction by C. J. Davisson and L. J. Germer[10] by using thermal neutrons or electrons instead of X-rays to probe the nature of crystalline materials.

The contribution of diffraction studies to our understanding of nature is really profound; the history and the role in modern physical sciences are described in a fascinating book entitled "Fifty Years of X-ray Diffraction" edited by P. P. Ewald[11].

2. The Geometrical Theory of X-ray Diffraction

The purpose of this section is to introduce some ideas, the appropriate nomenclature, and to present some of the basic equations used in X-ray crystal structure analysis. A series of excellent books are available which describe in detail the geometrical (or kinematical) theory of X-ray diffraction, its power and its limitations; some of them are given in references 12-20. The term geometrical distinguishes this theory from the dynamical by the power, depth and ability of the latter to predict and explain diffraction phenomena beyond those of the geometrical theory. The most complete account of the dynamical theory in the von Laue formulation (in the English literature) is an article written by James[21].

(a) The Atomic Scattering Factor

We consider a hydrogen-like atom with its nucleus located at the origin, and the position of its electron described by a vector $\vec{r}_n$ (Figure 1). A plane wave polarized perpendicular to the plane of the paper, and with direction of incidence described by the unit vector $\hat{s}_o$,

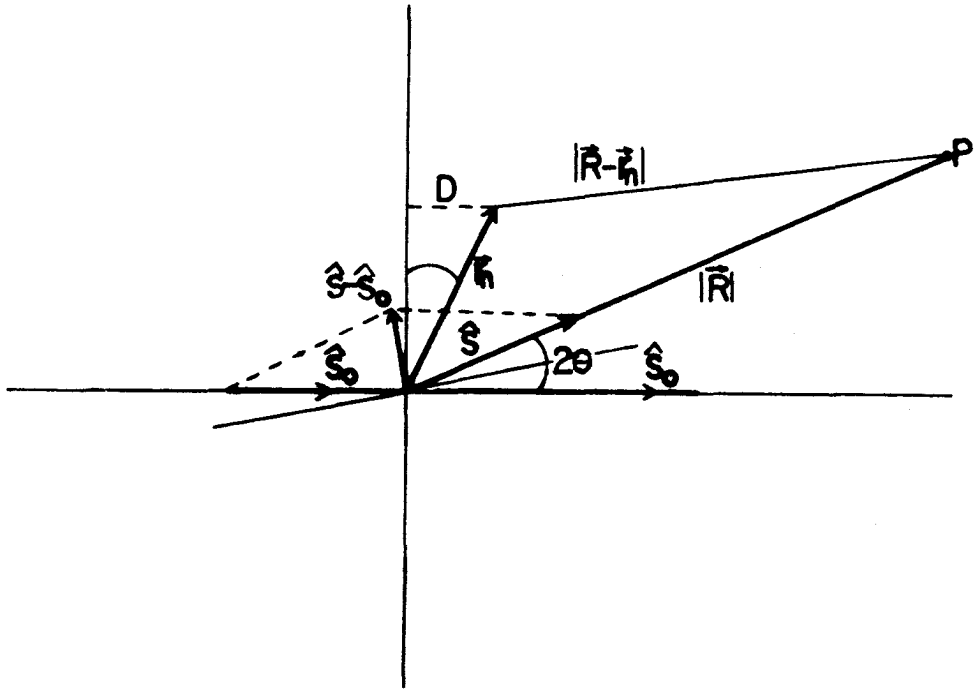


Figure 1. Scattering by a charge distribution described by the vector $\vec{r}_n$. The unit vectors $\hat{s}_0$ and $\hat{s}$ give the directions of primary and scattered beams respectively; P is the observation point.

is scattered from the charge distribution around the origin, and the scattered wave is observed at a point P, along the direction $\hat{s}$, and at a distance $|\vec{R}|$ ($\vec{R}=|\vec{R}|\hat{s}$) from the origin, with $|\vec{R}|>>|\vec{r}_n|$ ($|\vec{r}_n|\simeq 1A$, $|\vec{R}_n|\simeq 10cm$). The <u>instantaneous</u> value of the field in the primary beam on the electron is given by

$$\vec{E}=\vec{E}_o\exp\{i(2\pi\vec{\kappa}\cdot\vec{r}_n-\omega t)\}=\vec{E}_o\exp\{i(2\pi\kappa D-\omega t) \qquad (1)$$

where $\kappa=|\vec{\kappa}|=\frac{1}{\lambda}$, $\omega=2\pi\nu$ and $D=\vec{r}_n\cdot\hat{s}_o$. At the observation point P the magnitude and phase of the field $\vec{E}$ will be given by[17]

$$\vec{E}_n=\frac{\vec{E}_o e^2}{mc^2|\vec{R}-\vec{r}_n|}\exp\{2\pi i\kappa(D+|\vec{R}-\vec{r}_n|)-i\omega t\}, \qquad (2)$$

where e^2/mc^2 is the classical electron radius $\simeq 2.82\times10^{-13}$cm. Due to the fact that $|\vec{R}|>>|\vec{r}_n|$ equation (2) transforms to

$$\vec{E}_n=\frac{\vec{E}_o e^2}{mc^2|\vec{R}|}\exp\{i(2\pi\kappa|\vec{R}|-\omega t)\}\exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}_n\} \qquad (3)$$

Summing over all the instantaneous fields at P given by equation (3), an expression is obtained for the scattered wave $\vec{E}_s$ due to <u>one</u> electron

$$\vec{E}_s=\frac{\vec{E}_o e^2}{mc^2|\vec{R}|}\exp\{i(2\pi\kappa|\vec{R}|-\omega t)\}\sum_n\exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}_n\} \qquad (4)$$

Expression (4) is completely classical, but we have to make some concessions to the quantum mechanical nature of matter. In order to obtain the unmodified scattering, the summation over the instantaneous positions of $\vec{r}_n$ of the electron should be substituted by an integration over a volume $d\vec{r}$, of a quantum mechanical charge distribution[17], thus

$$\vec{E}_s=\frac{\vec{E}_o e^2}{mc^2|\vec{R}|}\exp\{i(2\pi\kappa|\vec{R}|-\omega t)\}\int\exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}\}\rho(\vec{r})d\vec{r} \qquad (5).$$

The integral quantity of equation (5) is called the "scattering factor

per electron" and is symbolized by f_e:

$$f_e=\int \exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}\}\rho(\vec{r})d\vec{r} \tag{6}$$

or

$$f_e=\int \exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}\}|\psi(\vec{r})|^2 d\vec{r}, \tag{7}$$

where $\psi(\vec{r})$ is the ground state wave function of a hydrogen-like atom, or more generally, a one electron quantum mechanical function. If $\psi(\vec{r})$ has spherical symmetry, equation (7) transforms easily to a simpler expression using spherical polar coordinates

$$f_e=\int 4\pi r^2|\psi(r)|^2 \frac{\sin\mu r}{\mu r}\, dr, \tag{8}$$

where $\mu=4\pi\sin\theta/\lambda$ (9)

(see Figure 1 for the definition of the angle θ).

For a polyelectronic atom we can infer by analogy to equation (7), that the "atomic scattering factor f" will be given by

$$f=\int |\psi(\vec{r}_1,\vec{r}_2,\ldots\vec{r}_z)|^2 \sum_j \exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}_j\}d\vec{r}, \tag{10}$$

where $\vec{r}_j$ is the vector to the volume element dv_j and the integration is over all the 3z electron coordinates. As an approximation to the wave function $\psi(\vec{r}_1,\vec{r}_2,\ldots\vec{r}_z)$ we can consider a Slater determinantal wave function

$$\Phi=\frac{1}{\sqrt{z!}}\begin{vmatrix} \phi_1(\xi_1) & \phi_1(\xi_2)\ldots & \phi_1(\xi_z) \\ \phi_2(\xi_1) & \phi_2(\xi_2)\ldots & \phi_2(\xi_z) \\ \ldots & \ldots & \ldots \\ \phi_z(\xi_1) & \phi_z(\xi_2)\ldots & \phi_z(\xi_z) \end{vmatrix}, \tag{11}$$

where $\phi_1(\xi_1),\ldots,\phi_Z(\xi_Z)$ are monoelectronic wave functions. Using equation (11), expression (10) transforms to

$$f=\sum_j^Z\int|\psi_j(r)|^2\exp\{2\pi i\kappa(\hat{s}-\hat{s}_o)\cdot\vec{r}_j\}d\vec{r}_j=\sum_j^Z f_{e_j} \qquad (12)$$

(see also reference 14).

The atomic scattering factor f is a basic quantity in X-ray crystallography, and an accurate structure determination requires an accurate set of f-factors for each atom of the structure.

(b) Periodic Lattices

The basic idea of a crystal is its translational invariance of some unit of atoms or molecules. The unit of the translational repetition can be defined by three, in general,nonorthogonal vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$, called the crystal axes. The parallelepipedon defined by the three axes $\vec{a}$, $\vec{b}$ and $\vec{c}$ is the smallest volume which, if repeated, will make up the crystal; it is called unit cell and its volume is given by

$$V=\vec{a}\cdot\vec{b}x\vec{c}\ . \qquad (13)$$

A lattice is defined using the fundamental lengths $\vec{a}$, $\vec{b}$, $\vec{c}$ as

$$\vec{R}=u\vec{a}+v\vec{b}+w\vec{c}, \qquad (14)$$

where u, v and w are integers. The position of an atom located at a certain point in space, inside any unit cell of a crystal will be given by the vector

$$\vec{R}_n=\vec{R}+\vec{r}_n, \qquad (15)$$

with the vector $\vec{r}_n$ expressed in fractional coordinates with respect to $\vec{a}$, $\vec{b}$ and $\vec{c}$.

According to the above,any property of the crystal, (electron density, conductivity, electrostatic potential,...) will be a periodic

function with respect to the lattice translation $\vec{R}$

$$\Phi(\vec{r}+\vec{R})=\Phi(\vec{r}). \tag{16}$$

However, a periodic function can be expressed as a Fourier series[22]; therefore

$$\Phi(\vec{r})=\sum_{\vec{r}*} C_{\vec{r}*}\exp(-2\pi i\vec{r}\cdot\vec{r}*) \tag{17}$$

where the quantity $\vec{r}*$ is a <u>reciprocal</u> <u>vector</u> and is defined by the following expressions

$$\vec{r}*=h\vec{a}*+k\vec{b}*+l\vec{c}*, \tag{18}$$

where h, k, l are integers, and

$$\vec{a}*=\frac{\vec{b}x\vec{c}}{\vec{a}\cdot\vec{b}x\vec{c}}=\frac{\vec{b}x\vec{c}}{V}$$

$$\vec{b}*=\frac{\vec{c}x\vec{a}}{\vec{a}\cdot\vec{b}x\vec{c}}=\frac{\vec{c}x\vec{a}}{V} \tag{19}$$

$$\vec{c}*=\frac{\vec{a}x\vec{b}}{\vec{a}\cdot\vec{b}x\vec{c}}=\frac{\vec{a}x\vec{b}}{V}$$

Reciprocal vectors were first introduced into vector analysis by Gibbs[23] and their usefulness to crystallography was first pointed out by Ewald[24]. They generate a <u>reciprocal</u> lattice of the original direct (physical) lattice. Some of its properties can be verified directly from the definitions (18) and (19)

$$\vec{a}\cdot\vec{b}*=0$$
$$\vec{a}\cdot\vec{a}*=1 \tag{20}$$

and all cyclic permutations;

$$\vec{R}\cdot\vec{r}^*=(u\vec{a}+u\vec{b}+w\vec{c})\cdot(h\vec{a}^*+k\vec{b}^*+l\vec{c}^*)=(uh+vk+wl)=\text{integer}. \qquad (21)$$

Another property of the reciprocal lattice is that the volume of a reciprocal unit cell is the reciprocal of the physical unit cell

$$V^*=\vec{a}^*\cdot\vec{b}^*x\vec{c}^*=\frac{1}{\vec{a}\cdot\vec{b}x\vec{c}}=V^{-1}. \qquad (22)$$

This property can be proved easily using the vector identity

$$\vec{A}x(\vec{B}x\vec{G})=\vec{B}(\vec{A}\cdot\vec{G})-\vec{G}(\vec{A}\cdot\vec{B}).$$

Expression (17) plays a prominent role in the theory of X-ray crystallography, being connected with the electron density of a crystal, whose exact distribution is the main problem of structural crystallography. The Fourier coefficients $C_{\vec{r}^*}$ can now be found due to the orthogonality properties of the plane waves $\exp(-2\pi i\vec{r}\cdot\vec{r}^*)$ and $\exp(2\pi i\vec{r}\cdot\vec{r}'^*)$, where $\vec{r}'^*=h'\vec{a}^*+k'\vec{b}^*+l'\vec{c}^*$[25]. Multiplying both sides of equation (17) with $\exp(2\pi i\vec{r}\cdot\vec{r}'^*)$ and integrating over the volume V we obtain

$$\int\Phi(\vec{r})\exp(2\pi i\vec{r}\cdot\vec{r}'^*)d\vec{r}=\sum_{hkl}C(hkl)x\int\exp\{2\pi i((h'-h)\vec{a}^*+(k'-k)\vec{b}^*+$$

$$+(l'-l)\vec{c}^*)\cdot\vec{r}\}dxdydz, \qquad (23)$$

where x, y, and z are the components of the Cartesian vectors $\hat{e}_x$, $\hat{e}_y$ and $\hat{e}_z$ respectively. The integration with respect to dxdydz can be replaced by an integration with respect to $d\xi d\eta d\zeta$ with ξ, η, ζ being components of the vector $\vec{r}$ and range from zero to one,

$$\vec{r}=\xi\vec{a}+\eta\vec{b}+\zeta\vec{c}.$$

Thus the volume element dxdydz can be substituted by

$$dxdydz = J d\xi d\eta d\zeta, \tag{24}$$

where J is the Jacobian of the transformation

$$J = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{vmatrix} = \begin{vmatrix} |\vec{a}|_x & |\vec{b}|_x & |\vec{c}|_x \\ |\vec{a}|_y & |\vec{b}|_y & |\vec{c}|_y \\ |\vec{a}|_z & |\vec{b}|_z & |\vec{c}|_z \end{vmatrix} = V. \tag{25}$$

Using (24) and (25) equation (23) becomes

$$\int \Phi(\vec{r}) \exp(2\pi i \vec{r}\cdot\vec{r}'^{*}) d\vec{r} = V \sum_{hkl} C(hkl) \int_0^1 \exp\{2\pi i (h'-h)\xi\} d\xi \times$$
$$\times \int_0^1 \exp\{2\pi i (k'-k)\eta\} d\eta \times \int_0^1 \exp\{2\pi i (l'-l)\zeta\} d\zeta =$$
$$= V \sum_{hkl} C(hkl) \delta(h-h') \delta(k-k') \delta(l-l'). \tag{26}$$

Therefore equation (26) can be written as

$$C(hkl) = V^{-1} \int \Phi(\vec{r}) \exp(2\pi i \vec{r}\cdot\vec{r}^{*}) d\vec{r}. \tag{27}$$

Renaming the fractional coordinates ξ, η and ζ along the crystal axes as x, y, and z in agreement with common crystallographic notation, expression (27) can be rewritten as

$$C(hkl) = V^{-1} \int \Phi(\vec{r}) \exp\{2\pi i (hx+ky+lz)\} d\vec{r}. \tag{28}$$

In essence expression (28) is the Fourier transform of equation (17); much of the theory and applications of X-ray crystallography centers around these two expressions.

(c) Phase Problem and Electron Density Representation of a Crystal

The structure factor F is another key quantity in the theory of diffraction; it is defined as follows:

$$F(hkl)=\sum_n f_n(hkl)\exp\{2\pi i(hx_n+ky_n+lz_n)\}, \tag{29}$$

where x_n, y_n, z_n are fractional components along the axial lengths $\vec{a}$, $\vec{b}$ and $\vec{c}$ respectively defining the position of an atom in the unit cell, and f_n is the atomic scattering factor of the nth atom as defined in the expression (12). The summation is over all the atoms in the unit cell. The structure factor F is the only quantity in which the positions x, y, z of the atoms in the unit cell appear; therefore its significance is unique in the theory of structure determination by diffraction methods. It can be seen that in general it is a complex quantity, and can be represented as a vector in the two dimensional Argand diagram:

$$F(hkl)=A(hkl)+iB(hkl) \tag{30}$$

where

$$A(hkl)=\sum_n f_n\cos 2\pi(hx_n+ky_n+lz_n)$$

$$B(hkl)=\sum_n f_n\sin 2\pi(hx_n+ky_n+lz_n)$$

and

$$|F(hkl)|=(A^2+B^2)^{\frac{1}{2}} \tag{31}$$

$$\phi(hkl)=\tan^{-1}\frac{B(hkl)}{A(hkl)}$$

In the above discussion and in what follows, the atomic scattering factor f is considered as a real quantity which is not in general true, but it is a good approximation. Using the last two equations of (31) the structure factor can also be written as

$$F(hkl)=|F(hkl)|\exp\{i\phi(hkl)\}=\{A^2(hkl)+B^2(hkl)\}^{\frac{1}{2}}\exp\{i\phi(hkl)\}.$$

The quantity $\phi(hkl)$ is called the phase angle and in general can assume any value between 0 and 2π. However, for a centrosymmetrical unit cell ,if the centre of symmetry is taken as the origin, for each triad of numbers (x, y, z) there is a centrosymmetric equivalent $(\bar{x}, \bar{y}, \bar{z})$ $(\bar{x}=-x, \bar{y}=-y, \bar{z}=-z$, again following common crystallographic notation). Therefore, for a centrosymmetrical crystal the expression (29) for the structure factor can be split into two parts

$$F(hkl)=\sum_{n=1}^{N/2} f_n(hkl)\exp\{2\pi i(hx_n+ky_n+lz_n)\}+\sum_{n=1}^{N/2} f_n(hkl)\exp\{2\pi i(h(-x)+$$
$$+k(-y)+l(-z))\}= 2\sum_{n=1}^{N/2} f_n(hkl)\cos 2\pi(hx_n+ky_n+lz_n). \quad (32)$$

Equation (32) shows that for a centrosymmetrical crystal the structure factor is a <u>real</u> quantity. From expressions (31) and (32) we deduce that

$$A(hkl)=2\sum_{n=1}^{N/2} f_n(hkl)\cos 2\pi(hx_n+ky_n+lz_n)$$
$$B(hkl)=0,$$

therefore $\tan\phi(hkl)=0$.

From this last equation we obtain

$$\phi(hkl)=0 \quad \text{or} \quad \phi(hkl)=\pi. \quad (33)$$

This severe restriction on the values of the phase angles $\phi(hkl)$ means that the phases $\exp i\phi(hkl)$ can take <u>only</u> two values +1 or -1, and the structure factor can be written as $F(hkl)=(\pm)|F(hkl)|$. This result is very helpful in crystal structure analysis.

A continuous quantity $\rho(\vec{r})$ can be defined, representing the electron density distribution of the unit cell. According to the previous discussion the electron density $\rho(\vec{r})$ should be a three dimensional periodic function with respect to the lattice vector $\vec{R}$

$$\rho(\vec{r}+\vec{R})=\rho(\vec{r}),$$

and according to equation (17)

$$\rho(\vec{r})=\sum_{\vec{r}*} C_{\vec{r}*}\exp(-2\pi i\vec{r}\cdot\vec{r}*) \tag{34}$$

or

$$\rho(\vec{r})=\sum_h\sum_k\sum_l C(hkl)\exp\{-2\pi i(hx+ky+lz)\}. \tag{35}$$

However, according to equation (28)

$$C(hkl)=V^{-1}\int_0^1\int_0^1\int_0^1 \rho(\vec{r})\exp\{2\pi i(hx+ky+lz)\}dxdydz. \tag{36}$$

The structure factor $F(hkl)$ defined in equation (29) can also be written employing the continuous electron density $\rho(\vec{r})$, by simply substituting the summation over all the atoms of the unit cell, by integration of the electron distribution over the volume V. Therefore

$$F(hkl)=\int_0^1\int_0^1\int_0^1 \rho(\vec{r})\exp\{2\pi i(hx+ky+lz)\}dxdydz. \tag{37}$$

A comparison of equations (35), (36) and (37) leads immediately to the fundamental result

$$\rho(\vec{r})=V^{-1}\sum_{h}\sum_{k}^{+\infty}\sum_{l}F(hkl)\exp\{-2\pi i(hx+ky+lz)\}. \qquad (38)$$
$$\qquad\qquad\qquad

$$\Phi(\vec{p}_1,\ldots,\vec{p}_z;t)=h^{-3z/2}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}\Psi(\vec{r}_1,\ldots,\vec{r}_z;t)\exp\{i/h\sum_j \vec{p}_j\cdot\vec{r}_j\ d\vec{r}_1\ldots d\vec{r}_z. \tag{41}$$

Equations (40) and (41) can be written in a more compact form as follows

$$\Psi(\vec{X};t)=h^{-3z/2}\int_{-\infty}^{+\infty}\Phi(\vec{P};t)\exp(-i/h\vec{P}\cdot\vec{X})d\vec{P} \tag{42}$$

$$\Phi(\vec{P};t)=h^{-3z/2}\int_{-\infty}^{+\infty}\Psi(\vec{X};t)\exp(i/h\vec{P}\cdot\vec{X})d\vec{X} \tag{43}$$

where the meaning of the new symbols is obvious.

In general, the wave function $\Psi(\vec{X};t)$ is complex and has a magnitude $|\Psi|$ and a phase angle ϕ. However, in the treament of stationary states with real Hamiltonians, the energy eigenfunction Ψ can be chosen to be real, therefore

$$\begin{aligned}\Psi(\vec{X})^{\dagger}&=\Psi(\vec{X})\\ \Phi(\vec{P})^{\dagger}&=\phi(-\vec{P})\end{aligned} \tag{44}$$

where the dagger means complex conjugation. If now we associate the wave function $\Psi(\vec{X};t)$ with the electron density $\rho(\vec{r})$, and the wave function $\Phi(\vec{P};t)$ in momentum space with the structure factor F(hkl) (which really is a quantity in momentum or $\vec{K}$ space) we can see the analogy in these quantities

$$\left(\begin{array}{ccc}\Psi(\vec{X};t) & \xleftrightarrow[\textit{transformation}]{\textit{Fourier}} & \Phi(\vec{P};t)\\ \rho(\vec{r}) & \xleftrightarrow[\textit{transformation}]{\textit{Fourier}} & F(hkl)\end{array}\right)$$

We also have

$$\rho(r)^{\dagger}=\rho(r)$$
$$F(hkl)^{\dagger}=F(\bar{h}\bar{k}\bar{l}) \tag{45}$$

in complete analogy with relations (44). In order to obtain the wave function $\Psi(\vec{X};t)$ from equation (42) we need to know <u>both</u> the amplitude and the phase of the momentum function $\Phi(\vec{P},t)$. Experimentally the only quantity measured is $|\Phi(\vec{P},t)|$ in striking analogy with the quantity $|F(hkl)|$ which is obtained experimentally in X-ray crystallography. The phase of the wave function cannot be obtained directly from the solution of Schrodinger equation due to the nature of equation itself. Only a knowledge of the initial conditions $\Psi(\vec{X};t_0)$ can provide the complete wave function, obtained by a unitary transformation

$$\Psi(\vec{X};t)=\mathcal{U}(t,t_0)\Psi(\vec{X};t_0),$$

where $\mathcal{U}$ is an evolution operator. Of course the quantum mechanical phase problem is of a dynamical nature, in contrast to the crystallographic phase problem which is a static one; a knowledge of the initial conditions in the latter case means a knowledge of the charge distribution in space, that is, the structure itself (for a short discussion of phase problem in quantum mechanics see reference 26).

(d) <u>Bragg's Law and Intensity Relations</u>

Consider a beam of X-ray photons impinging on a crystal; the electromagnetic waves interact strongly with the negatively charged electrons, and if the appropriate geometrical conditions are fulfilled, the crystal emits light mainly of the same wavelength as that of the incident radiation. This emission is usually called "reflection" and

the geometrical conditions are described by Bragg's law which can be derived as follows. If the direction of the incident and diffracted beams are specified by the unit vectors $\hat{s}_0$ and $\hat{s}$ respectively (Figure 1), the following equations give the directions and energy of the two beams:

$$\vec{\kappa}_0=\hat{s}_0\frac{1}{\lambda}, \quad \vec{\kappa}=\hat{s}\frac{1}{\lambda}$$

$$|\vec{\kappa}_0|=|\vec{\kappa}|=\kappa=\frac{1}{\lambda}$$

$$E=\frac{hc}{\kappa} .$$

Employing the first Born approximation, we can describe a transition with the matrix element[27]

$$M_{\vec{\kappa}\vec{\kappa}_0}=(\psi_{\vec{\kappa}}(\vec{r}),\rho(\vec{r})\psi_{\vec{\kappa}_0}(\vec{r})), \tag{46}$$

where $\psi_{\vec{\kappa}}$ and $\psi_{\vec{\kappa}_0}$ are the incident and diffracted waves respectively. We can consider that they are adequately described by plane waves, therefore

$$\begin{aligned}\psi_{\vec{\kappa}}&=C_1\exp(2\pi i\vec{\kappa}\cdot\vec{r})\\ \psi_{\vec{\kappa}_0}&=C_2\exp(2\pi i\vec{\kappa}_0\cdot\vec{r}) \quad ,\end{aligned} \tag{47}$$

where C_1, C_2 are constants. Expanding the electron density $\rho(\vec{r})$ in a Fourier series according to equation (28), equation (46) becomes

$$\begin{aligned}M_{\vec{\kappa}\vec{\kappa}_0}&=\int C_1^{\dagger}\exp(-2\pi i\vec{\kappa}\cdot\vec{r})V^{-1}\sum_{hkl}F(hkl)\exp(2\pi i\vec{r}\cdot\vec{r}^*)C_2\exp(2\pi i\vec{\kappa}_0\cdot\vec{r})d\vec{r}=\\ &=C_1^{\dagger}C_2V^{-1}\sum_{hkl}F(hkl)\int\exp\{2\pi i(\vec{\kappa}_0-\vec{\kappa}+\vec{r}^*)\cdot\vec{r}\}d\vec{r}\end{aligned} \tag{48}$$

The integral (48) is different than zero only if

$$\vec{\kappa}_0-\vec{\kappa}+r^*=0$$

or

$$\vec{\kappa}=\vec{\kappa}_0+\vec{r}^*. \tag{49}$$

For fixed $\vec{\kappa}$ we can only observe diffracted beams in directions dictated by the right hand side of equation (49). Under those conditions we obtain from equation (48)

$$M_{\vec{\kappa}\vec{\kappa}_0}=C_1^{\dagger}C_2V^{-1}F(hkl)$$

or

$$|M_{\vec{\kappa}\vec{\kappa}_0}|^2=|C_1C_2V^{-1}|^2|F(hkl)|^2. \tag{50}$$

Equation (50) shows that the intensity of the diffracted radiation should be proportional to $|F(hkl)|^2$, which is more or less true for imperfect crystals.

For a static crystal the energy of the diffracted beam is the same as the incident beam, i.e., $|\vec{\kappa}_0|=|\vec{\kappa}|$; this last condition imposes immediately a condition on the scattering angle θ. To satisfy the geometrical conditions of equation (49) and the isoenergetic condition $|\vec{\kappa}_0|=|\vec{\kappa}|$ at the same time, we construct the Ewald sphere (Figure 2), with radius CO equal to the incident wave vector. The lattice shown is the reciprocal lattice, C is the position of the diffracting crystal and O the origin. From Figure 2 and equation (49) we obtain:

$$|\vec{\kappa}-\vec{\kappa}_0|=|\vec{r}^*|=|(h\vec{a}^*+k\vec{b}^*+l\vec{c}^*)|=2\frac{1}{\lambda}\sin\theta(hkl).$$

However, $|\vec{r}^*(hkl)|=\frac{1}{d(hkl)}$, where d(hkl) is the distance between the "reflecting" planes (hkl); from these two equations we obtain

$$2d(hkl)\sin\theta(hkl)=\lambda \tag{51}$$

which is Bragg's law in its most convenient form.

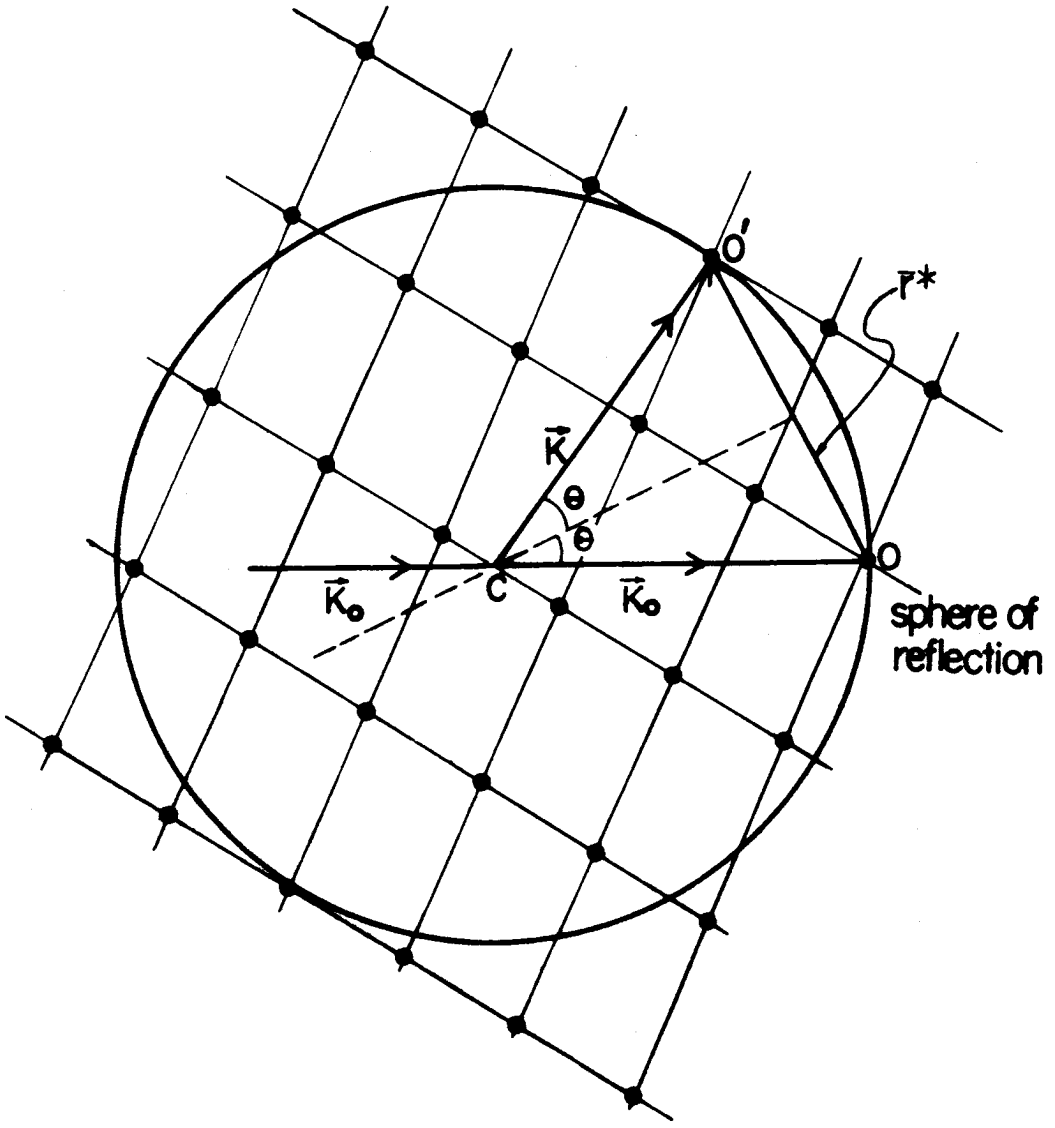


Figure 2. Ewald's construction.

We consider now a pencil of monochromatic X-ray beams, polarized perpendicular to the plane of the paper of intensity I_0 and wavelength λ, falling on a very small crystal. The conditions are illustrated in Figure 1; the crystal is located at the origin O, the vector $\vec{r}_n$ represents the location of the nth atom in a certain unit cell, and the directions of primary and diffracted beams are described by the unit vectors $\hat{s}_0$ and $\hat{s}$ respectively. For simplicity it will be assumed that the crystal has the shape of parallelopipedon with edges $N_a|\vec{a}|$, $N_b|\vec{b}|$, $N_c|\vec{c}|$ parallel to the crystal axes $\vec{a}$, $\vec{b}$, $\vec{c}$. Assuming also that $|\vec{r}_n|<<|\vec{R}|$, which holds for a small crystal, the instantaneous electric field at the observation point P, according to Warren[17], will be given by

$$\vec{E}=\vec{E}_0\frac{e^2}{mc^2|\vec{R}|}\exp\{i(2\pi\kappa|\vec{R}|-\omega t)\}F\,\frac{\exp\{2\pi i\kappa(\hat{s}-\hat{s}_0)\cdot N_a\vec{a}\}-1}{\exp\{2\pi i\kappa(\hat{s}-\hat{s}_0)\cdot\vec{a}\}-1}\times$$

$$\times\frac{\exp\{2\pi i\kappa(\hat{s}-\hat{s}_0)\cdot N_b\vec{b}\}-1}{\exp\{2\pi i\kappa(\hat{s}-\hat{s}_0)\cdot\vec{b}\}-1}\times\frac{\exp\{2\pi i\kappa(\hat{s}-\hat{s}_0)\cdot N_c\vec{c}\}-1}{\exp\{2\pi i\kappa(\hat{s}-\hat{s}_0)\cdot\vec{c}\}-1}\ . \qquad (52)$$

The diffracted intensity at the point P will be given by

$$I_p=\frac{|\vec{E}|^2}{8\pi} \qquad (53)$$

in terms of the intensity in the primary beam

$$I_0=\frac{|\vec{E}_0|^2}{8\pi}\ . \qquad (54)$$

From equations (52)-(54) and for an unpolarized beam, we obtain for the diffracted intensity

$$I_p=I_0\left(\frac{e^2}{mc^2}\right)^2\frac{1}{|R|^2}|F|^2\,\frac{1+\cos^2 2\theta}{2}\times\frac{\sin^2\pi/\lambda(\hat{s}-\hat{s}_0)\cdot N_a\vec{a}}{\sin^2\pi/\lambda(\hat{s}-\hat{s}_0)\cdot\vec{a}}\times$$

$$\times\frac{\sin^2\pi/\lambda(\hat{s}-\hat{s}_0)\cdot N_b\vec{b}}{\sin^2\pi/\lambda(\hat{s}-\hat{s}_0)\cdot\vec{b}}\times\frac{\sin^2\pi/\lambda(\hat{s}-\hat{s}_0)\cdot N_c\vec{c}}{\sin^2\pi/\lambda(\hat{s}-\hat{s}_0)\cdot\vec{c}}\ . \qquad (55)$$

Expression (55) assumes significant values only if

$$(\hat{s}-\hat{s}_o)\cdot\vec{a}=\lambda h$$
$$(\hat{s}-\hat{s}_o)\cdot\vec{b}=\lambda k \qquad (56)$$
$$(\hat{s}-\hat{s}_o)\cdot\vec{c}=\lambda l$$

Equations (56) are called the Laue equations and are entirely equivalent with Bragg's law. Under the conditions of (56) equation (55) becomes

$$I_{P(Max)}=I_o\left(\frac{e^2}{mc^2}\right)^2 \frac{1}{|\vec{R}|^2} N_a^2N_b^2N_c^2|F(hkl)|^2 . \qquad (57)$$

However, for rather technical reasons, $I_{P(Max)}$ is not a very useful quantity. The primary beam is never perfectly parallel, and therefore it is never true that all of the primary beam has the correct direction $\hat{s}_o$ to satisfy the three Laue equations exactly. In general, a crystal is "mosaic" and an exact setting on one part of the crystal would not be correct for other parts[17]. A more useful quantity which can be measured relatively easy is the "integrated intensity", given by

$$E=\iint I_P dt dA, \qquad (58)$$

where I_P is given by equation (55), t is the duration of measurement and A is the area crossed by the diffracted beam. The evaluation of the above integral is not trivial and we quote only the result here[17]

$$E(hkl)=\frac{I_o}{\Omega}\left(\frac{e^2}{mc^2}\right)^2 \frac{\lambda^3 V_x}{V^2}|F(hkl)|^2 \frac{1+\cos^2 2\Theta}{2\sin 2\Theta} , \qquad (59)$$

where V_x is the volume of the small crystal, Ω is the angular velocity with which the crystal is rotated through the Bragg angle and $\frac{1}{2\sin 2\Theta}$ is a geometrical term (the quantity $\frac{1+\cos^2 2\Theta}{2\sin 2\Theta}$ is called the Lorentz-Polarization term). The important thing of equation (59) is that the

reflected energy E(hkl) is proportional to the square of the structure amplitude $|F(hkl)|$. It applies equally to very small crystals, for which dynamical effects are almost absent, or for ideally imperfect crystals. Equation (59) can be rewritten as

$$|F(hkl)|^2 LP = \frac{E(hkl)\Omega}{I_o}\left(\frac{mc^2}{e^2}\right)^2 \frac{V^2}{\lambda^3 V_x}$$

or

$$|F(hkl)|^2 LP = Q(hkl), \qquad (60)$$

where $LP = \frac{1+\cos^2 2\theta}{2\sin 2\theta}$.

The dimensionless quantity $Q(hkl)$ is the measurement in a diffraction experiment, and therefore the <u>relative</u> structure amplitudes can be obtained directly from (60)

$$|F(hkl)|_{rel} = \left\{\frac{Q(hkl)}{LP}\right\}^{\frac{1}{2}} . \qquad (61)$$

Although equation (61) is almost always used in structure determinations, care should be exercised if very accurate structure amplitudes are required, because dynamical effects are always present and sometimes can completely disrupt the validity of the last equation.

(e) <u>The Patterson Function</u>

The most severe obstacle in the investigation of a crystal structure is that the phases of F(hkl) cannot be experimentally observed. However, in 1934-35 A. L. Patterson[28,29] approached this problem in a very ingenious way, and showed that a structure could be solved, or that much of the information concerning the phase problem can be obtained by a direct use of the experimentally obtainable structure amplitudes.

The Patterson function $P(\vec{X})$ is defined as the self-convolution of the electron density $\rho(\vec{r})$

$$P(\vec{X}) \equiv V\int \rho(\vec{r})\rho(\vec{r}+\vec{X})d\vec{r}. \tag{62}$$

The electron density $\rho(\vec{r})$ can be written in an integral form as

$$\rho(\vec{r})=V^{-1}\int F(\vec{r}'^{*})\exp(-2\pi i\vec{r}\cdot\vec{r}'^{*})d\vec{r}'^{*}. \tag{63}$$

A density $\rho(\vec{r}+\vec{X})$ at a point $\vec{r}+\vec{X}$ will be given by

$$\rho(\vec{r}+\vec{X})=V^{-1}\int F(\vec{r}^{*})\exp\{-2\pi i(\vec{r}+\vec{X})\cdot\vec{r}^{*}\}d\vec{r}^{*}. \tag{64}$$

Substituting (63) and (64) into (62) we obtain

$$P(\vec{X})=V^{-1}\iint F(\vec{r}'^{*})\exp(-2\pi i\vec{r}\cdot\vec{r}'^{*})d\vec{r}'^{*}\int F(\vec{r}^{*})\exp\{-2\pi i(\vec{r}+\vec{X})\cdot\vec{r}\}d\vec{r}^{*}d\vec{r}=$$

$$=V^{-1}\iint F(\vec{r}'^{*})F(\vec{r}^{*})\exp(-2\pi i\vec{X}\cdot\vec{r}^{*})d\vec{r}'^{*}d\vec{r}^{*}\int\exp\{-2\pi i(\vec{r}^{*}+\vec{r}'^{*})\cdot\vec{r}\}d\vec{r}$$

or

$$P(\vec{X})=V^{-1}\iint F(\vec{r}'^{*})F(\vec{r}^{*})\exp(-2\pi i\vec{X}\cdot\vec{r}^{*})\delta(\vec{r}^{*}+\vec{r}'^{*})d\vec{r}'^{*}d\vec{r}^{*}. \tag{65}$$

Expression (65) is other than zero only if $\vec{r}'^{*}+\vec{r}^{*}=o$, or $\vec{r}'^{*}=-\vec{r}^{*}$, therefore

$$P(\vec{X})=V^{-1}\int F(\vec{r}^{*})F(-\vec{r}^{*})\exp(-2\pi i\vec{X}\cdot\vec{r}^{*})d\vec{r}^{*}=V^{-1}\int|F(\vec{r}^{*})|^{2}\exp(-2\pi i\vec{X}\cdot\vec{r}^{*})d\vec{r}^{*} \tag{66}$$

because $F(-\vec{r}^{*})=F(\vec{r}^{*})^{\dagger}$.

If the fractional components of vector $\vec{X}$ are u, v and w along the crystallographic axes, equation (66) can also be written as

$$P(uvw)=V^{-1}\sum_{h}\sum_{k}^{+\infty}\sum_{l}|F(hkl)|^{2}\exp\{-2\pi i(hu+kv+lw)\}=$$

$$V^{-1}\sum_{h}\sum_{-\infty}^{+\infty}\sum_{l}|F(hkl)|^{2}\cos 2\pi i(hu+kv+lw). \tag{67}$$

From equation (67) it can be seen that the Patterson function is a cosine Fourier series, hence centrosymmetric, whose coefficients can be directly determined from experiment. As a crystal may be regarded

as a built up from N atoms, each of which is repeated on an infinite lattice, the Patterson fuction consists of N^2 "Patterson Peaks" repeated in the same way[15]. However, peaks in the Patterson function do not represent the position of atoms; they are the terminal points of a set of vectors, each vector representing the displacement of some atom from some other atom. The deconvolution of a Patterson synthesis is not a simple task and special techniques have been developed for this purpose[30].

(f) The Temperature Factor

It was tacitly assumed up to now that all the atoms in a crystal occupy definite positions. However, these positions are only average positions about which atoms oscillate. The experimentally obtained structure amplitudes correspond to a double average; one over the whole crystal and another over the time of the measurement, the latter being very large with respect to the period of vibration of an atom.

It has been proven by Bloch[31] that if an atom vibrates at any temperature T in a harmonic potential, the probability density of its displacement is given by the Gaussian

$$G(x_1x_2x_3)= \frac{1}{(2\pi)^{3/2}u_1u_2u_3} \exp\{-\frac{1}{2}(\frac{x_1^2}{u_1^2} + \frac{x_2^2}{u_2^2} + \frac{x_3^2}{u_3^2})\}, \qquad (68)$$

where x_1, x_2, x_3 are displacements along orthogonal axes coincident with the principal axes of the family of ellipsoids

$$\frac{x_1^2}{u_1^2} + \frac{x_2^2}{u_2^2} + \frac{x_3^2}{u_3^2} = 2\Gamma \quad (\Gamma>o), \qquad (69)$$

which represent surfaces of constant probability in direct space.

u_1^2, u_2^2, u_3^2 are the mean square displacements in the x_1, x_2, x_3 directions respectively; the mean square displacement is defined by

$$\langle u_i^2\rangle = \iiint_{-\infty}^{+\infty} x_i^2 G(x_1 x_2 x_3)\,dx_1 dx_2 dx_3, \qquad i=1,2,3.$$

Due to the convolution theorem which says that "the Fourier inverse transform of a product of Fourier transforms is the convolution of the original functions," the atomic scattering factor at temperature T will be given by

$$f_T(hkl)=f(hkl)\exp\{-2\pi^2(u_1h_1^2+u_2h_2^2+u_3h_3^2)\}, \tag{70}$$

where the exponential factor is the Fourier transform of expression (68), and h_1, h_2, h_3 are the projections of the reciprocal lattice vector $\vec{r}^*$ on axes parallel to the principal axes of the ellipsoid[15]. If $u_1^2=u_2^2=$ $=u_3^2=\langle u^2\rangle$, expression (70) transforms to

$$f_T(hkl)=f(hkl)\exp\{-2\pi^2\langle u^2\rangle|\vec{r}^*|^2\},$$

where $$|\vec{r}^*| = \frac{1}{d} = \frac{2\sin\theta}{\lambda}\ ;\ \text{therefore,}$$

$$f_T(hkl)=f(hkl)\exp\{-B(\sin\theta/\lambda)^2\} \tag{71}$$

where $B=8\pi^2\langle u^2\rangle$ is the Debye-Waller factor. If it is considered the same for all atoms and all directions in a crystal, it is given by

$$B=\frac{6h^2T}{mk\theta^2}\{(T/\theta)\int_0^{\theta/T}\frac{\xi d\xi}{\exp\xi-1} + (\theta/4T)\},$$

where k is Boltzmann's constant, m an average mass, and θ is the "characteristic temperature" defined as $k\theta=h\nu_D$, with ν_D the Debye frequency. It should be noticed that the exponential factor of

equation (71) is wavelength independent.

When referred to the reciprocal axes of the crystal, the anisotropic temperature factor takes the form[15]

$$\exp\{-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{23}kl+\beta_{13}hl)\}, \tag{72}$$

where the β's are constants to be determined by the method of least squares in a refinement procedure. Expression (72) can also be written as[15]

$$\exp\{-2\pi^2(U_{11}(h|\vec{a}^*|)^2+U_{22}(k|\vec{b}^*|)^2+U_{33}(l|\vec{c}^*|)^2+2U_{12}(hk|\vec{a}^*||\vec{b}^*|)^2+$$

$$+2U_{13}(hl|\vec{a}^*||\vec{c}^*|)^2+2U_{23}(kl|\vec{b}^*||\vec{c}^*|)^2)\}. \tag{73}$$

The U_{ij} are the elements of a symmetric tensor $\overleftrightarrow{U}$ and form a matrix

$$\mathbf{U} = \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ U_{12} & U_{22} & U_{23} \\ U_{13} & U_{23} & U_{33} \end{pmatrix} . \tag{74}$$

By comparison of (72) and (73) we obtain $U_{11} = \dfrac{\beta_{11}}{2\pi^2|\vec{a}^*|^2}$ etc.

The problem now reduces to finding an orthogonal set of coordinates in which the matrix (74) has zero for all off-diagonal elements, hence reducing the expression (73) to the form which appears in equation (70). The rules for this are given by Cruickshank et. al.[32], and outlined by Lipson and Cochran[15]. The final result is that of the values of u_1^2, u_2^2, u_3^2, the mean square displacements along the directions of the principal axes.

II. PORPHYRINS

Porphyrins are macrocyclic tetrapyrrolic structures; they are derived formally from porphine (Figure 3) by substitution of some or all of the hydrogen atoms 1-8 by various side chains. The number of substitutions which can be obtained, and the names which have been assigned to the resultant porphyrins are bewildering, and many times serious errors creep into the porphyrin literature[33]. The name porphyrin will be used in this general discussion indiscriminately, either for porphyrins or porphines (for the porphyrin nomenclature see reference 34).

The biological importance of the porphyrin ring system is obvious from the fact that it is a basic constituent of the blood, and of the plant pigments, chlorophyll and bacteriochlorophyll. Porphyrins have a remarkable ability to form complexes with a variety of metal ions acting as tetradentates, to produce "metalloporphyrins", which have the most diverse biological abilities ever encountered in chemical systems (reference 35, p. 7).

Many porphyrin-riddles have been answered since 1912 when Küster first suggested the tetrapyrrolic structure of these molecules[36], but as usual, nature plays its asymptotic game with knowledge, and much experimental and theoretical work has yet to be carried out in this field.

The question of the existence of different -NH tautomers of porphyrins was long argued, and much of the early work aimed at

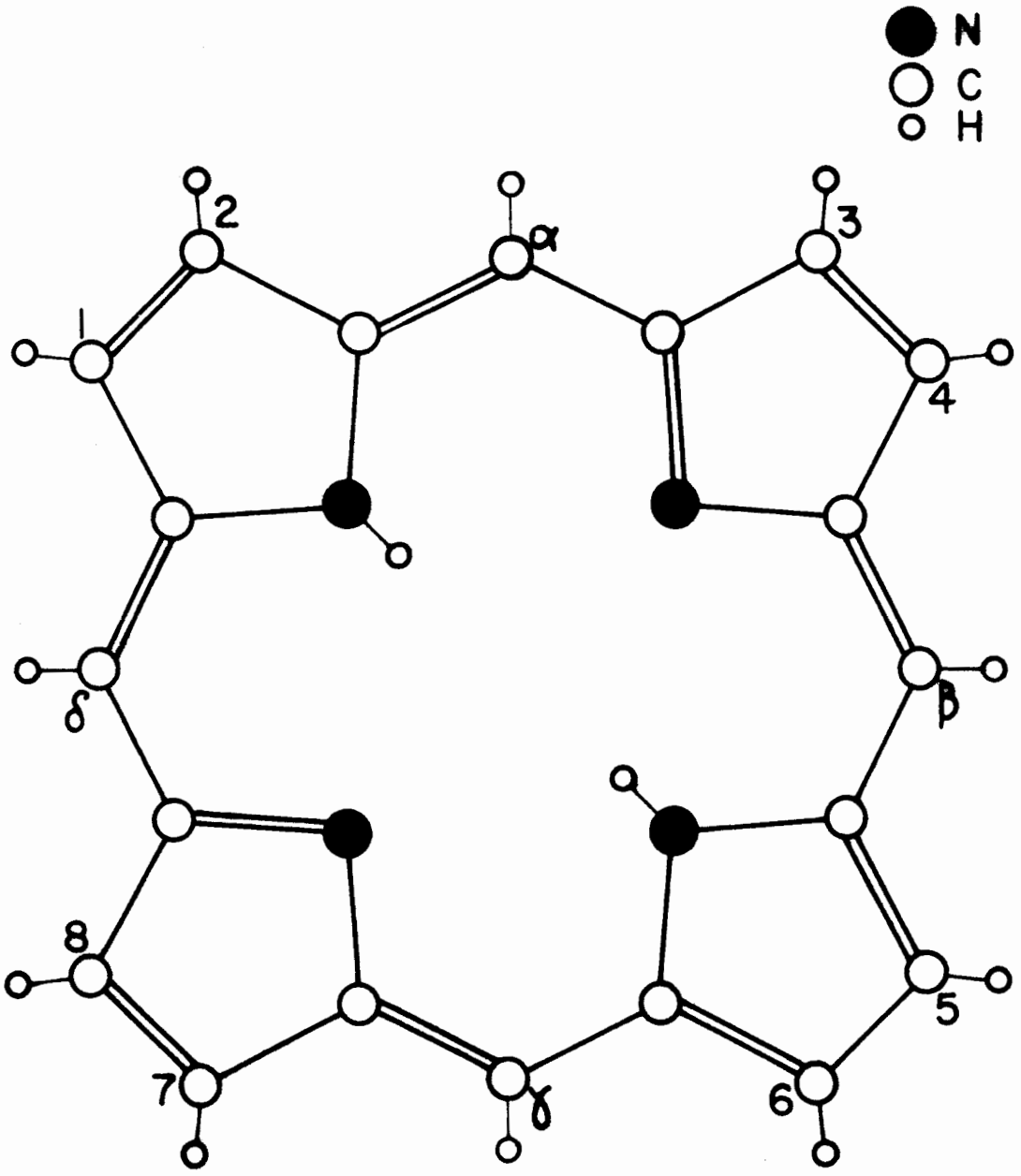
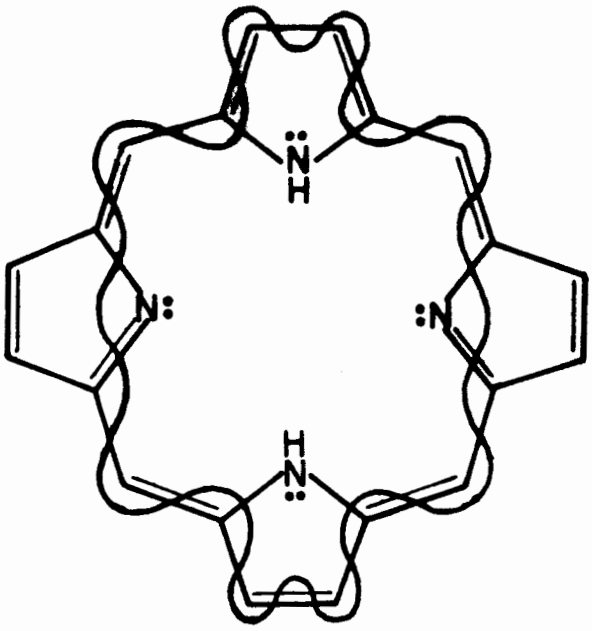


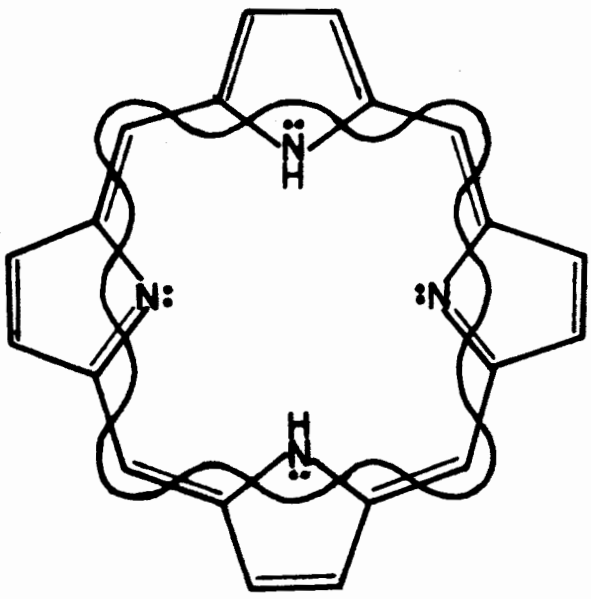
Figure 3. Free base porphine.

isolating separate tautomeric forms has been conveniently summarized[37]. A definite answer to this problem (at least in the solid state) was given with the determination of the crystal structures of tetraphenylporphine and of the free base porphine by Silvers and Tulinsky[38] and by Chen and Tulinsky[39] respectively. They found two well-defined imino hydrogen atoms located on pyrroles opposite to each other; the same imino hydrogen localization was also found by Codding and Tulinsky in the crystal structure of tetra-n-propylporphine[40].

Porphyrins and metalloporphyrins are in general planar molecules showing D_{4h} symmetry. It can also be said that they are aromatic, because they can exhibit a strong diamagnetic ring current caused by the "circulation" of the π-electrons (the two imino hydrogens of porphyrins are magnetically shielded and they show NMR shifts comparable to those of 18 annulenes[41-44]). The "path" of this electron movement is not well defined and at least two structures are compatible with some of the experimental facts, both of which are appealing to the chemical intuition of the experimental organic chemist. These two possibilities are illustrated in Figure 4; in Figure 4b,16 atoms form a closed ring (wavy line) in which 6 double bonds and 18 electrons participate. This ring cannot be a conjugate cyclic polyene since a 16-membered polyene has only 16_π electrons. The 6 double bonds contain 12_π electrons and the addition of both the lone-electron pairs would only bring the total electron count to 16. To arrive at an 18-electron count it is necessary to interrupt the conjugation, placing two positive charges on the nitrogens bonded to hydrogens, and assigning the two balancing countercharges to the remaining 14 central ring atoms (reference 35, p. 201). On the other hand ,the model on Figure 4a is an 18-atom system with 9 double bonds and thus 18 electrons and it



(a)



(b)

Figure 4. Two different "electron paths" for the porphyrin system.

possesses Kekule´ type resonance. The feature of this model is that Huckel's 4n+2 rule imposes the requirement that the lone electron pairs must be different from the π-electrons of the system (reference 35, p. 201). If the two nitrogen atoms involved are regarded to a first approximation as equivalent to methine groups, the 18 annulene could be taken as a model to describe porphyrins. However, both of the described models cannot account in a consistent way for the crystallographically obtained bond distances or absorption and emission spectra. A linear mixing of both forms would be more realistic but such an approach is poor from a model point of view.

The absorption spectra of porphyrins have a very typical pattern. Almost all of them have four bands between 500 and 700nm, generally increasing in intensity from red to green, and a very strong band in the blue-violet region ($\sim$400nm) called the Soret band[45]. Gouterman and his coworkers have published a long series of papers[46-68], both theoretical and experimental trying to explain the origin and the intensity of the forementioned bands, and the changes to these bands accompanied by metal complexation. However, their efforts have been hampered by the great complexity of the porphyrin system which leads inevitably to the use of highly approximate semiempirical methods (extended Hückel calculations) and therefore unreliable results. A very simple model based on the "particle in a circle" idea proposed as early as 1949 by Simpson[69] can explain in a qualitative manner the porphin spectrum; it seems that it is better in this case to use a very simple model whenever possible, than a complicated one based on various approximations and assumptions.

Although the crystal structures of more than forty metallo-porphyrins have been solved in the past fifteen years, it has only

been recently that some interest was shown to the Group IV metalloporphyrins. Thus, independent X-ray crystallographic structure determinations have been reported for dichloridetetraphenylporphineSn(IV)[70], and dichlorideoctaethylporphineSn(IV)[71] and extended Hückel calculations have also been carried out with Si, Ge, Sn and Pb metalloporphyrins by Gouterman and his coworkers[62].

We undertook the investigation of an octahedrally germanium-substituted porphine for a number of reasons. A relatively accurate structure could be obtained because the electron content of the germanium is sufficiently small. The radius of the germanium atom in oxidation state IV is small enough to be accomodated in the central hole of the porphine core with minimal perturbation. Coordinates from this structure could then be used in theoretical calculations. This would also be the first metallo-porphyrin without a substitution in the porphine frame (Figure 3), and therefore a direct comparison could be made with the free base porphine[39]. In addition, interest is centered about the metallic character of germanium, which is not very pronounced. For instance,it can easily form tetrahedral compounds with sp^3 hybridization in a manner similar to silicon or carbon[72], or it can be octahedral, as in the present system with a possible d^2sp^3 hybridization.

III. EXPERIMENTAL

1. Crystals

Dihydroxyporphinato Ge(IV) ($PGe(OH)_2$) used in the X-ray crystallographic structure determination was purchased from the M and J Chemical Co.[73]. Well formed purple crystals exhibiting a rectangular-prismatic morphology were obtained by diffusing methanol into a nearly saturated solution of the compound in chloroform. At the time, we were unaware that the hydroxyl groups of $PGe(OH)_2$ are quite reactive and can easily be replaced with methoxy groups by treatment with methanol[74]. Indeed, the structure determination of crystals grown this way showed that the $PGe(OH)_2$ had been converted to dimethoxyporphinato Ge(IV) ($PGe(OMe)_2$) during the crystallization procedure.

2. Intensity Data Collection

Preliminary X-ray studies of a single crystal of $PGe(OMe)_2$ by photographic[75,76] and diffractometric[77] techniques showed the crystal system to be monoclinic and systematic absences fixed the space group to be $P2_1/c$. A suitable crystal fragment with approximate dimensions of 0.08x0.10x0.35mm (crystal volume $V_x \approx 2 \times 10^{-6} cm^3$) was used for recording diffracted intensities. The crystal was much smaller than the cross section of X-ray beam, and although small, of sufficient size and quality to produce a strong diffraction pattern. The lattice parameters were obtained from diffractometer measurements by the least squares fit of the angular coordinates of twelve reflections

well distributed in the reciprocal space in the range $52^{o}<2\theta<90^{o}$. The density of the crystal was measured in a mixture of carbon tetrachloride and methylene iodide. The crystal and unit cell data are summarized in Table I.

The intensity data collection was carried out with CuKα radiation (λ=1.5418A) using a Picker four circle diffractometer controlled by a Digital Equipment Corporation (DEC) 4K PDP-8 computer (FACS I system) coupled to a DEK 32K Disc File and an Ampex TMZ 7-Track Tape Transport. Intensities of reflections were measured by a wandering ω-step-scan procedure using balanced Ni/Co filters[78]. The purpose of Ni/Co balanced filters is to supress the unwanted components of the X-ray beam[79]. The step scan was performed in 0.03^{o} increments of the ω-angle of the diffractometer and extended $\pm 0.075^{o}$ on either side of the calculated peak position. Each step was measured for a duration of four seconds and the four largest measurements were summed to give the integrated intensity of the reflection[80]. When the observed peak position did not coincide with the calculated ω-value, one or two additional steps were taken to assure centering of the scan. The background was measured with a Co filter at the ω-value of the maximum intensity for a time interval of four seconds and this count was multiplied by four to give the total background intensity. Since the step scan procedure is essentially a stationary crystal-stationary counter measurement, in order to avoid Kα splitting effects, the intensity data collection was confined within the range $2\theta<110^{o}$. Therefore the minimum interplanar spacing d_m given by Bragg's law will be

$$d_m = \frac{1.5418}{2\sin 55} \approx 0.94 \text{ A}$$

TABLE I. Crystal and Unit Cell Data of $PGe(OMe)_2$

$\lvert\vec{a}\rvert$	=	15.015(5) A
$\lvert\vec{b}\rvert$	=	14.441(5) A
$\lvert\vec{c}\rvert$	=	8.414(4) A
β	=	$91.85(2)^{\circ}$
z	=	4
MW	=	443.0 a.u.
D_c	=	1.61 $gmcm^{-3}$
D_o	=	1.60 $gmcm^{-3}$
μ	=	27.33 cm^{-1}
μl_{mx}	=	0.96 (longest direction)
F(000)	=	904 electrons
$\frac{F(000)}{V}$	=	0.495 eA^{-3}

so that details on scale smaller than

$$r=0.61xd_m=0.61x0.94\simeq0.57A$$

cannot be resolved (reference 14, p. 400). In this, it is assumed, of course, that a complete series of reflections will be measured to the limiting value d_m.

During the intensity data collection, the alignment of the crystal was monitored with the use of an automatic realignment sub-routine by measuring the intensities of three standard reflections: (006) at $X=90^\circ$ and two Φ-values 100° apart, and the ($\bar{8}42$) reflection[78]. The monitored reflections also served to monitor X-ray damage to the crystal; no decrease in their intensities was observed. Before the onset of intensity data collection, the mosaic spreads of two reflections were measured to ensure crystal quality and to help select the quadrant to be used for data collection.

The intensities of the reflections were corrected for absorption and lack of balance[78]. For the absorption correction, an empirical method was used based on the variance of the relative transmission (T) with the azimuthal angle Φ[81] (the problem of absorption is discussed more completely in the next section). The intensities of a total of 2396 independent reflections were recorded, of which 598 were taken to be unobserved; 120 reflections were systematically absent. The observable limit was fixed from the average value of the measured intensities of the systematically absent reflections and this gave 1798 reflections for the structure analysis; this number corresponds to a data/parameter ratio of 6.7 if hydrogen atoms are not included (9 parameters for each non-hydrogen atom). Finally, the corrected intensities were converted to relative structure amplitudes using equation (61).

3. Absorption

X-rays are absorbed when passed through matter mainly due to the photoelectric effect, and absorption is one of the main sources of error in structural X-ray crystallographic work. If we denote by A(hkl) the amount by which the intensity of a reflection is reduced due to absorption, the reciprocal of A(hkl) is the transmission factor, T(hkl), by which the observed intensity is to be multiplied to obtain the correct intensity. T is given by

$$T=\frac{1}{A}=\frac{V_x}{\int_{V_x}\exp\{-\mu(t_1+t_2)\}dV}, \qquad (75)$$

where μ is the linear absorption coefficient of the crystal, V_x its volume, t_1 the path of the incident beam inside the crystal and t_2 the path of the diffracted beam inside the crystal. The linear absorption coefficient can be computed by the following expression:

$$\mu=D\sum_j P_j\left(\frac{\mu}{D}\right)_j, \qquad (76)$$

where D is the density of the crystal, P_j the fraction by the weight of element j in the crystal and $(\mu/D)_j$ its mass absorption coefficient. Values of mass absorption coefficients for all the elements are tabulated in the International Tables for X-ray Crystallography Volume III[82]. For $PGe(OMe)_2$ μ is 27.33 cm^{-1} (λ=1.5418A).

For a nonspherical crystal the integral of equation (75) can only be evaluated using numerical methods[83,84]. However, these methods require the precise measurement of the crystal dimensions, which may be difficult if not impossible to achieve, especially if the crystal morphology is unfavourable or ill-defined. An empirical method for

the correction of absorption was suggested by Furnas[85] and was improved by modification by North et. al.,[81].

Suppose that a crystal is aligned at $X=90^{\circ}$ in the four-circle diffractometer; as the crystal rotates around the Φ axis (at $X=90^{\circ}$) the intensity of a reflection varies as a function of the azimuthal angular setting Φ for the corresponding reciprocal lattice level. Such a variation in intensity provides a measure of the relative absorption suffered by X-rays passing through the crystal in mean directions perpendicular to the rotation axis. According to Furnas, the relative absorption correction for a general (hkl) reflection is given by

$$A(hkl)= \frac{I_{max}(\Phi)}{I(\Phi_{hkl})} = \frac{1}{T(hkl)} , \qquad (77)$$

where I_{max} is the maximum intensity observed for an axial reflection at $X=90^{\circ}$ as Φ is varied over 360°, and Φ_{hkl} is the angle at which the hkl reflecting planes are parallel to the incident X-ray beam. In the Phillips absorption correction[81] the mean of the absorption corrections is used for beams passing through the crystal in the directions of the incident and reflected rays. Hence, in terms of the transmission factor

$$T(hkl)=\frac{T(\Phi_{inc})+T(\Phi_{ref})}{2} , \qquad (78)$$

where Φ_{inc} and Φ_{ref} define the orientations of the crystal in which the incident and reflected beams for the (hkl) reflection coincide with or lie in the same plane as the incident X-ray beam. Φ_{inc} and Φ_{ref} can be expressed in terms of Φ_{hkl},

$$\Phi_{inc}= (\Phi_{hkl}-\varepsilon_{hkl})$$

$$\Phi_{ref}=(\Phi_{hkl}+\delta_{hkl}),$$

where $$\varepsilon_{hkl}=\delta_{hkl}=\sin^{-1}\{\sin\Theta(hkl)\cos\chi(hkl)\}$$

for the equi-inclination geometry of the four-circle diffractometer.

The $T(\Phi)$ curves were constructed for $PGe(OMe)_2$ by measuring the variation of the absorption of reflections at $\chi=90^\circ$. Since the $\vec{c}^*$ axis occurred at $\chi=90^\circ$, the (001) reflections were used to correct the general reflections in terms of Φ, 2Θ and reciprocal lattice level (l-index). The absorption of three reflections was measured: (002), (004), and (006), with $2\Theta=21.13^\circ$, 43.02° and 66.73° respectively. Plots of I_{max}/I as a function of Φ are shown in Figure 5.

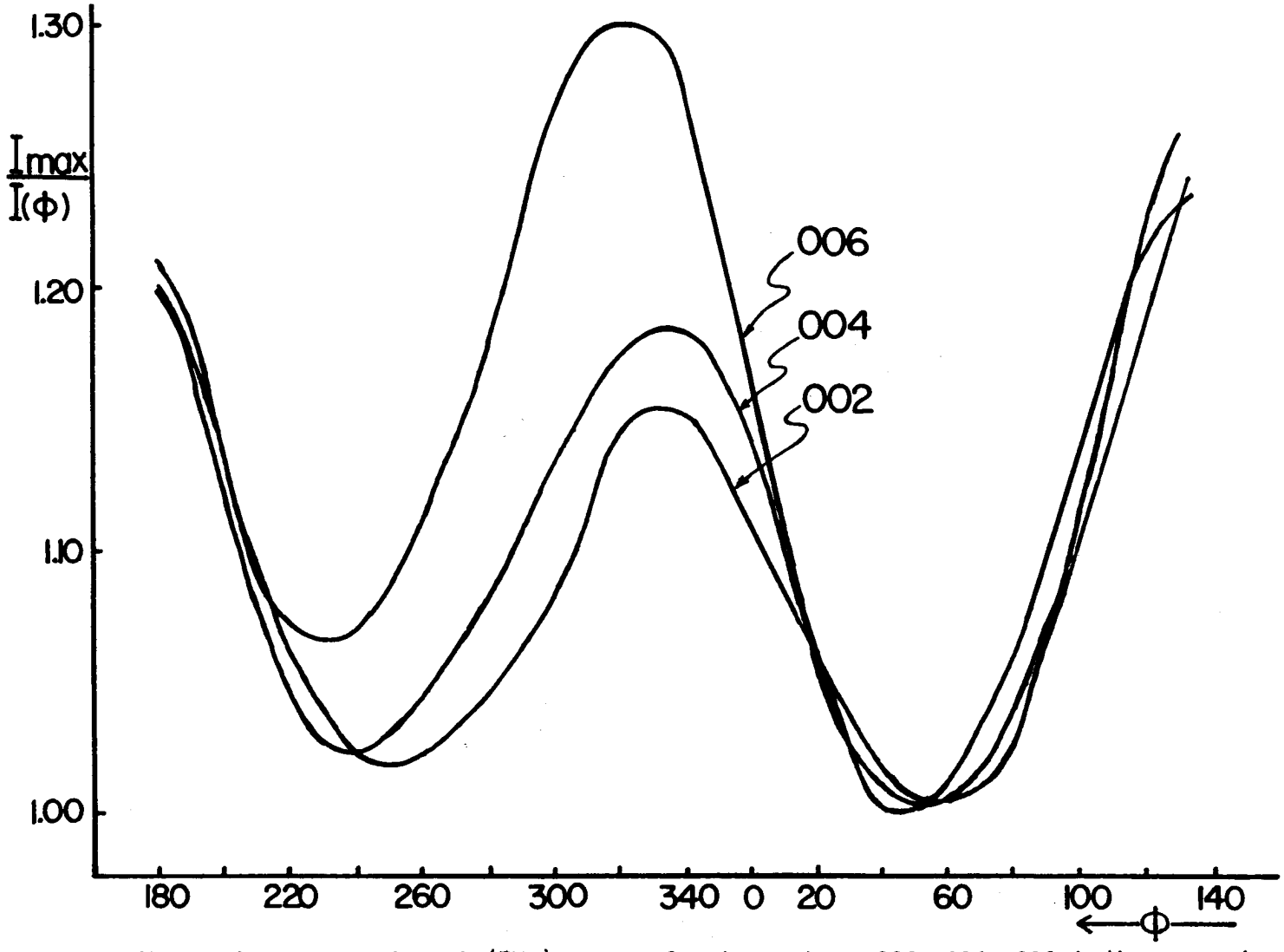


Figure 5. Absorption curves for $PGe(OMe)_2$ crystal; the numbers 006, 004, 002 indicate reciprocal lattice levels.

IV. STRUCTURE ANALYSIS

The <u>unitary</u> structure factor U is defined by

$$U(hkl)=\frac{F(hkl)}{\sum_{j}^{N} f_j(hkl)} \tag{79}$$

and the <u>normalized</u> structure factor by

$$E(hkl)=\frac{U(hkl)}{\langle|U|^2\rangle^{\frac{1}{2}}} . \tag{80}$$

From (79) and (80) we obtain

$$E(hkl)=\frac{F(hkl)\Big/\sum_{j}^{N} f_j(hkl)}{\langle|F(hkl)|^2\rangle^{\frac{1}{2}}\Big/\sum_{j}^{N} f_j(hkl)}=\frac{F(hkl)}{\langle|F(hkl)|^2\rangle^{\frac{1}{2}}} .. \tag{81}$$

From the definition of the structure factor we also obtain

$$|F(hkl)|^2=\sum_{j}^{N} f_j^2+\sum_{j\mu}^{N}{}' f_j f_\mu \exp\{2\pi i h(x_j-x_\mu)+k(y_j-y_\mu)+l(z_j-z_\mu)\} .$$

If the average of the above expression is taken over a reasonable number of F(hkl)'s with no greatly different values of $\sin\theta(hkl)/\lambda$, so that the f's can be treated as constants, the exponential terms will largely cancel[18], so that

$$\langle|F(hkl)|^2\rangle \simeq \varepsilon^2 \sum_{j}^{N/\varepsilon} f_j^2 , \tag{82}$$

where ε is a symmetry factor depending on the space group[18] (for the validity of equation (82) see also reference 86). Substituting expression (82) into (81) we obtain an <u>approximate</u> normalized structure factor given by

$$E(hkl) = \frac{F(hkl)}{\varepsilon\left(\sum_{j}^{N/\varepsilon} f_j^2\right)^{1/2}}$$

or

$$|E(hkl)| = \frac{|F(hkl)|}{\varepsilon\left(\sum_{j}^{N/\varepsilon} f_j^2\right)^{1/2}} . \qquad (83)$$

Equation (83) was used to convert the structure amplitudes to normalized values employing an approximate absolute scale ($\sim$1.3) and an average isotropic temperature B factor ($\sim$2.8A^2) determined by Wilson's method[87]. Since the crystal density corresponds to four molecules per unit cell and the space group requires four equivalent positions, two related by centers of symmetry, it was assumed that the germanium atoms were located in general positions. At the time, we were also considering the compound to be $PGe(OH)_2$ and the discrepancy between the calculated and observed density (1.51 $gmcm^{-3}$ and 1.61 $gmcm^{-3}$, respectively, or 60 a.u. per unit cell) was attributed to localized solvent of crystallization.

A sharpened, origin removed, three dimensional Patterson ($|E|^2$-1) map was synthesized. Figure 6 illustrates the Harker section[88] at v=1/2. Harker sections are simply planes of the three dimensional Patterson map, which take advantage of the symmetry of the space group, and many times carry useful information concerning the location of certain atoms (usually heavy) in the unit cell.

The Harker section in this case is obtained from equation (66) by simply substituting v=1/2;

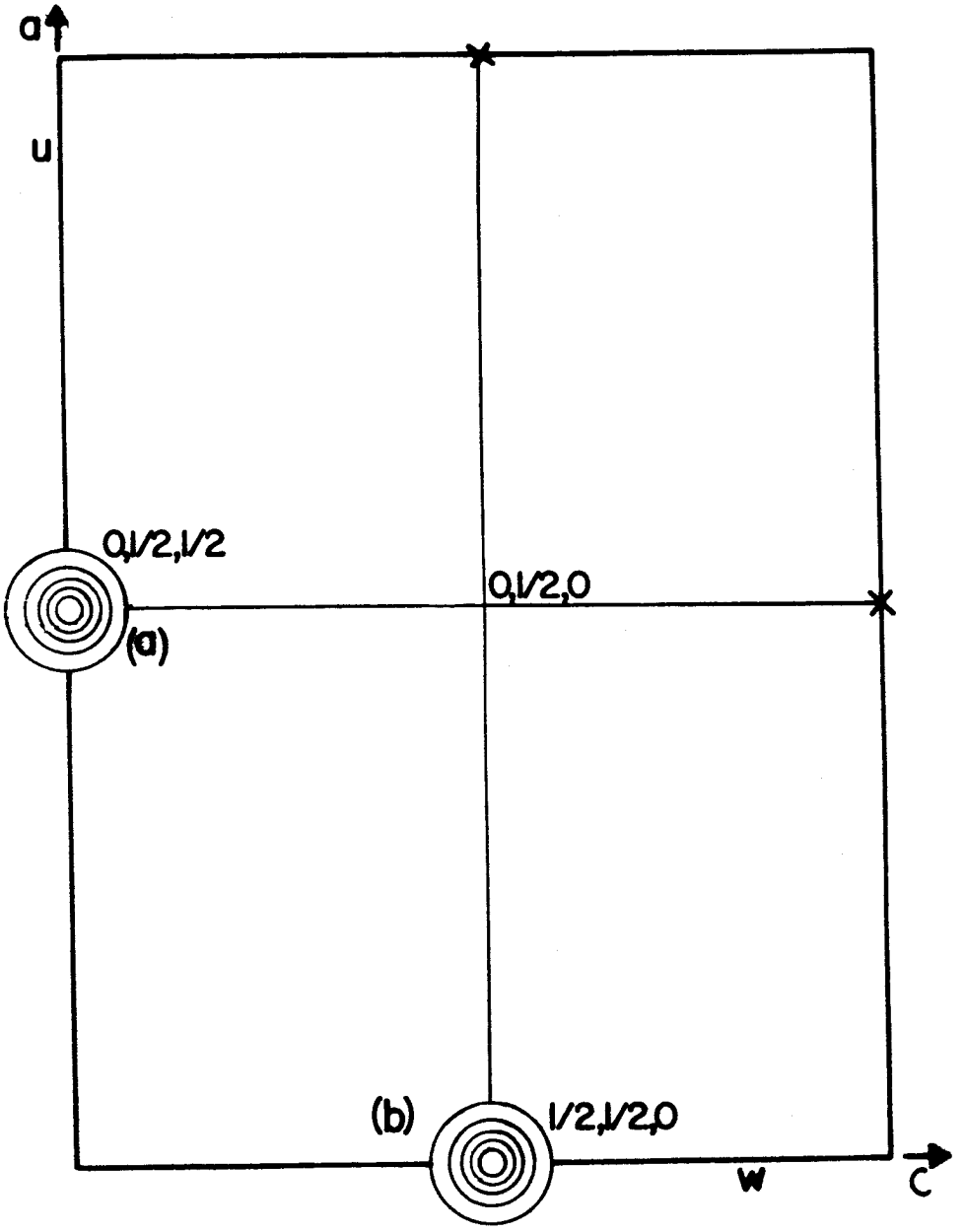


Figure 6. Schematic representation of the Harker section of $PGe(OMe)_2$ at V = 1/2.

$$P(u,1/2,w)=V^{-1}\sum_{h}\sum_{k=-\infty}^{+\infty}\sum_{l}|F(hkl)|^2\cos 2\pi(hu+k/2+lw)=$$

$$=V^{-1}\sum_{h}\sum_{k=-\infty}^{+\infty}\sum_{l}(-1)^k|F(hkl)|^2\cos 2\pi(hu+lw)=$$

$$=V^{-1}\sum_{h}\sum_{l=-\infty}^{+\infty}C_{hl}\cos 2\pi(hu+lw) \qquad (84)$$

where

$$C_{hl}=\sum_{k=-\infty}^{+\infty}(-1)^k|F(hkl)|^2 .$$

The Harker line is based on the same idea, the difference being that higher symmetry is employed for its synthesis. For the group $P2_1/c$ the Harker line is obtained by substituting u=0, w=1/2 in the expression (66)

$$P(0,v,1/2)=\sum_{k=-\infty}^{+\infty}B_k\cos 2\pi ky \quad \text{with} \quad B_k=\sum_{h}\sum_{l=-\infty}^{+\infty}(-1)^l|F(hkl)|^2 . \qquad (85)$$

For space group $P2_1/c$, there are four equivalent positions[82]:

(*i*) x,y,z
(*ii*) $\bar{x},\bar{y},\bar{z}$
(*iii*) $\bar{x}$,1/2+y,1/2-z
(*iv*) x,1/2-y,1/2+z .

Subtracting (*iii*) from (*i*) we obtain

$$\begin{matrix} 2x \\ -1/2 \\ 2z-1/2 \end{matrix} \quad \text{or} \quad \begin{matrix} 2x \\ +1/2 \\ 2z+1/2 \end{matrix} . \qquad (86)$$

The above set of difference coordinates correspond to the Harker section at v=1/2 (see equation (83)). Subtracting (*iv*) from (*i*), the Harker line u=0, w=1/2 coordinates are obtained (see equation (84))

$$\begin{matrix} 0 \\ -1/2+2y \\ -1/2 \end{matrix} \quad \text{or} \quad \begin{matrix} 0 \\ 2y+1/2 \\ 1/2. \end{matrix} \qquad (87)$$

By equating the coordinates of peak (a) in Harker section with the difference coordinates (86), the x,z coordinates of the germanium atom can be deduced:

$$\begin{aligned} & u=2x=0 \longrightarrow x=0 \\ \text{or} \quad & 2x=1 \longrightarrow x=1/2 \\ & w=2z+1/2=1/2 \longrightarrow z=0 \\ \text{or} \quad & 2z-1/2=1/2 \longrightarrow z=1/2 \ . \end{aligned}$$

From the difference coordinates (87) and a peak at v=1/2 in the Harker line the y coordinate of the germanium is deduced:

$$\begin{aligned} & v=-2y+1/2=1/2 \longrightarrow y=0 \\ \text{or} \quad & -2y-1/2=1/2 \longrightarrow y=1/2 \ . \end{aligned}$$

Combining the above results the coordinates of the germanium atom obtained from peak (a) of Harker section and the Harker line are as follows:

x=0	x=0	x=1/2	x=1/2	x=0	x=0	x=1/2	x=1/2
y=0	y=1/2	y=0	y=1/2	y=1/2	y=0	y=1/2	y=0
z=0	z=1/2	z=0	z=1/2	z=0	z=1/2	z=0	z=1/2

(a)

The coordinates (a) correspond to the eight centers of symmetry in space group $P2_1/c$ and are four, symmetry independent pairs. Any pair

can be chosen for the location of two germanium atoms, but then the second pair should be selected in such a way as to give the position of the peak (b) in the Harker section. For instance if the first choice is (0,0,0), then the second should be (1/2,1/2,0).

Following exactly the same procedure, but now using the coordinates of peak (b) in the Harker section, another set of coordinates is deduced for the germanium atom

$$
\begin{array}{rrcl}
 & u=2x=1/2 & \longrightarrow & x=1/4 \\
\text{or} & 2x=-1/2 & \longrightarrow & x=3/4 \\
 & w=2z-1/2=0 & \longrightarrow & z=1/4 \\
\text{or} & 2z+1/2=0 & \longrightarrow & z=3/4 \ ,
\end{array}
$$

and $y=0$ or $y=1/2$ from the Harker line. Thus

x=1/4	x=3/4	x=3/4	x=1/4
y=0	y=0	y=1/2	y=1/2
z=1/2	z=3/4	z=1/4	z=3/4

(b1)

x=3/4	x=1/4	x=1/4	x=3/4
y=0	y=0	y=1/2	y=1/2
z=1/4	z=3/4	z=1/4	z=3/4 .

(b2)

From both sets of coordinates found for the heavy atom, the coordinates of the two Harker peaks can be deduced by taking the appropriate differences. In addition a non-Harker peak of the same size with peaks (a) and (b) located at (1/2,0,1/2) can be explained by using either coordinates (a) or coordinates (b1) or (b2). Therefore, there is ambiguity for the position of the metal in the unit cell.

Structure factor calculations based on positions (a) and (b) for the germanium atom gave exactly the same residual (R) factor ($R=\sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$, $|F_c|$ are the observed and calculated structure amplitudes respectively) of 0.55. This included only 590 reflections since the Ge atom contributes only to those reflections with all indices even or all indices odd. Because the first set of coordinates (a) corresponds to two independent centers of symmetry, at first, it was considered unreasonable and although the second set of coordinates (b) is also special, they do not coincide with a symmetry element as the first. Hence the second set was selected as the positions of the germanium atoms.

A Fourier map based on the phase angles (simply signs in this case because the space group is centrosymmetric) obtained from the set (b1) for the germanium atom revealed some new and additional peaks and although structure factors with R-values as low as 0.33 were obtained with the inclusion of these as atoms of the structure, subsequent Fourier maps based on these coordinates and those of the metal, did not converge to any reasonable structure. The same results were obtained from the set (b2). Since the coordinates (a) were considered unreasonable , a Fourier map based on these coordinates for the metal was never synthesized; instead it was decided to proceed with direct methods of phase determination.

The Patterson synthesis can be characterized as a direct method, because it uses the observed structure amplitudes directly to obtain information about the phases. However, for each crystal structure a new interpretation is required in which the ingenuity and imagination of the crystallographer play a very prominent role. Therefore, the

adjective "direct" has come to be reserved for those methods which attempt to derive the phases of the structure factors directly by mathematical means from X-ray diffraction data[89]. They are usually cast in the form of a mathematical problem which, once formulated, may be solved by a routine sequence of steps in which decisions are of a purely mathematical nature. As early as 1928, Ott[90] derived relations among the structure factors and atomic positions by means of algebraic manipulations on the structure factor equations. His results were mainly of academic interest. In 1948, Harker and Kasper[91], obtained inequality relationships from the application of Schwarz's and Cauchy's inequalities to equations (29) and (37). However, the real breakthrough in direct methods started with the probabilistic approach of Hauptman and Karle, whose work up to 1953 is summarized in a monograph[92], and with Sayre's equation published in 1952[93]. Since then the progress has been enormous and now most structures of moderate complexity, say 40-50 independent non-hydrogen atoms, are solved via direct methods in an almost routine way. The most recent and complete reference to the subject is the book by Hauptman[94].

The basic idea of the direct method is the "structure invariant" and "seminvariant", which are the tools for clarifying the nature of the relationship between the values of the individual phases and the choice of origin. Assume that the position of the nth atom in a unit cell is described by the vector $\vec{r}_n$; if the origin is moved at a different point described by the vector $\vec{r}_o$ with respect to the first origin, the position vector of the nth atom with respect to the new origin will be

$$\vec{r}_n' = \vec{r}_n - \vec{r}_o \ . \tag{88}$$

Then the structure factor F(hkl) will be transformed to

$$F'(hkl)=\sum_{n}^{N} f_n(hkl)\exp(2\pi i\vec{r}_n'\cdot\vec{r}^*) \qquad \text{or}$$

$$F'(hkl)=\sum_{n}^{N} f_n(hkl)\exp\{2\pi i\vec{r}^*\cdot(\vec{r}_n-\vec{r}_o)\}=$$

$$=\exp(-2\pi i\vec{r}^*\cdot\vec{r}_o)\sum_{n}^{N} f_n(hkl)\exp(2\pi i\vec{r}^*\cdot\vec{r}_n)$$

$$F'(hkl)=\exp(-2\pi i\vec{r}^*\cdot\vec{r}_o)F(hkl). \tag{89}$$

Equation (89) shows that the structure amplitude is an origin independent quantity, hence a structure invariant. However, the phase angle $\phi'_{\vec{r}^*}$ with respect to the new origin will be given by

$$\phi'_{\vec{r}^*}=\phi_{\vec{r}^*}-2\pi\vec{r}^*\cdot\vec{r}_o, \tag{90}$$

and it is clearly origin dependent. If now a linear combination of both sides of equation (90) is taken with integer coefficients $\mathcal{D}_{\vec{r}^*}$ which depend upon $\vec{r}^*$[94], we obtain

$$\sum_{\vec{r}^*}\mathcal{D}_{\vec{r}^*}\phi'_{\vec{r}^*}=\sum_{\vec{r}^*}\mathcal{D}_{\vec{r}^*}\phi_{\vec{r}^*}-2\pi\vec{r}_o\cdot\left(\sum_{\vec{r}^*}\mathcal{D}_{\vec{r}^*}\vec{r}^*\right) . \tag{91}$$

Clearly if $\sum_{\vec{r}^*}\mathcal{D}_{\vec{r}^*}\,\vec{r}^*=o$, then

$$\sum_{\vec{r}^*}\mathcal{D}_{\vec{r}^*}\phi'_{\vec{r}^*}=\sum_{\vec{r}^*}\mathcal{D}_{\vec{r}^*}\phi_{\vec{r}^*} , \tag{92}$$

no matter what the vector $\vec{r}_o$ may be, and the linear combination of the phases is a structure invariant (sometimes called "universal" structure

invariant), since it is entirely origin independent (for the definition of seminvariants see reference 94, p. 29)

Depending on the space group, an origin can be defined by specifying the signs (for centrosymmetrical cases) to certain reflections in such a way as to avoid a conflict with the "boundary conditions" of the invariance property of certain linear combinations. The origin fixes the geometrical part of structure factor and then provided that the normalized structure amplitudes are known, an appropriate technique is employed for the generation of phases.

In the present case, Sayre's equation[93] was used, which also results from the $\sum_2$ equation of Hauptman and Karle[92]; this equation states that

$$\mathrm{Sign}(E_{\vec{A}})=\mathrm{Sign}\Big(\sum_{\vec{A}=\vec{B}+\vec{C}} E_{\vec{B}}E_{\vec{C}}\Big) \tag{93}$$

where $\vec{A}$, $\vec{B}$, $\vec{C}$ are the vectors (hkl) for the reflections A, B and C and $E_{\vec{A}}$, $E_{\vec{B}}$, and $E_{\vec{C}}$ are the normalized structure factors for these reflections. The probability that $E_{\vec{A}}$ is positive, $P_+(E_{\vec{A}})$, is given by[95]

$$P_+(E_{\vec{A}})\simeq 1/2+(1/2)\tanh(\sigma_3/\sigma_2^{3/2})\,|E_{\vec{A}}|\sum_{\vec{A}=\vec{B}+\vec{C}} E_{\vec{B}}E_{\vec{C}}\ , \tag{94}$$

where $\sigma_m=\sum_j z_j^m$, with z_j the atomic number of atom j and N the total number of atoms in the unit cell. In order to define the origin in the space group $P2_1/c$, three signs may be arbitrarily assigned. If the values of (hkl) for these three reflections are called $(hkl)_1$, $(hkl)_2$, and $(hkl)_3$, then these three origin-determining reflections must be chosen such that the following six vectors are not (even, even, even): $(hkl)_1$, $(hkl)_2$, $(hkl)_3$, $\{(hkl)_1+(hkl)_2\}$, $\{(hkl)_1+(hkl)_3\}$,

$\{(hkl)_2+(hkl)_3\}$. This may be accomplished by choosing the three reflections as follows: none may have the parity (even, even, even), the parity of the second reflection chosen must be different from the parity of the first reflection, the parity of the third must be different from the parities of the first, the second, and the sum of the first two. Then the general procedure for sign-determination can be outlined as follows. A starting set of reflections is selected. This set includes the origin-determining reflections and n other reflections, usually four or five, which <u>are not</u> structure invariants. The origin-determining signs are arbitrarily assigned, and the other n signs may be + or -; therefore 2^n possible starting sign sets are considered. Sayre's equation is reiteratively applied to each of these starting sign sets, yielding 2^n solutions. When Sayre's equation is applied to a starting sign set, additional signs are determined, and these are used to determine more signs, and to redetermine those already predicted. This process is reiterated until there are no new additions or no changes in the list. The signs in the starting set are predicted (during each reiteration) from the other signs which have been predicted, but the signs in the staring set are not allowed to change.

Once all the possible sets of phases are calculated only one or the true solution must be determined. A method used is to calculate all 2^n electron density maps, and to select the one which is chemically most reasonable. A more efficient method is to use a "consistency" index as a test for the correct structure. The consistency index C, is defined as

$$C=\frac{\left\langle\left|E_{\vec{A}}\sum_{\vec{A}=\vec{B}+\vec{C}}E_{\vec{B}}E_{\vec{C}}\right|\right\rangle}{\left\langle\left|E_{\vec{A}}\right|\sum_{\vec{A}=\vec{B}+\vec{C}}\left|E_{\vec{B}}\right|\left|E_{\vec{C}}\right|\right\rangle}, \qquad (95)$$

where the sums are overall pairs of $\vec{B}$ and $\vec{C}$ for which $\vec{B}+\vec{C}=\vec{A}$. If C is equal to one, the particular solution is said to be completely consistent. The true solution will usually be the most consistent one, i.e., it will have the highest consistency index.

A basic difficulty which we encountered with the $PGe(OMe)_2$ structure was that almost all of the largest normalized structure amplitudes were of the type (hkl) = (even, even, even) or (odd, odd, odd), and therefore it was not possible to find three appropriate origin determining reflections with which to initiate the direct method of sign determination. In an effort to circumvent the problem, the E-distribution was changed to correspond approximately to that of the free base macrocycle. This was approximated by subtracting the amplitude of the germanium atom contribution from the 590 germanium atom-affected observed structure amplitudes $|F_o(hkl)|$ (the amplitude of the contribution is the same for both sets of coordinates (a) or (b)). These new structure amplitudes, $|F_p(hkl)|$, where

$$|F_p(hkl)|=||F_o(hkl)|-|F_{Ge}(hkl)_{cal}||, \qquad (96)$$

and $|F_{Ge}(hkl)_{cal}|$ is the calculated structure amplitude of the germanium atom, were then converted to normalized values ($|E_p|$).

If we call F_{P+Ge} the structure factor of the $PGe(OMe)_2$, then the following four combinations are possible for F_p,the structure factor of the free macrocycle:

$$(\alpha) \quad F_{P+Ge}=(+)|F_p|+(+)|F_{Ge}|$$

$$(\beta) \quad F_{P+Ge}=(-)|F_p|+(-)|F_{Ge}|$$

$$(\gamma) \quad F_{P+Ge}=(+)|F_p|+(-)|F_{Ge}|$$

$$(\delta) \quad F_{P+Ge}=(-)|F_p|+(+)|F_{Ge}| \ .$$

For the cases (α), (β), the structure amplitude of the free macrocycle will be given by

$$|F_P| = ||F_{P+Ge}| - |F_{Ge}||,$$

which is equation (96). For the cases (γ), (δ) the structure amplitude $|F_{Ge}|$ will be given by

$$|F_{P+Ge}| = ||F_P| - |F_{Ge}|| \ . \qquad (97)$$

Now if $\quad |F_P| < |F_{Ge}| \longrightarrow |F_{Ge}| - |F_P| > 0, \quad$ hence

$$|F_{P+Ge}| = |F_{Ge}| - |F_P| \ , \quad \text{or}$$

$$-|F_P| = |F_{P+Ge}| - |F_{Ge}|, \text{ or}$$

$$|F_P| = ||F_{P+Ge}| - |F_{Ge}||$$

which again is equation (96). However, if $|F_P| > |F_{Ge}|$ (which is not very probable for most of the reflections), $|F_P| - |F_{Ge}| > 0$ and therefore equation (97) becomes

$$|F_{P+Ge}| = |F_P| - |F_{Ge}| \qquad \text{or}$$

$$|F_P| = ||F_{P+Ge}| + |F_{Ge}|| \ ,$$

and in this last case the approximation (96) fails.

The statistical distributions of the normalized structure amplitudes, both with and without the metal, along with the theoretical distributions for centric and acentric cases are given in Table II, where the normalized structure factors were scaled such that $\langle E^2 \rangle = 1.0$. A set of 200 $|E_P|$'s with values greater than 1.5 were used in a sign determination program written by Long[96]. A starting set of seven signs was used; these corresponded to the three origin-defining reflections, ($\bar{5}24$), ($\bar{2}12$), (423) and the four general reflections, ($\bar{4}11$), (323), ($\bar{3}12$) and ($\bar{7}35$).

TABLE II. Statistical Distribution of $|E|$

	$PGe(OMe)_2$	$PGe(OMe)_2$-Ge	Centric	Acentric
$\langle\|E\|\rangle$	0.80	0.85	0.80	0.89
$\langle E^2\rangle$	1.00	1.00	1.00	1.00
$\langle E^2-1\rangle$	1.04	0.85	0.97	0.74
%>1.0	30.65	32.95	32.00	37.00
%>2.0	4.71	3.49	5.00	1.80
%>3.0	0.11	0.22	0.30	0.01

In each of the sixteen possible solutions (2^4) all 200 signs were determined. The solution with the highest consistency index was selected (C=0.72), and an E_p Fourier map revealed the positions of all the non-hydrogen atoms of the porphyrin <u>except</u> that of the metal, which was conspicuously absent. The solution proved to be that of the <u>two independent molecules centered around the independent centers of symmetry</u> and subsequent refinement verified the correctness of the structure. Such unusual behavior has been observed in other structures as well[97-99].

The coordinates of 26 non-hydrogen atoms corresponding to $PGe(OH)_2$ were determined from the E_p map. A structure factor calculation based on these with an average isotropic thermal parameter for all atoms gave an R-value of 0.21. At this stage, full-matrix, <u>unit</u> weight least squares refinement was initiated. The function most commonly minimized is

$$R=\sum W(hkl)\{|F_o(hkl)-|F_c(hkl)|\}^2,$$

where the sum is over the set of crystallographically independent planes and the W(hkl) are weights.

Three cycles of refinement, one varying the coordinates and the scale factor, another varying all isotropic thermal parameters and the last varying the coordinates again along with the scale factor, reduced R to 0.15. An observed electron density map was synthesized employing all but about 100 coefficients which did not satisfy a rejection test, ($|F_o|$xr.r.-$|F_c|$), where r.r.=0.5; whenever the foregoing expression was negative the corresponding reflection was not used in the Fourier synthesis. This map revealed additional electron density near that of

each of the independent oxygen atoms. From the peak height and the distance to the oxygen atom (∿1.4A), it became clear that the hydroxyl groups had been replaced by methoxy groups of methanol during crystallization, and that the additional density was due to a carbon atom corresponding to a methyl group. An infrared spectrum of the compound did not show an absorption in the -OH stretching region but it did show a very strong absorption band at 1200 cm^{-1}, which is close to the region of the $O-CH_3$ stretch.

A structure factor calculation including the coordinates of the assumed methyl carbon atoms reduced R significantly to 0.108 and a difference electron density map was synthesized. All the expected hydrogen atoms appeared including those of the methyl group thus confirming the correctness of the original assumption. The hydrogen atoms were assigned isotropic temperature factors which were 1.25 greater than the isotropic temperature factors of the carbon atoms to which they were bonded and the resulting structure factor calculation had an R of 0.094.

A correction for anomalous dispersion was introduced at this point for the germanium atom in the zero ionization state ($f'=-1.3$ for CuK_α). Anisotropic thermal parameters were introduced for all the non-hydrogen atoms and the refinement was continued. The thermal parameters were varied separately for the inside atoms of molecules 1 and 2 followed by the outside atoms. According to the notation of Figure 7, the "inside" atoms were taken to be:

1Ge, 1NA, 1A1, 1A4, 1AB, 1NB, 1B1, 1B4, 1BC,

2Ge, 2NA, 2A1, 2A4, 2AB, 2NB, 2B1, 2B4, 2BC,

with the remainder of the atoms being taken as the outside atoms.

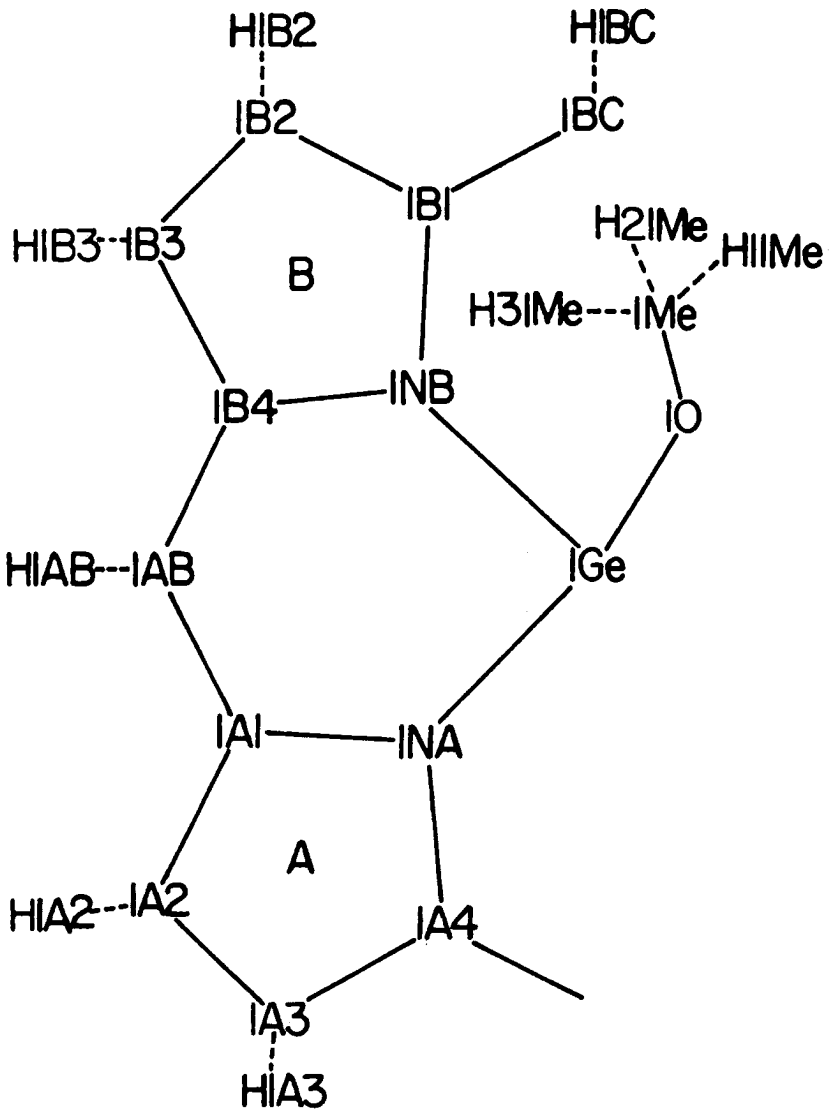


Figure 7. Labelling of $PGe(OMe)_2$.

Three cycles of refinement, one varying the coordinates of the non-hydrogen atoms and the scale factor, another varying the anisotropic thermal parameters, first for the eighteen inside atoms and then for the twelve outside atoms, and the last varying the coordinates of all the non-hydrogen atoms and the scale factor, reduced R to 0.059. At this stage, it was clear that 6 of the largest low order reflections were affected by extinction: (111), ($\bar{1}$11), (020), (200), ($\bar{2}$02) and (040); the average percent discrepancy, 100x $<||F_o|-|F_c||>/|F_o|$, for these reflections was 12.5. Consequently, they were removed from the structure factor calculation and the R-value decreased to 0.056. Two more cycles of refinement without these reflections, first on anisotropic thermal parameters of the eighteeen and then on the twelve non-hydrogen atom sets, and then on coordinates of all non-hydrogen atoms, reduced the R-value to 0.052. Relocation of the hydrogen atoms form a new difference electron density map improved the R value to 0.048. Three more cycles of refinement , first on anisotropic thermal parameters in the way described above, then on coordinates of all non-hydrogen atoms and finally on thermal parameters again, gave an R-value of 0.043. A cycle of refinement in the coordinates of hydrogen atoms, followed by a cycle of refinement on the coordinates of non-hydrogen atoms did not change the R-factor; furthermore, the parameter shifts were insignificant compared to estimated standard deviations. Therefore, the refinement of the structure was terminated.

V. RESULTS

Tables III and IV list the final atomic coordinates, anisotropic temperature factors, and peak heights of all the non-hydrogen atoms of the two independent centrosymmetrical molecules of the asymmetric unit of the unit cell. Table V lists the final coordinates, isotropic thermal parameters and peak heights of hydrogen atoms. The atom notation is according to Figure 7. The standard deviations of the atomic coordinates are in parentheses and are those of the final cycle of the least squares refinement; the errors in the hydrogen atom coordinates are about ten times greater than those of the atom to which they are bonded. A least squares plane was calculated for the atoms of the porphine of each molecule. The deviations of the atoms from these least squares planes are listed in Table VI. Best least squares planes were also computed for the atoms of each pyrrole ring separately and the deviations from these planes are listed in Table VII. From Table VII, it can be seen that the individual pyrrole rings are planar within the error of their determination (± 0.01A). Table VIII lists the equivalents of the principal mean-square displacements from equilibrium positions of the non-hydrogen atoms. In addition, an average isotropic temperature factor is listed for each atom based on the principal mean-square displacements. From Table VIII it can be seen that molecule 1 is considerably more disordered than molecule 2 of the asymmetric unit. This can also be seen from the peak heights listed in

Tables III and IV and it is also the reason for the generally larger standard deviations of the atomic coordinates of molecule 1. The bond distances and angles of the two independent molecules are presented in Figures 8 and 9, and Figures 10 and 11 show distances between non-bonded atoms. Figure 12 shows the $PGe(OMe)_2$ molecule drawn in perspective in terms of its vibration ellipsoids (ORTEP, reference 100). Finally, Figure 13 illustrates in perspective the manner in which the four molecules are packed in the unit cell (ORTEP, reference 100). The molecules (0,0,0) and (0,1/2,1/2), as well as (1/2,1/2,0) and (1/2,0,1/2) are related with the symmetry operations of the space group $P2_1/c$ while the molecules (0,0,0) and (1/2,1/2,0) or (0,1/2,1/2) and (1/2,0,1/2) are symmetry independent. The dihedral angle between the two symmetry dependent molecules (0,0,0), (0,1/2,1/2) is 33.34°, while the dihedral angle between the other two symmetry related molecules (1/2,1/2,0), (1/2,0,1/2) is 25.55°. The angle between the independent molecules is 75.31° and the angle between molecules (1/2,1/2,0), (0,1/2,1/2) or (0,0,0), (1/2,0,1/2) is 67.64°.

TABLE III. Final Atomic Coordinates, Thermal Parameters and Peak Heights. Molecule 1.

Atom	Fractional Coordinates			Anisotropic Thermal Parameters($\times 10^4$)						Peak Height(eA^{-3})
	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}	
1Ge	.5	.5	0	35	48	92	-4	-2	-2	49.0
1NA	.4436(3)	.6268(3)	-.0122(5)	50	49	131	0	-16	-6	6.7
1A1	.4743(4)	.7043(4)	-.0850(6)	63	57	136	-7	-32	-3	5.1
1A2	.4190(4)	.7812(4)	-.0517(6)	77	55	179	9	-52	-22	4.9
1A3	.3527(4)	.7496(4)	.0382(7)	62	80	205	16	-41	-36	4.7
1A4	.3673(3)	.6539(4)	.0685(6)	51	59	152	9	-27	-23	4.9
1AB	.3177(3)	.5969(4)	.1593(6)	46	77	157	0	0	-30	5.2
1NB	.4042(2)	.4552(3)	.1427(4)	42	66	111	-11	4	-5	6.5
1B1	.3350(3)	.5068(4)	.1972(6)	43	92	110	-14	-5	24	5.2
1B2	.2841(3)	.4509(5)	.3033(6)	49	105	119	-24	16	-29	5.0
1B3	.3226(4)	.3675(4)	.3106(6)	68	86	140	-33	0	-14	4.7
1B4	.3972(3)	.3684(4)	.2107(6)	48	73	115	-21	-2	-2	5.0
1BC	.4524(4)	.2948(4)	.1796(6)	63	67	118	-14	-14	6	5.1
1O	.5698(2)	.5430(2)	.1650(3)	49	57	87	-14	-18	2	7.9
1Me	.5754(4)	.5020(4)	.3129(6)	51	80	114	-14	-12	-1	5.2

TABLE IV. Final Atomic Coordinates, Thermal Parameters and Peak Heights. Molecule 2.

Atom	Fractional Coordinates			Anisotropic Thermal Parameters($\times 10^4$)						Peak Height(eA^{-3})
	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}	
2Ge	0	.5	.5	34	34	88	1	5	0	58.0
2NA	-.0948(2)	.5888(2)	.4181(4)	38	36	111	3	4	0	7.4
2A1	-.1022(3)	.6807(3)	.4589(5)	40	44	125	6	13	14	5.9
2A2	-.1809(3)	.7185(3)	.3796(6)	47	48	171	11	17	16	5.9
2A3	-.2109(3)	.6512(4)	.2956(6)	44	57	159	10	7	20	5.7
2A4	-.1664(3)	.5685(3)	.3171(6)	40	52	113	3	8	11	5.7
2AB	-.1847(3)	.4832(4)	.2516(6)	42	61	134	-2	-4	1	6.1
2NB	-.0592(2)	.3976(2)	.3738(4)	40	34	102	0	1	-4	7.2
2B1	-.1350(3)	.4046(3)	.2775(6)	45	47	121	-2	3	6	6.0
2B2	-.1537(3)	.3161(4)	.2056(6)	54	56	146	-9	-4	-15	5.6
2B3	-.0891(3)	.2577(4)	.2563(6)	56	52	134	-15	20	-13	5.8
2B4	-.0306(3)	.3081(3)	.3631(5)	48	43	109	-5	15	-6	5.6
2BC	.0450(3)	.2720(3)	.4398(6)	55	40	131	3	22	-3	6.3
2O	.0732(2)	.5330(2)	.3410(3)	38	45	83	-3	16	4	9.0
2Me	.0594(4)	.5067(4)	.1814(6)	57	78	99	-11	18	-1	4.7

TABLE V. Final Atomic Coordinates, Isotropic Temperature Factors, and Peak Heights of the Hydrogen Atoms.

Atom	x	y	z	$B(A^2)$	Peak Height (eA^{-3})
Molecule 1					
H1A2	.4312	.8502	.0911	5.2	.50
H1A3	.3143	.7824	.0874	5.7	.41
H1AB	.2670	.6249	.2135	5.3	.38
H1B2	.2311	.4838	.3599	5.3	.45
H1B3	.3234	.3076	.3694	5.7	.36
H1BC	.4513	.2327	.2210	5.0	.39
H11Me	.6064	.5304	.3744	4.4	.40
H21Me	.5203	.4940	.3536	4.4	.51
H31Me	.5974	.4427	.3002	4.4	.48
Molecule 2					
H2A2	-.1947	.7766	.3966	4.6	.54
H2A3	-.2727	.6488	.2292	4.6	.44
H2AB	-.2407	.4754	.1930	4.4	.41
H2B2	-.1987	.2978	.1359	4.8	.43
H2B3	-.1024	.2072	.2435	4.7	.44
H2BC	.0594	.2093	.4162	4.5	.49
H12Me	.1010	.5368	.1131	4.4	.51
H22Me	.0009	.5155	.1515	4.4	.44
H32Me	.0734	.4384	.1690	4.4	.51

TABLE VI. Deviations of Porphine Skeleton From Least Squares Plane

Molecule 1		Molecule 2	
Atom	d(A)	Atom	d(A)
1Ge	0	2Ge	0
1NA	-.05	2NA	-.03
1A1	.03	2A1	-.01
1A2	.09	2A2	.01
1A3	.01	2A3	.01
1A4	-.06	2A4	-.01
1AB	-.07	2AB	.00
1NB	-.02	2NB	.02
1B1	-.03	2B1	.01
1B2	.03	2B2	.00
1B3	.07	2B3	-.02
1B4	.03	2B4	.01
1BC	-.01	2BC	-.01
1O	1.81	2O	1.81
1Me	2.67	2Me	2.68
$\sigma = \pm .05A$		$\sigma = \pm .02A$	

TABLE VII. Deviations of Pyrroles From Least Squares Planes

Molecule 1

Pyrrole A

Atom	d(A)
1NA	.003
1A1	-.008
1A2	.009
1A3	.008
1A4	.003

$\sigma = \pm.007A$

Pyrrole B

Atom	d(A)
1NB	.004
1B1	-.003
1B2	.000
1B3	.002
1B4	.004

$\sigma = \pm.003A$

Molecule 2

Pyrrole A

Atom	d(A)
2NA	-.002
2A1	.002
2A2	-.002
2A3	.001
2A4	.000

$\sigma = \pm.002A$

Pyrrole B

Atom	d(A)
2NB	-.001
2B1	-.003
2B2	.007
2B3	-.008
2B4	.006

$\sigma = \pm.006A$

TABLE VIII. Principal Mean-Square Displacements (A^2) in 'Isotropic B' Notation

Molecule 1					Molecule 2				
Atom	$8\pi^2\bar{u}_1^2$	$8\pi^2\bar{u}_2^2$	$8\pi^2\bar{u}_3^2$	<B>	Atom	$8\pi^2\bar{u}_1^2$	$8\pi^2\bar{u}_2^2$	$8\pi^2\bar{u}_3^2$	<B>
1Ge	2.53	3.08	4.13	3.24	2Ge	2.43	2.77	3.20	2.80
1NA	3.04	4.07	4.93	4.01	2NA	2.93	3.13	3.56	3.20
1A1	2.73	4.79	6.89	4.80	2A1	2.89	3.15	4.76	3.60
1A2	2.97	4.34	9.44	5.58	2A2	3.15	3.92	6.00	4.36
1A3	3.53	4.98	9.75	6.09	2A3	3.30	4.02	5.88	4.40
1A4	2.90	4.10	6.89	4.63	2A4	2.94	3.48	4.70	3.71
1AB	3.64	4.20	7.17	5.00	2AB	3.48	4.10	5.11	4.23
1NB	3.13	3.37	5.86	4.12	2NB	2.66	3.05	3.64	3.12
1B1	2.50	3.91	8.23	4.88	2B1	3.26	3.90	4.20	3.79
1B2	2.97	3.54	9.97	5.49	2B2	3.28	4.83	5.55	4.55
1B3	3.27	4.44	9.55	5.75	2B3	3.28	3.40	6.37	4.35
1B4	2.99	3.47	7.21	4.56	2B4	2.80	3.34	4.72	3.62
1BC	3.07	4.53	7.07	4.89	2BC	2.96	3.49	5.47	3.97
1O	2.00	3.72	5.99	3.90	2O	1.90	3.65	3.89	3.15
1Me	2.85	4.46	7.20	4.84	2Me	2.52	4.89	7.01	4.81

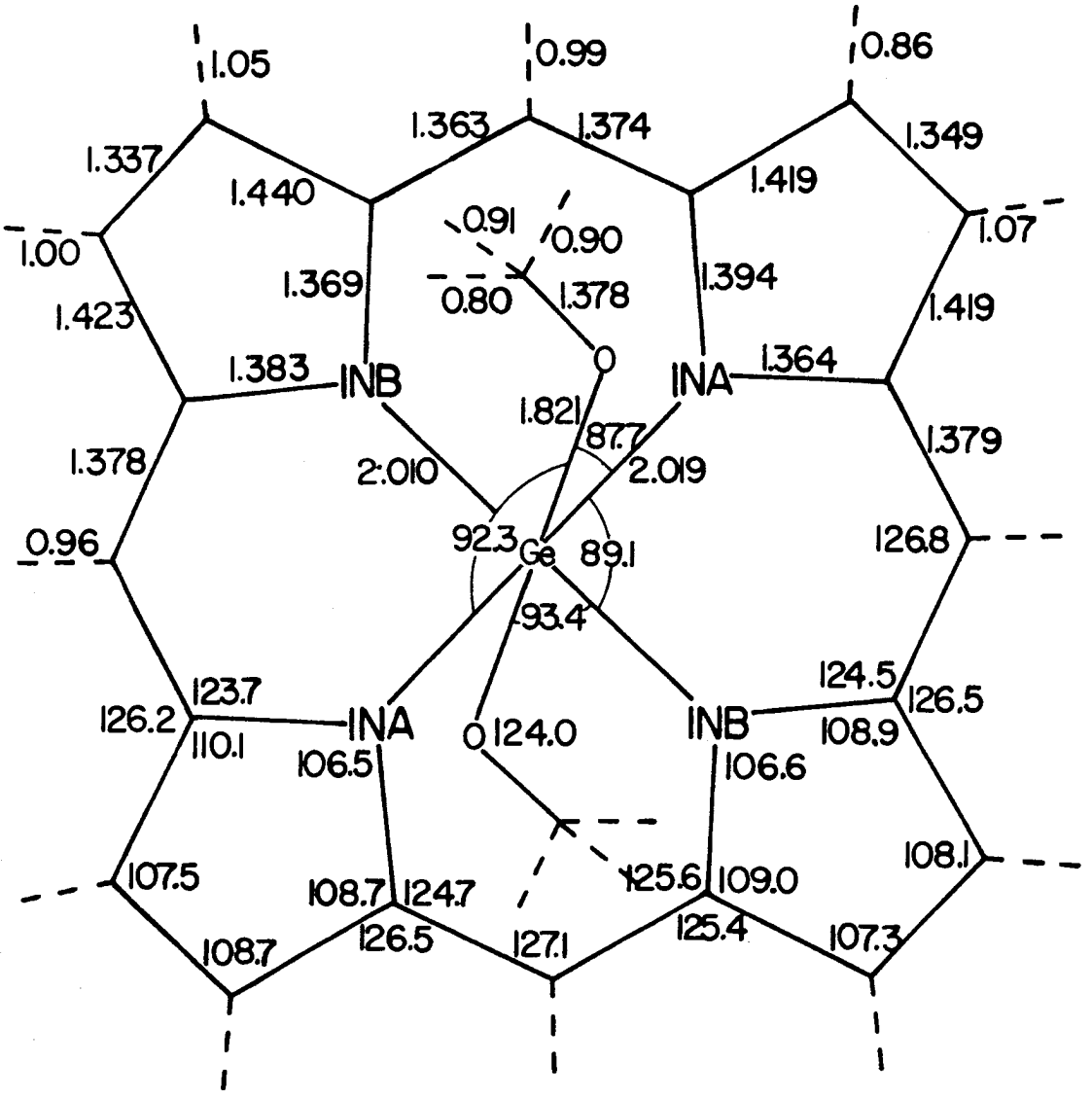


Figure 8. Bond distances (in A) and angles (in degrees); broken lines indicate C-H distances; molecule 1.

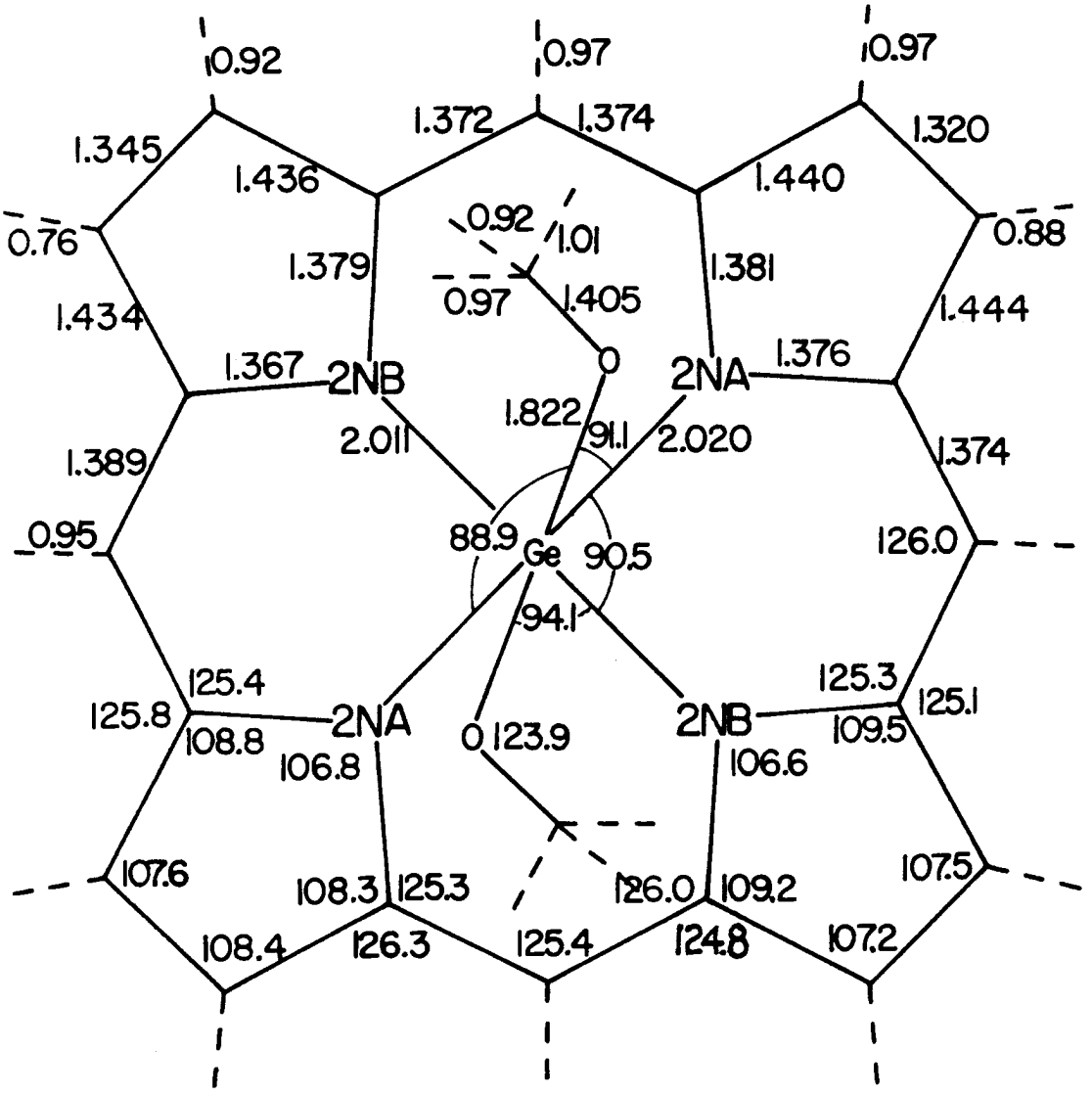


Figure 9. Bond distances (in A) and angles (in degrees); broken lines indicate C-H distances; molecule 2.

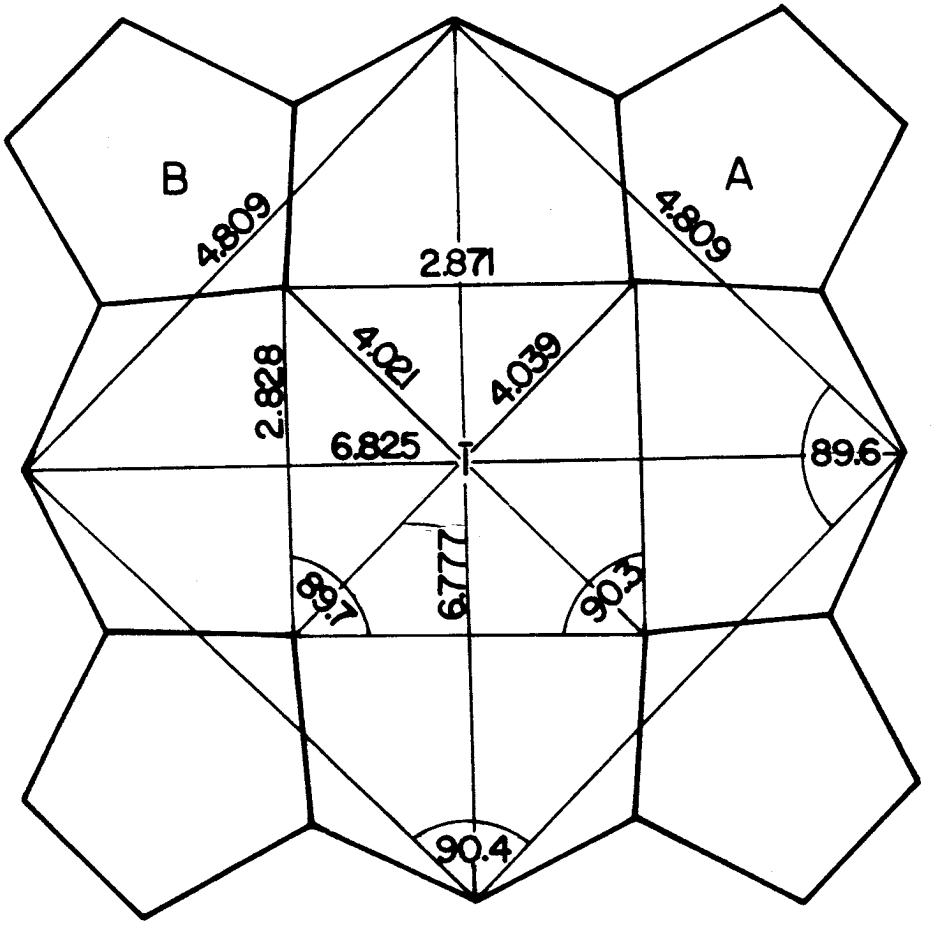


Figure 10. Intramolecular distances and angles of the central core region; molecule 1.

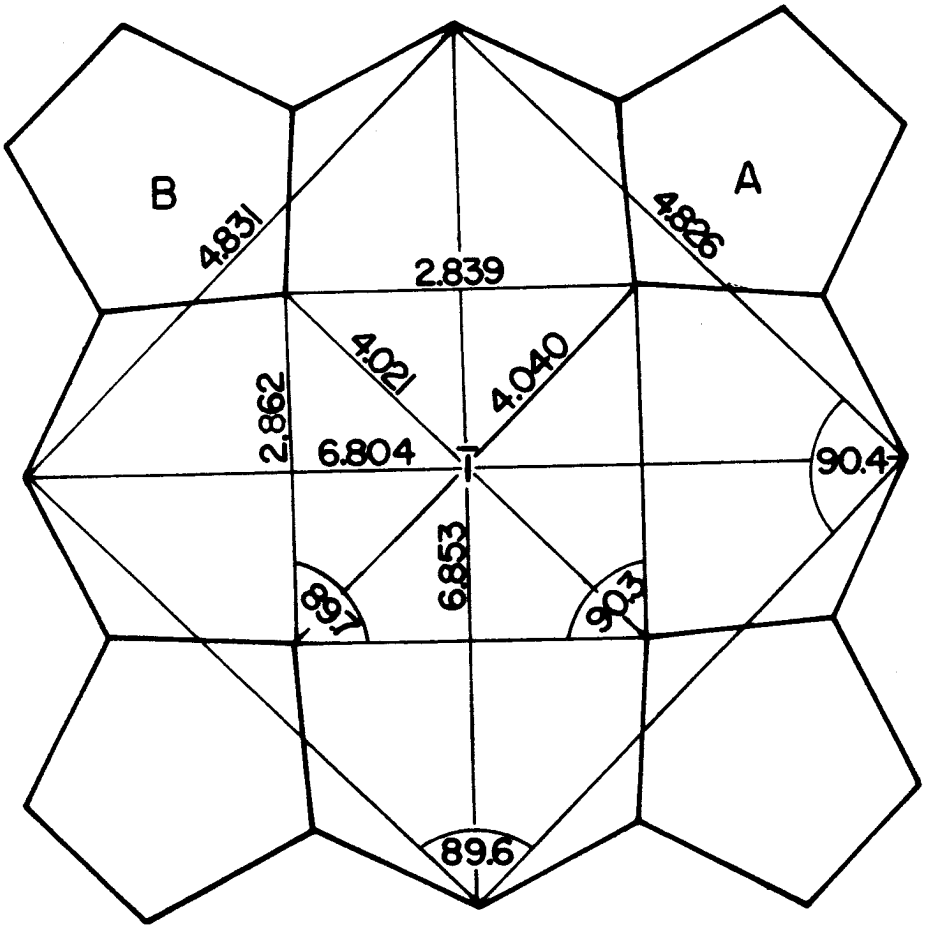


Figure 11. Intramolecular distances and angles of the central core region; molecule 2.

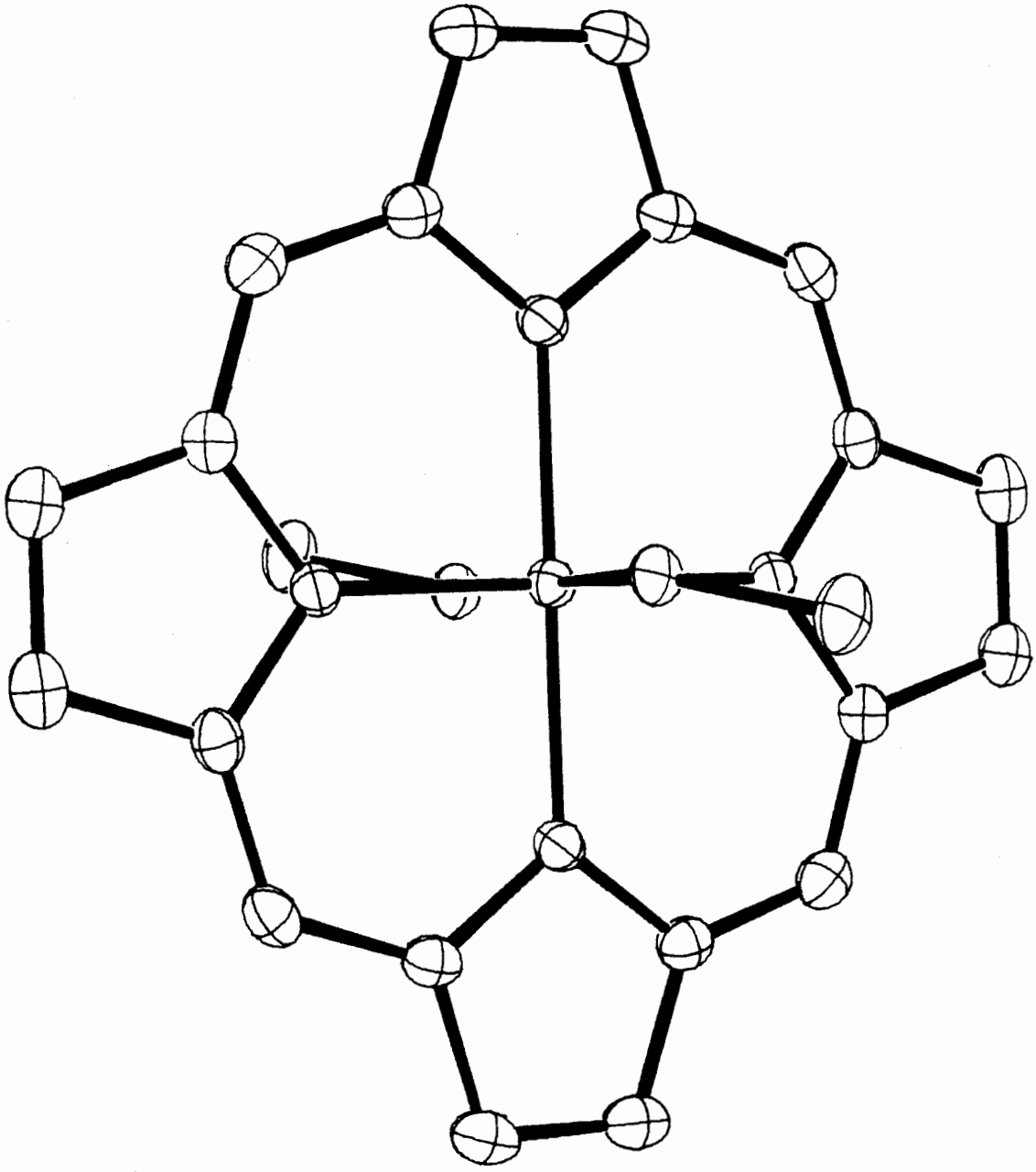


Figure 12. Computer plot of the $PGe(OMe)_2$ molecule (ORTEP, reference 100).

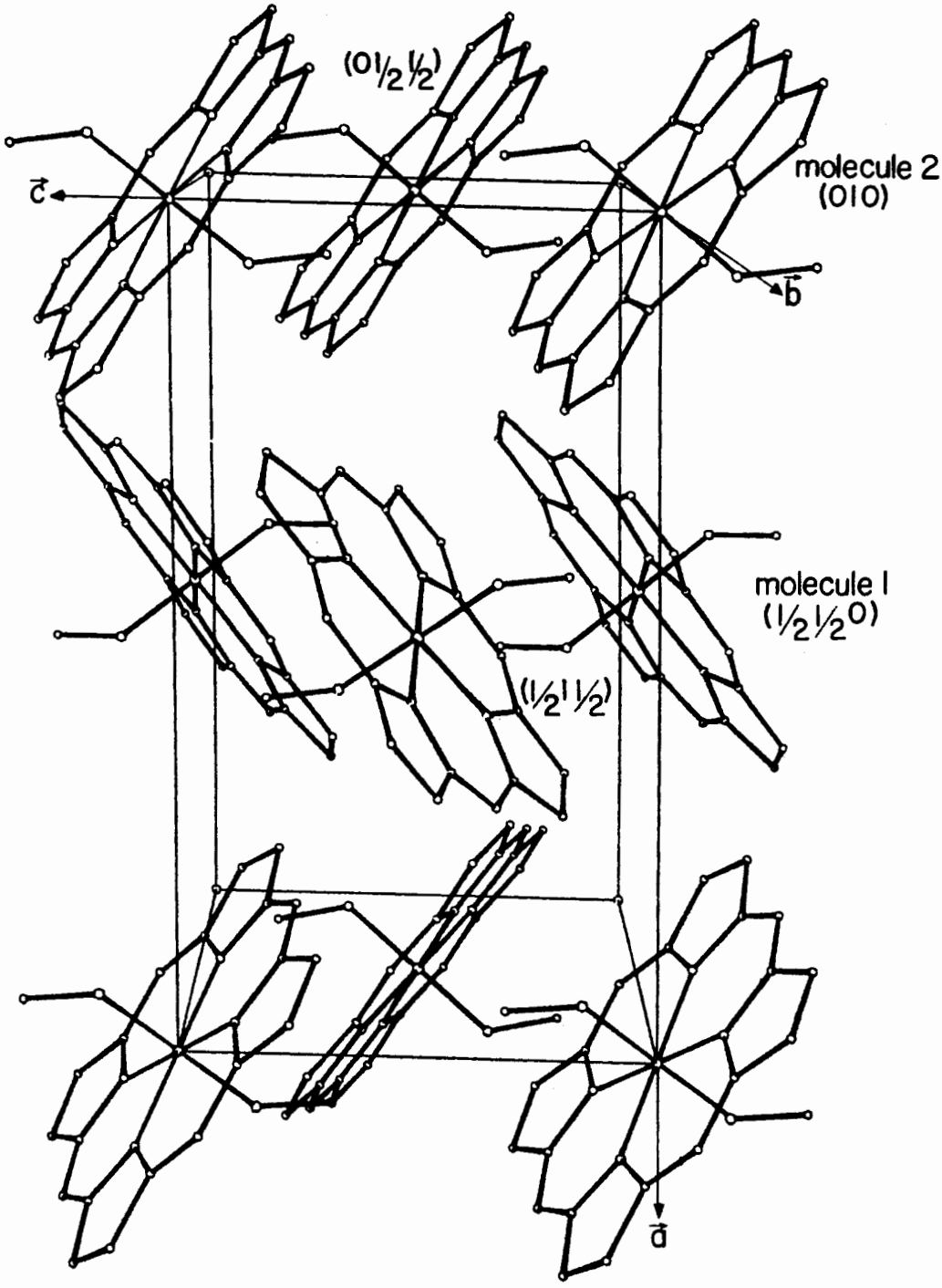


Figure 13. Molecular packing of $PGe(OMe)_2$ (ORTEP, reference 100).

VI. DISCUSSION

It is unusual and also interesting that there are two molecules of $PGe(OMe)_2$ located on independent centers of symmetry in the unit cell. Although the reason for such a phenomenon remains uncertain, one thing that is clear is the independence of the two molecules, in this case, can lead to more efficient molecular packing. Four molecules in general positions in space group $P2_1/c$ are related to each other by the symmetry elements of the space group. However, in the present case, the orientation of the one pair of molecules is not dependent on the orientation of the other pair and the only restrictions between the two pairs are "chemical and/or physical" in nature. Thus, the plane of molecule 2 is tilted about 75° with respect to the plane of molecule 1 (Figure 13). In view of the large metal-normalized density of $PGe(OMe)_2$ (see below), this angle is probably close to an optimum with respect to molecular packing.

The packing of $PGe(OMe)_2$ is exceptionally efficient compared to other porphyrin systems. The density of $PGe(OMe)_2$ is the highest among all porphines and porphyrins for which a structure has been determined or unit cell dimensions have been measured. Due to the rather special geometry of porphines and porphyrins (planar with a central hole), macroscopic mass density is not the best quantity to compare for relative effectiveness of packing. A better quantity is $\{F(000)-M\}/V$, where F(000) is the number of electrons per unit cell, M is the contribution to F(000) from a metal which might be located in the central hole,

and V is the volume of the unit cell. The values of this "normalized" electron density term for about 35 porphines and porphyrins are given in Table IX. From Table IX, it can be seen that these fall conspicuously into two groups (a) one with an average normalized density of (0.382± 0.006)eA^{-3} and (b) the other with an average normalized density of (0.402±0.003)eA^{-3}. The corresponding value for $PGe(OMe)_2$ is 0.425 eA^{-3}, which differs significantly from the averages of the respective groups by many standard deviations. The compact packing of $PGe(OMe)_2$ is achieved with no unusually close intermolecular contacts and is probably related to the location of the molecules on independent centers of symmetry in the crystal and the concomitant freedom derived thereof. The closest van der Waal's approach is about 3.49A (1Me-2Me), so that the efficiency of the molecular packing is probably accomplished in a cooperative way.

The recently determined crystal structure of dipyridinateocta-ethylporphyrinRu(II)($Ru(Py)_2OEP$)[99], which also has two independent molecules per unit cell located on independent centers of symmetry (space group $P2_1/c$), does not show a high normalized density (0.380eA^{-3}). The very high normalized density observed in $PGe(OMe)_2$ might be due to the more appropriate geometry of the ligands around the metal, combined with the independence of the molecules in the unit cell.

The $PGe(OMe)_2$ metalloporphine can be considered as a centrosymmetric distorted octahedral complex with the oxygen atoms located at apical positions. As was expected from the metal radius and its coordination[120], the germanium atom is located on the plane of the macrocycle (Table VI). The macrocycles of both 1 and 2 are essentially planar, although molecule 2 shows a higher degree of planarity than molecule 1; in molecule 2 the

TABLE IX(a). Comparison of Intrinsic Packing-Density of Porphyrins

Compound	Space Group	$\rho(eA^{-3})$*
TPP[38]	$P\bar{1}$	0.401
Ag(II)TPP[101]	$P\bar{1}$	0.399
Cd(IV)TPP[101]	P1 (or $P\bar{1}$)	0.399
α-Chlorohemin[102]	$P\bar{1}$	0.409
Porphine[39]	$P2_1/c$	0.403
TPrP[40]	$P2_1/c$	0.400
Cu(II)TPrP[103]	$P2_1/c$	0.398
Mn(III)ClTPP[101]	$P2_1/c$	0.398
Bis-imidazole Fe(III)TPPCl[104]	$P2_1/c$	0.401
VO-DPEEP[105]	$P2_1/c$	0.397
PyZn(II)TPyP[106]	C2/c	0.404
TPP[107]	$I\bar{4}2d$	0.404
Cu(II)TPP[108]	$I\bar{4}2d$	0.403
Pd(II)TPP[108]	$I\bar{4}2d$	0.405
Ni(II)TPP[108]	$I\bar{4}2d$	0.405

$$\langle (F(000)-M)/V \rangle = (0.402 \pm 0.003)eA^{-3}$$

TABLE IX(b). Comparison of Intrinsic Packing-Density of Porphyrins

Compound	Space Group	$\rho(eA^{-3})$*
Bis-piperidine-Fe(II)TPP[109]	$P\bar{1}$	0.381
OEP[110]	$P\bar{1}$	0.390
Ni(II)OEP[111]	$P\bar{1}$	0.388
Etio I[112]	$P2_1/c$	0.384
Co(II)OEP-1Me-imidazole[113]	$P2_1/c$	0.376
H_2OMg(II)TPP[114]	I4/m	0.377
H_2OZn(II)TPP[115]	I4/m	0.377
H_2OFe(III)(OH)TPP[116]	I4/m	0.374
Sn(IV)Cl_2TPP[70]	I4/m	0.380
VOTPP[101]	I4/m	0.375
Ni(II)OEP[117]	$I4_1/m$	0.373
MeOFe(III)-Porphyrin IXester[118]	I2/m	0.383
Ni-etio I[119]	$I4_1/m$	0.387

$$\langle (F(000)-M)/V \rangle = (0.380 \pm 0.006)eA^{-3}$$

$$*\rho = (F(000)-M)/V$$

angle between the planes of adjacent pyrrole rings is 1.2° while in molecule 1 the corresponding angle is 6.2°.

From Figures 8 and 9 it can be seen that the distances and angles of the two independent molecules are not exactly the same, and in general molecule 2 shows better internal consistency than molecule 1. It can also be seen from Tables III and IV that the peak heights for the atoms of molecule 2 are in general higher than the corresponding peak heights of molecule 1, as well as from Table VIII which lists the mean square displacements of the atoms along the principal axes of the ellipsoid. The thermal parameters of the molecule 1 are higher than those of molecule 2. Those differences could be due to packing effects which can more easily influence the more flexible peripheral part of the molecule but not the dimensions of the central hole. Indeed the germanium-nitrogen distances are exactly the same in both molecules, as well as the germanium-oxygen distances. The latter agree fairly well with the distance obtained from GeO_2 (space group $P4_2/mnm$) which is 1.85A(reference 25, p. 321) and with the distance obtained by simply adding the atomic radii of germanium and oxygen (1.85A, reference 25, p. 321). The carbon-oxygen bond of the methoxy group is short, particularly in molecule 1; the corresponding distance in simple alcohols and ethers is 1.43A, as compared with an average of 1.391A in the $PGe(OMe)_2$ structure.

The geometry of the germanium-methoxy group with respect to the macrocycle is different in the two independent molecules (Figure 14); the different interactions of the methoxy group with the atoms of the porphine is probably another reason why the molecules are on independent centers of symmetry and are not exactly alike.

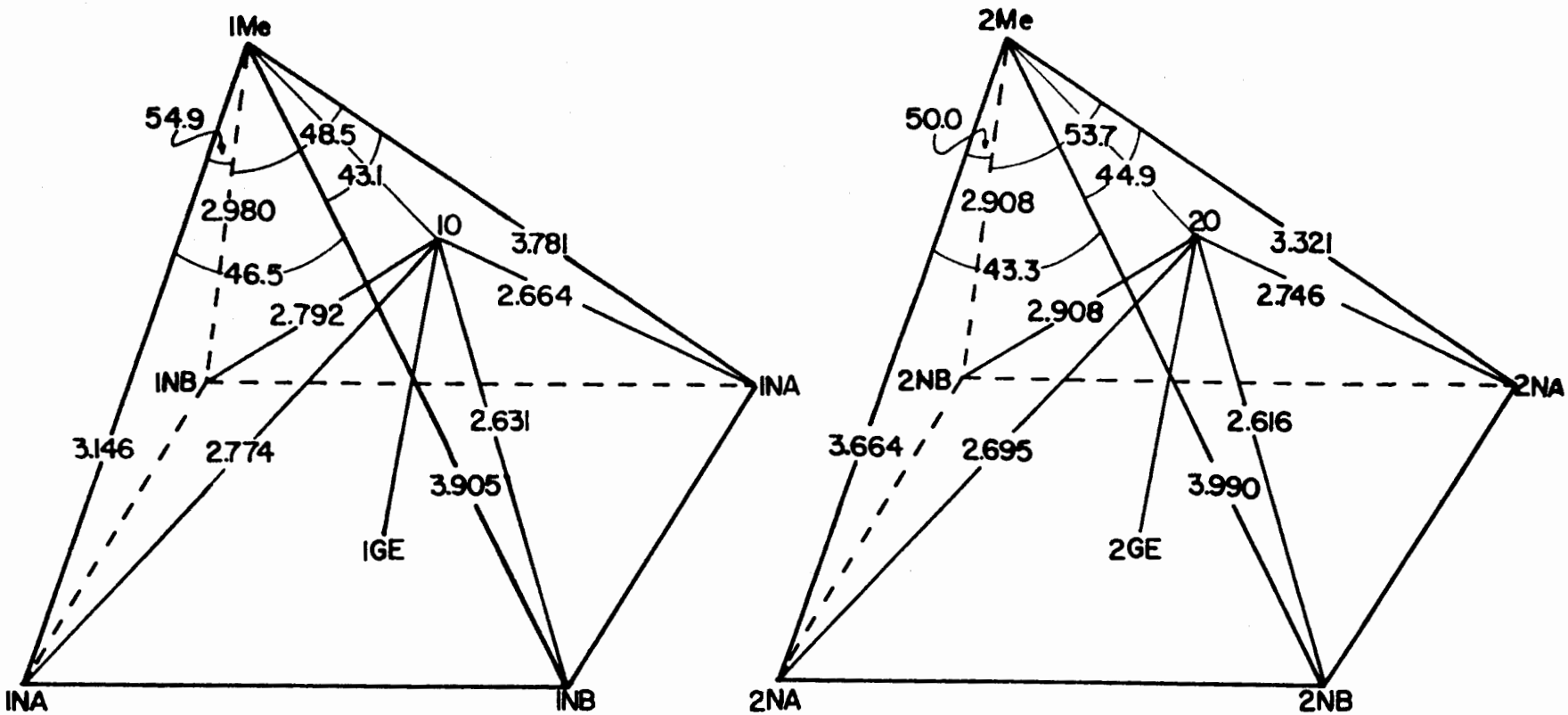


Figure 14. Geometry of the Ge-OMe group with respect to the pyrollic nitrogens of the porphine macrocycle.

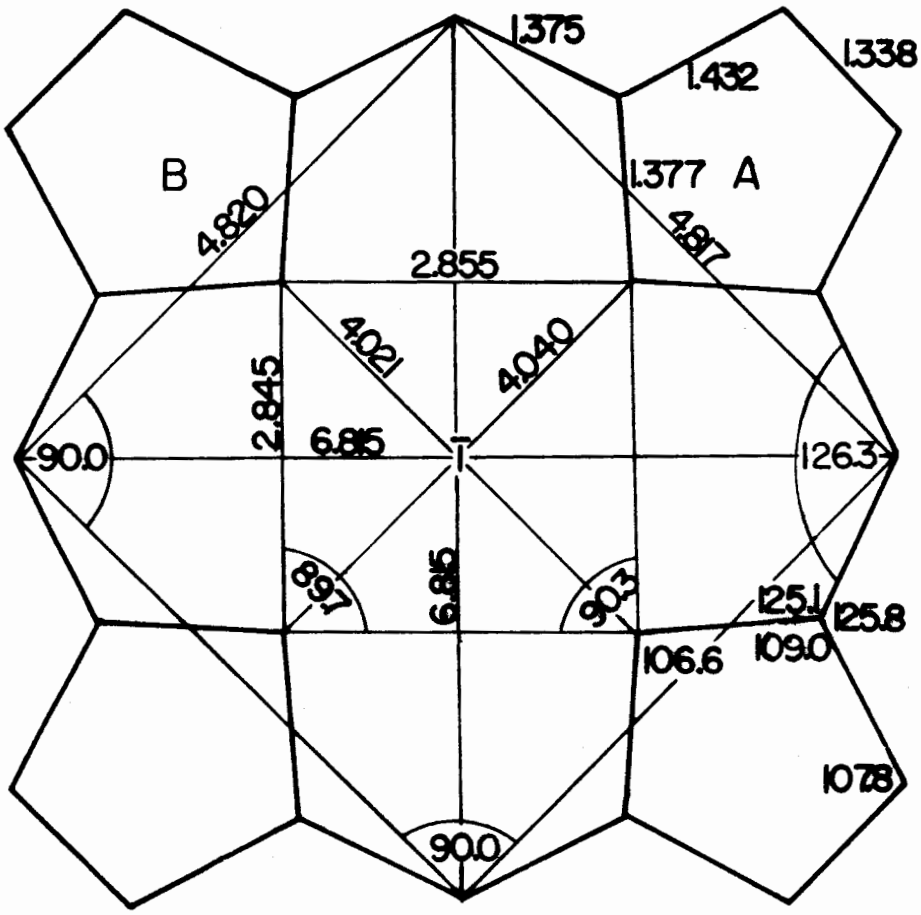


Figure 15. Average bond distances and angles of $PGe(OMe)_2$.

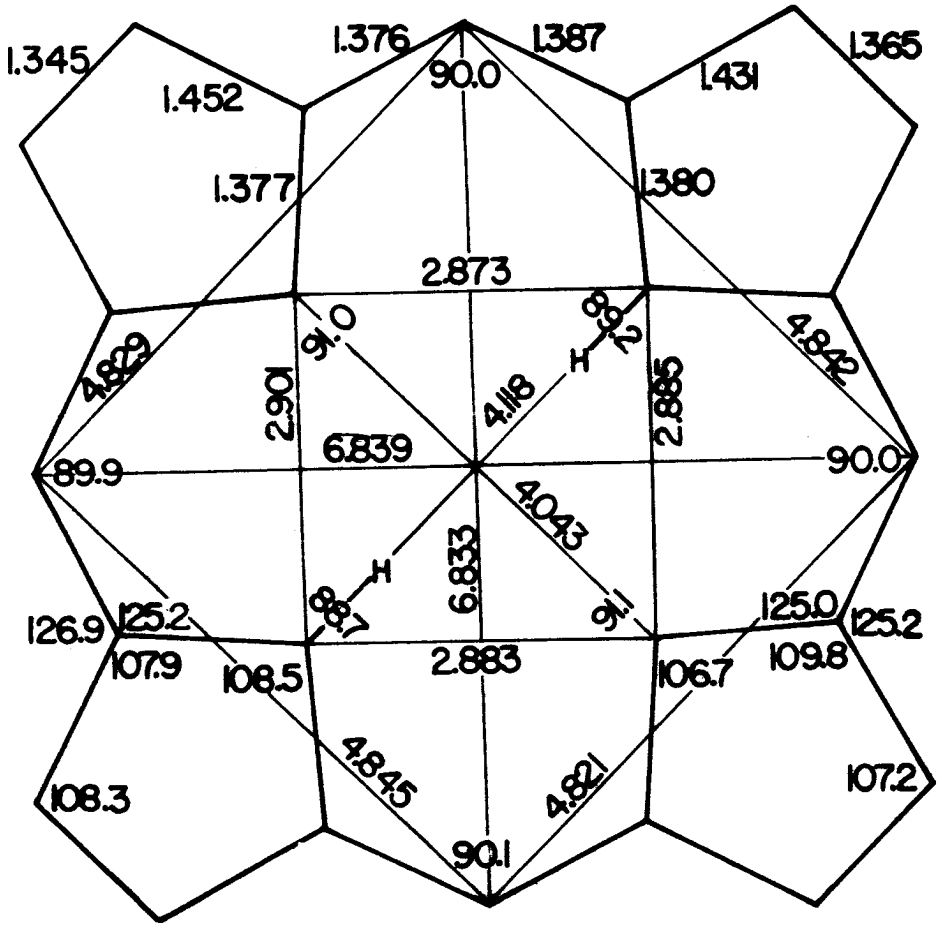


Figure 16. Average bond distances and angles of the free base porphine (from reference 39).

For a comparison of the $PGe(OMe)_2$ structure with similar systems, an average of the bond distances and angles over the two molecules has been taken. This average structure is shown in Figure 15, and an average structure of the free base porphine[39] is illustrated in Figure 16. By a direct comparison of the two average structures it can be seen that the angles of the pyrrole ring of the $PGe(OMe)_2$ structure are the same (within the experimental error) with the aza pyrrole of the porphine. In addition the average distances of the aza pyrrole of the porphine, with the exception of the C_α-C_β bond, are the same with the corresponding bond distances of the $PGe(OMe)_2$ porphine. The slight decrease of about 0.02A in the C_α-C_β bond in the latter might be an indication that the "electron path" a in Figure 4 is enhanced when the germanium atom substitutes the two pyrrolic hydrogens. In general the substitution of the two inner pyrrole hydrogens with a relatively non heavy metal tends to make all the pyrrole rings of the porphine equivalent and equivalent to the aza pyrroles of the free base as far as their bond distances and especially their angles are concerned. This phenomenon was also observed in the structures of n-propylporphineCu(II)[103] and octaethylporphine-Ni(II)[117], despite the fact that the metals Ge, Cu, and Ni are of entirely different nature. The complexation process of a metal by a porphine or porphyrin can be naively described by a two step process; first the two pyrrole hydrogens are removed and the two electrons are delocalized all over the macrocycle making all four pyrroles equivalent, and then, the positively charged metal simply neutralizes the negatively charged molecule. The geometrical constraints of the porphyrin systems are so overwhelming, that metals with a relative low electron content do not disturb the system and show the same effect of increasing the symmetry

of the macrocycle. The effects due to internal electronic distributions are very difficult to detect in the crystal structure. On the other hand, the size of the metal severely effects the geometry of the core. Metals with large radii which are forced due to symmetry reasons (for instance octahedral geometry) to remain on the plane of the macrocycle, can cause large changes in bond distances and angles. This was shown in the two recently determined structures of dichloridetetraphenylporphineSn(IV)[70] and dichlorideoctaethylporphineSn(IV)[71]. Because of the large tin-nitrogen distance (∿2.1A), the C_aNC_a angle is forced to increase by approximately 3° and the C_bC_b bond distance to increase by 0.02-0.03A in order to minimize the reduction of the already strained NC_aC_b angle. Thus the D_{4h} approximate symmetry is retained but the pyrroles have been changed from their free base state because of the inherent strain of the porphine macrocycle.

PART II

THREE DIMENSIONAL STUDY OF α-CHYMOTRYPSIN AT pH 8.7 AND 2.7 WITH DIFFERENCE FOURIER METHOD

PROLEGOMENA

Myoglobin and hemoglobin were the first globular proteins to have their three-dimensional structures solved by means of X-ray diffraction. J. C. Kendrew and M. F. Perutz received the Nobel Prize in chemistry in 1962 for this achievement[121].

Fifteen years later, the field of protein crystallography has grown to such an extent, that at least forty other protein structures have been solved to near-atomic resolution[122]. Electron density maps at lower resolution have been calculated for numerous other proteins; even a conservative estimate of these would be unwise to give here, because the number changes so rapidly with time[122,123]. The exponential growth is primarily due to the two following reasons: first the number of scientists involved in protein crystallography has been increased dramatically the last ten years, and second, the increased availability of very fast digital computers and automatic, almost self-controlled diffractometers has made the routine examination of such problems practical.

The three dimensional structure determination of a molecule of small molecular weight can still be difficult and tedious today due to the "phase problem" which is discussed in part I of this work. The determination of the phase angles for a protein crystal is much more difficult and time consuming because essentially none of the methods which have been developed for small molecules are applicable to protein structure determination[124]. In addition to the technical difficulties of the

interpretation of the Fourier image of a "solved protein structure"[124], the chemist-crystallographer also attempts to answer questions relating to the function of the protein molecule in living systems. He tries to unambiguously locate the "active site" within the molecule, that is, those parts of the protein which are mainly responsible for its specific action, and with the aid of the three-dimensional arrangement and additional experiments to propose a "mechanism" for that action. Moreover, he attempts to rationalize the overall topological features of the molecule, and if possible explain the evolutionary changes that have taken place in various species, by a careful comparison of the structural and chemical characteristics in a series of similar functioning proteins originating from different biological species, and by locating "invariances" around sensitive and critical parts of protein molecules. Trying to unfold even one of the forementioned "protein mysteries" is not an easy task so that pure empiricism is interwoven with sophisticated theory in an effort to achieve a reasonable answer[125,126]. However, one thing is certainly clear. Protein molecules are composed of the same fundamental constituents as other molecules and they should obey the basic principles that we already know. Therefore it should be a question of time, patience and a little luck before some of the current major problems are solved.

This work has been conducted in the spirit of providing missing pieces in the experimental picture of the proteolytic enzyme α-chymotrypsin, whose three-dimensional structure has been determined in our laboratory.

VII. INTRODUCTION

1. Enzymes: Historical

Enzymes are substances of large molecular weight that direct numerous chemical reactions that occur in all living organisms. We can envisage that the living organism functions through a matrix of chemical reactions of extraordinary complexity, which are controlled and catalyzed by many enzymes. The results of the action of enzymes were observed long before the nature of the responsible agents was recognized. Lazzaro Spallanzani noted in 1765 that gastric juice dissolved chunks of meat[127], but it was only after 68 years that the first clear recognition of an enzyme was made, when in 1833 A. Payen and J. F. Persoz conducted a more detailed study of the process of starch solubilization and showed that the responsible agent was something which they called "diastase". The name enzyme was introduced 45 years later by the physiologist W. Kuhne, and is derived from the Greek preposition en (εν), meaning in, and the Greek word zyme (ζύμη), meaning yeast[128]. In 1907 the German chemist Eduard Büchner was awarded the Nobel Prize in chemistry for his work on enzymes.

The first pure enzyme, which was named urease because it causes the hydrolysis of urea to ammonia and carbon dioxide, was obtained in 1926 by J. B. Summer. He also showed that urease was a protein, something of tremendous significance at that time. After this cornerstone discovery things proceeded at a faster rate. J. H. Northrop and M. Künitz

isolated pure crystalline pepsin and trypsin and showed that these enzymes are also proteins. By 1946, hundreds of enzymes had been purified and it had been shown beyond a doubt that all enzymes are of proteinic nature. For their work on enzymes, in 1949 Summer and Northrop shared the Nobel Prize in chemistry. In the following years, the study of enzymes became more systematic and thorough. In 1963, the primary structure of the enzyme bovine pancreatic ribonuclease had been established[129]; in 1965 and 1967, the three-dimensional structures of lysozyme[130] and pancreatic ribonuclease[131] were reported respectively, and finally in 1969, the laboratory synthesis of bovine pancreatic ribonuclease was announced[132].

All enzymes are proteins, ranging in molecular weight from about 10,000 to several million amu. They are very sensitive to changes in temperature and chemical environment and particularly sensitive to changes in pH: their action can be completely inhibited with relatively small changes in pH. They show a remarkable degree of specificity for the reactions which catalyze. Their action is usually attributed to a small portion of their total content, the so called "active site". It seems that the rest of the molecule is inactive, that is, it does not participate directly in chemical catalysis, but rather, it helps in maintaining the geometry of the active site and its immediate environment in a manner appropriate for binding substrate for reaction; this inactive part is also connected with regulatory and protective properties.

The exact manner in which an enzyme acts has not yet been elucidated. It is well known that the specific action of enzymes is related to certain chemical groups which are "catalytically active" (from a completely empirical point of view), such as hydroxyl (-OH), carboxyl (-COOH)

or amino ($-NH_2$) groups. For instance, an aliphatic group has never been classified as necessary for the activity of an enzyme. It seems that the mobility and the special properties of the hydrogen ion (the H^+ has an exceptionally small radius $\sim 10^{-13}$ cm and due to this property it can cause very large polarizations in the nearby groups, and it can also act as a wave packet rather than as a particle) play a fundamental role in the mechanism of enzymatic action. It is also known that the geometrical structure or the charge distribution in space (commonly called three-dimensional structure) of an enzyme is very critical for its function. All else concerning mechanisms of enzymatic action is speculative. For this problem to be approached more effectively, it probably should be examined in a less traditional way. As Baron Francis Bacon said: "It would be an unsound fancy and self-contradictory to expect that things which have never yet been done, can be done except by means which have never yet been tried".

2. Alpha Chymotrypsin: A General Discussion

Alpha Chymotrypsin (hereafter denoted as α-CHT) is a proteolytic enzyme of chemical composition $C_{1113}N_{300}O_{349}H_{1752}S_{12}$, consisting of three peptide chains A, B, C of 13, 131 and 97 residues, respectively. It has a molecular weight of 25,305 amu and its sequence was determined in 1964 by Hartley[133,134] and Meloun, et. al.,[135] and it was revised by Blow, et. al., in 1969[136]. Figure 17 shows the sequence of α-CHT. It belongs to a more general class of proteins which are called "Serine Proteinases" where a reactive serine plays an important role in activity; this residue reacts with diisopropylfluorophosphate $\{((CH_3)_2CHO)_2\text{-}\underset{\underset{O}{\|}}{P}\text{-}F\}$ leading to complete loss of enzymatic activity. α-CHT is formed from an

Figure 17. The sequence of α-CHT.

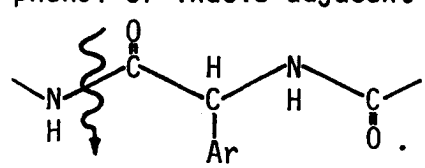
Val·Leu·Ser·Gly·Leu————Ile·Val·Asn·Gly·Glu·Glu·Alg·Val·Pro·Gly·Ser·Trp·Pro·Trp
21
Pro Gln
Gln Glu·Asn·Ile·Leu·Ser·Gly·Gly·Cys·Phe·His·Phe·Gly·Thr·Lys·Asp·Gln·Leu·Ser·Val
S 40
Ilu Asn S
Ala Trp·Val·Val·Thr·Ala·Ala·His·Cys·Gly·Val·Thr·Thr·Ser·Asp·Val·Val·Val
57
Pro Ala
Val Ile·Lys·Leu·Lys·Gln·Ile·Lys·Glu·Ser·Ser·Ser·Gly·Gln·Asp·Phe·Gly·Gly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

inactive precursor, chymotrypsinogen, by cleavage of the peptide bond by trypsin between the residues ARG15 and ILE16. The molecule which is formed, known as π-chymotrypsin, is completely active (Figure 18). Removal of the dipeptide SER14-ARG15 from π-chymotrypsin gives δ-chymotrypsin and further cleavage and removal of the dipeptide THR147-ASN148 gives α or γ chymotrypsin; all the subsequent chymotrypsins are also active.

The structural differences between α- and γ- chymotrypsin are not well understood. Under acidic conditions (pH∿ 4), α-CHT forms a dimer which crystallizes in the monoclinic system, but at pH=5.6 the molecule is a monomer and crystallizes in the tetragonal system[137]. The monoclinic form was named α and the tetragonal γ . Chemically the two forms are identical. α-CHT is stable at pH∿4 but converts easily to γ-chymotrypsin under the appropriate pH conditions. The reverse transition is very slow in comparison with the former; apparently under the influence of the pH conditions, the molecule transforms to a conformation which is more favorable thermodynamically.

Although all four forms of chymotrypsin are catalytically active and very similar in their specificity and action, the α-form is the one that has been studied most extensively, because relatively speaking, it is the most readily obtained free of contamination[138].

The biological function of α-CHT is the catalysis of peptide bond hydrolysis of polypeptide substrates in the mammalian small intestine. It has a preferential specificity for peptide bonds with an aromatic residue, like phenyl, phenol or indole adjacent to the carbonyl[139]



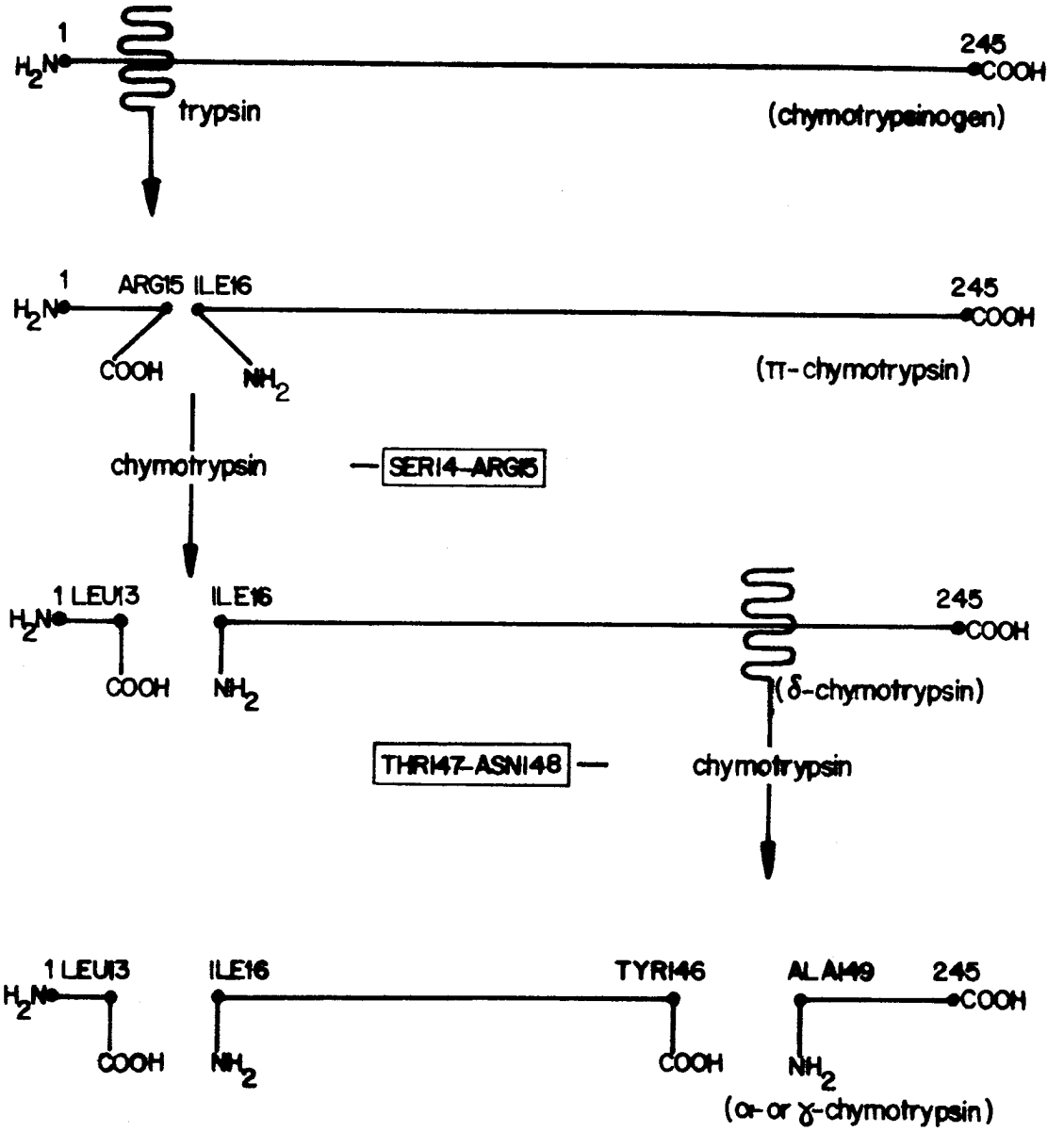


Figure 18. Formation of α-(or γ-) CHT from its inactive precursor chymotrypsinogen.

However, the enzyme does not have a unique specificity, for it can more or less hydrolyze almost any kind of a peptide bond, although at a much slower rate. This lack of specificity probably renders α-CHT an ideal model for study, because it can reveal the general qualifications of an active site which is able to work effectively with many different kinds of substrates.

The proposed reaction pathway for α-CHT is a three step process

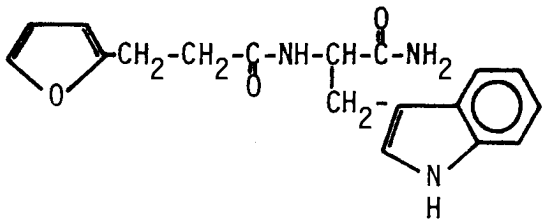
$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} ES' \xrightarrow{k_3} E+P_2 \ ,$$

$$\begin{array}{ccc} ES & \xrightarrow{k_2} & P_1 \end{array}$$

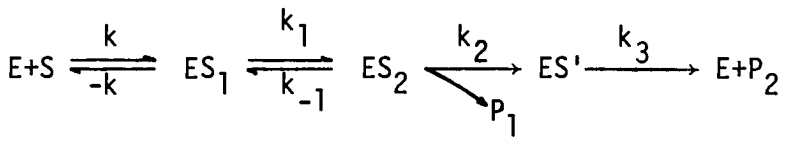
① ② ③

where E is the enzyme, S the substrate, ES a noncovalent enzyme-substrate complex and ES' a covalent acyl-enzyme intermediate. P_1 is the initially released product of hydrolysis and P_2 is the carboxylic moiety. There is a vast amount of literature in support of the above sequence of reactions, or at least that a three step process adequately describes the gross features of the hydrolytic reaction[140-142]. The last step is believed to be the rate determining step ($k_3 << k_2$).

The simplicity of the foregoing process is appealing, even though it is understood how complicated the phenomenon can be, and how difficult it is to deconvolute a mechanism, for which one of the reactants has a molecular weight of 25,305 amu. Indeed, more recent work by Hess et. al.,[143] with the substrate furyl acryloyltryptophanamide



shows another intermediate between steps (2) and (3) shown previously,



where ES_1 and ES_2 are different <u>conformers</u> of the substrate-enzyme complex.

The activity of almost any enzyme depends (among other things) on the pH of the surrounding medium; α-CHT conforms to the rule. Optimal catalytic activity is between pH 7.5-8.5; above and below this range the activity falls off sharply. The effect of pH on α-CHT will be discussed in the last section of this introduction.

3. The Three Dimensional Structure of α-CHT

At about pH 3.6, α-CHT crystallizes in the monoclinic crystal system, space group $P2_1$, with four molecules per unit cell or two molecules per asymmetric unit. The two molecules in the asymmetric unit are related to each other by a local (non-crystallographic) two fold rotation axis, which by the very nature of being non-crystallographic is not exact. It runs approximately parallel to the $\vec{a}^*$ axis along the interface of the independent molecules which form a dimeric unit. Figure 19 shows a schematic representation of the dimeric structure in terms of the monomeric units; the view is parallel to the local two fold axis.

The three dimensional structure of the enzyme was first reported by D. M. Blow and his collaborators at Cambridge, England (group I) in 1967-1968[144,145]. The papers that followed[136,146-148] describe their

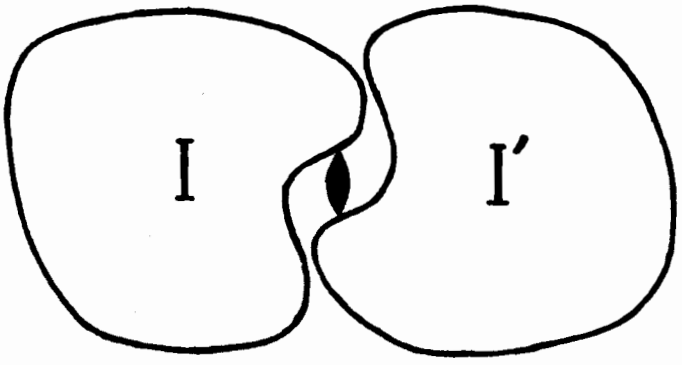


Figure 19. Schematic representation of the dimeric unit of α-CHT. View parallel to the local two fold axis.

work in detail. The structure of α-CHT was also determined independently by A. Tulinsky and his coworkers here at Michigan State University (group II)[149,150]. The two groups followed a different approach and philosophy, the latter's approach essentially being more general and more careful. Some of the differences between the work of the two groups are summarized below:

Group I

(a) They used tosyl-α-CHT instead of native α-CHT; tosyl fluoride (H_3C—⟨○⟩—SO_2-F) reacts stoichiometrically with α-CHT and forms a sulfonic acid ester with the hydroxyl group of SER195 which is located in the active site and is directly associated with the activity of the enzyme. For obvious reasons, the insertion of the tosyl group detracts from the determination of the geometry of the active site region of the native enzyme.

(b) The electron density distribution of the monomeric units was averaged about the local two fold axis and the average map was used to obtain a structure for monomeric α-CHT[151].

(c) Dioxane was added to the mother liquor of crystallization to reduce twinning of the crystals[144,152]; dioxane was found to bind in the active site region.

(d) The collection of X-ray intensity data was by the use of photographic methods.

Group II

(a) Native α-CHT was used for and throughout the structure determination.

(b) The structure determination was not biased by the use of non-crystallographic symmetry,and the structure of the dimeric unit was deduced.

(c) Dioxane was not used during the crystallization process; instead,crystals with minimal twinning were selected for intensity data collection, and the data were corrected for the twinning.

(d) The X-ray intensity data were collected by diffractometric techniques[78], which in general, can be considered superior to photographic techniques[77].

The two independent structure determinations agree well in the grosser aspects, but differences exist, mainly on the surface of the molecules and in the interface region of the dimer, close to the active site[150]. It is difficult at present to assess the importance of these differences in terms of activity of the enzyme; time should eventually prove to be the judge.

The three dimensional structure of α-CHT has been discussed in excellent papers by Blow[147,148]; Vandlen[153] also discusses the structure in a fairly detailed manner in a Ph.D. thesis. Due to the fact that the results of this work are intimately related to the three dimensional structure of α-CHT, an attempt will be made to summarize the most important features of the molecule revealed by the crystallographic work.

As mentioned already, α-CHT crystallizes in the monoclinic crystal system, space group $P2_1$ (C_2^2) and the unit cell parameters for the native enzyme (pH∿3.6) determined from diffractometer measurements are[149]

$$|\vec{a}| = 49.24(7) \text{ A}$$

$$|\vec{b}| = 67.20(10) \text{ A}$$

$$|\vec{c}| = 65.94(9) \text{ A}$$

$$\beta = 101.79(8)^{\circ} .$$

The manner in which the four molecules are packed in the unit cell is shown in Figure 20[154]. The dimeric unit which is formed from

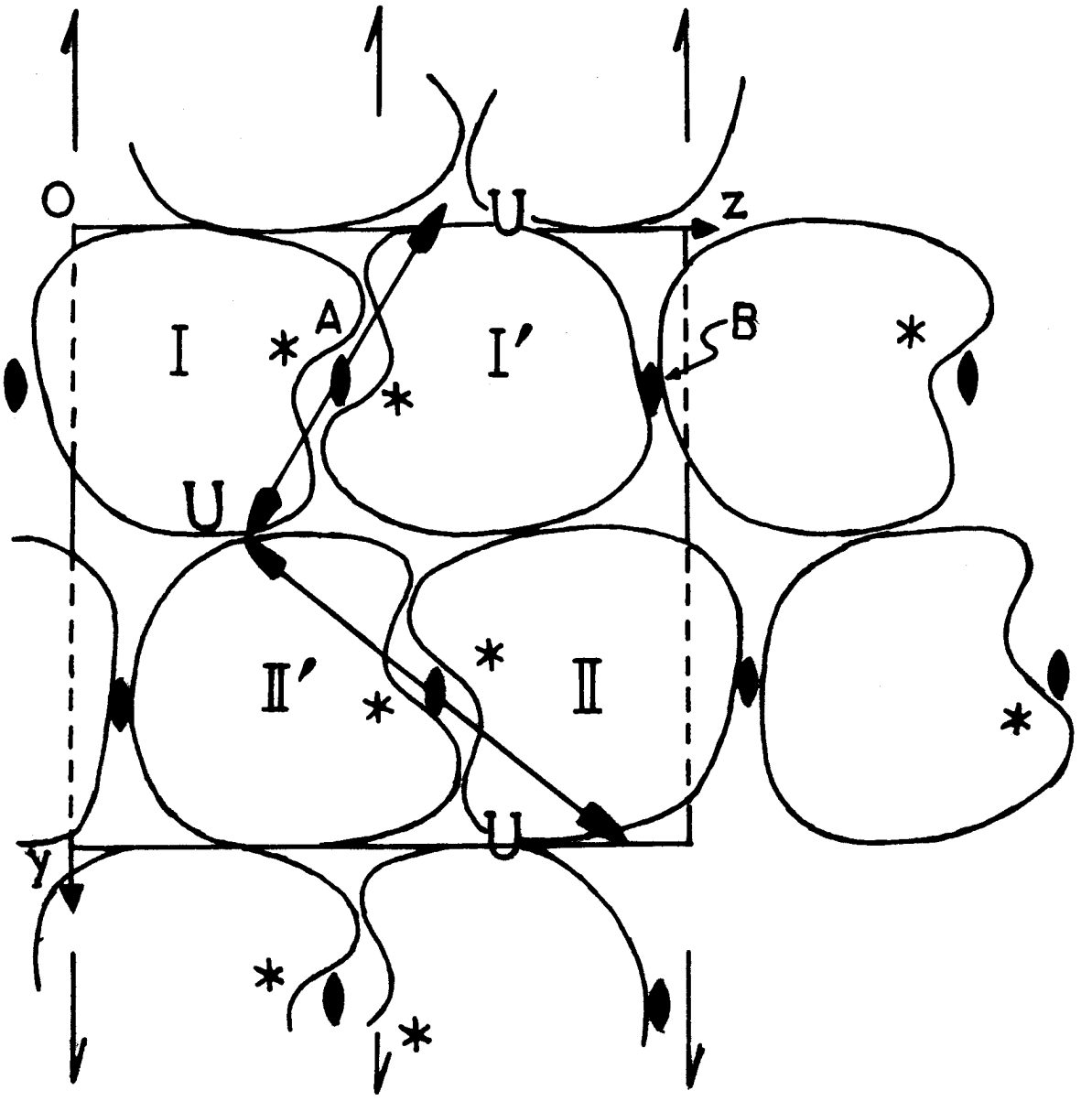


Figure 20. Schematic packing diagram of α-CHT viewed along the $\vec{a}$* direction. Molecules I and I' form an asymmetric unit and are related by non crystallographic two fold rotation axes A and B; active site regions near center of dimer designated with asterisks; uranyl binding region designated by U (reference 154).

two α-CHT molecules can be described as an ellipsoid with approximate dimensions of 45 by 33 by 67A, with the axes of the ellipsoid corresponding to the crystallographic directions $\vec{a}^*$, $\vec{b}$ and $\vec{c}$. The monomeric units are held together through electrostatic interactions, van der Waals forces and hydrogen bonding interaction. Table X summarizes some of the main interactions of the monomeric units about the local two fold axis (dyad A, Figure 20). The monomeric unit has an overall spherical shape of approximate dimensions 45x33x34A with the active site residing in a crevice about 7A deep.

The folding of the molecule is complex; only ten residues (235-245) at the carboxyl terminal of C-chain are in a relatively undistorted α-helical conformation. A distorted segment of α-helix is also formed in the C-chain with the residues 164-173. The antiparallel pleated-sheet structure[155] is much more common in α-CHT, with rather extensive regions of distorted sheets[148]. The non polar groups are buried in the interior of the molecule; a manifestation of this is the formation of a spherical cavity in the interior of the molecule, whose surface is composed of the non polar side chains of 18 residues[150,153]. Almost all polar groups are on the surface of the α-CHT molecule being accessible to solvent molecules, except for three catalytically important residues ILE16, ASP194 and ASP102, which are relatively buried. The positively charged α-amino group of ILE16 interacts electrostatically with the negatively charged carboxyl group of ASP194 to form an ion pair located in the interior of the molecule. Fersht[156] found the free energy of the removal of a charge from this ion pair to be +2.9 kcal $mole^{-1}$ from the ratio of equilibrium constants of active and inactive conformers. The distance between the $-NH_3^+$ of ILE16 and the $-COO^-$ of ASP194 is approximately 4A

TABLE X. Some Molecular Interactions About the Local Two Fold Axis

α-NH_3^+ (ALA149)	$\longleftrightarrow$	β-COO^- (ASP64)
α-COOH (TYR146')	$\longleftrightarrow$	carbonyl of the peptide bond 57 (hydrogen bond)
or		
α-COO^- (TYR146')	$\longleftrightarrow$	Imidazole(+) of HIS57
side chain MET192	$\longleftrightarrow$	side chain of MET192'
Phenyl group of PHE39	$\longleftrightarrow$	phenyl group of PHE39'

(as measured directly from the Kendrew model of α-CHT), and the electrostatic energy calculated from Coulomb's law and assuming a dielectric constant of one for the interior of the molecule, is approximately -80 $kcalmole^{-1}$. Allowing for a higher dielectric constant of ∿5 for the protein matrix, an approximate electrostatic energy of -16 $kcalmole^{-1}$ is obtained, which is very significant. Thus, it is believed that this buried ion pair and its conformational ramifications play a unique role in the overall activity of the enzyme.

The catalytic site of α-CHT involves three residues, ASP102, HIS57 and SER195, whose importance for activity has been confirmed by chemical and crystallographic studies[146,157-159].

4. Effect of pH on α-CHT

There are numerous factors upon which the three dimensional structure and activity of a protein depends[160]. Ignoring solvent, some of the most important interactions encountered in proteins are the following:

(a) Hydrogen bonding

(b) van der Waals interactions

(c) Dipole-dipole interactions

(d) Electrostatic interactions.

From a biochemical point of view the above list is by no means exhaustive[160], even though it is difficult to distinguish interaction (a) from (d)[161], and (c) from (d), and it seems that the inclusion of more "potentials" can lead only to circular arguments and confusion. The properties of the aqueous environment of a protein are inextricably related with the forementioned interactions, so that by monitoring

various parameters of the solvents, such as the ionic strength or the hydrogen ion concentration, these complicated systems can be studied in an efficient way.

α-CHT shows a complex dependence upon the pH of the surrounding medium; the transformation of α-CHT to γ-chymotrypsin and vice versa by merely changing the pH has already been mentioned. Hess and co-workers[162,163] have shown that α-CHT is catalyticaly active only if the α-amino group of ILE16 is protonated, that is, if the salt bridge between ILE16 and ASP194 is maintained. In addition, the activity of the enzyme sets in with the deprotonation of HIS57 around pH 6.8[164], it then reaches a maximum around pH 7.5-8.5, depending on the substrate and the solvent, and finally, it falls off sharply because of deprotonation of the α-amino terminal of ILE16 and the disruption of the internal salt bridge. Figures 21a, 21b show the almost Gaussian-type dependence of the activity of α-CHT upon pH[164], and Figure 22 shows the dependence of the dimerization of the enzyme upon pH[166].

It has also been shown that the optical rotation of α-CHT is pH dependent[167]; Figure 23 shows the pH dependence of the specific rotation at 340nm and 12°C (protein concentration 1.2mg/ml and ionic strength of 0.14M). The "jump" of optical rotation at pH∿8 has also been attributed to the deprotonation of the positively charged amino group of ILE16 and the disruption of ILE16⟷ASP194 salt bridge.

It was mentioned previously that almost all the ionizable groups of α-CHT are on the surface of the molecule and the surface charge is heavily pH dependent. A quantitative treatment of this problem is very difficult but a qualitative approach based on a simple model is easy. α-CHT can be represented as a sphere with a uniform distribution

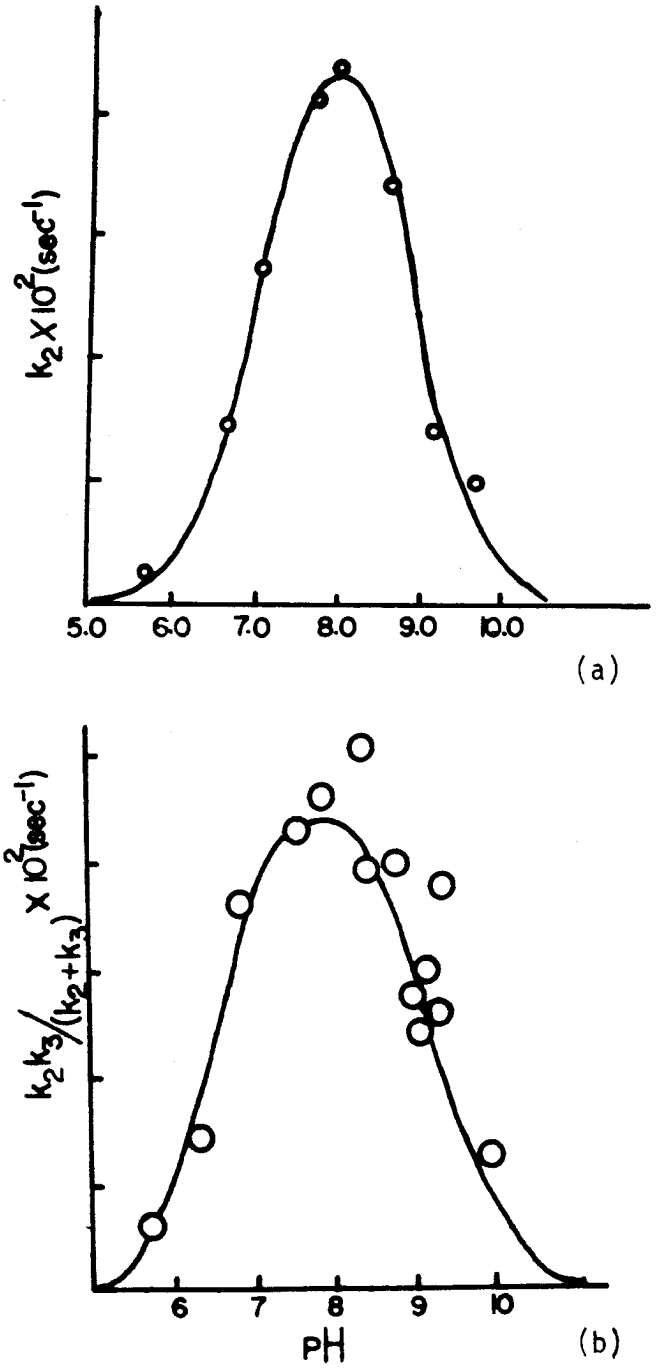


Figure 21. The α-CHT-catalyzed hydrolysis of N-acetyl-L-tryptophan amide (a), and of N-acetyl-L-phenylalanine (b) in water at 25°C as a function of pH; kinetic constants are defined in the following equation:

$$E+S \overset{K_s}{\rightleftharpoons} ES \xrightarrow{k_2} ES' \xrightarrow{k_3} E+P_2$$
$$ES \rightarrow P_1$$

(reference 165).

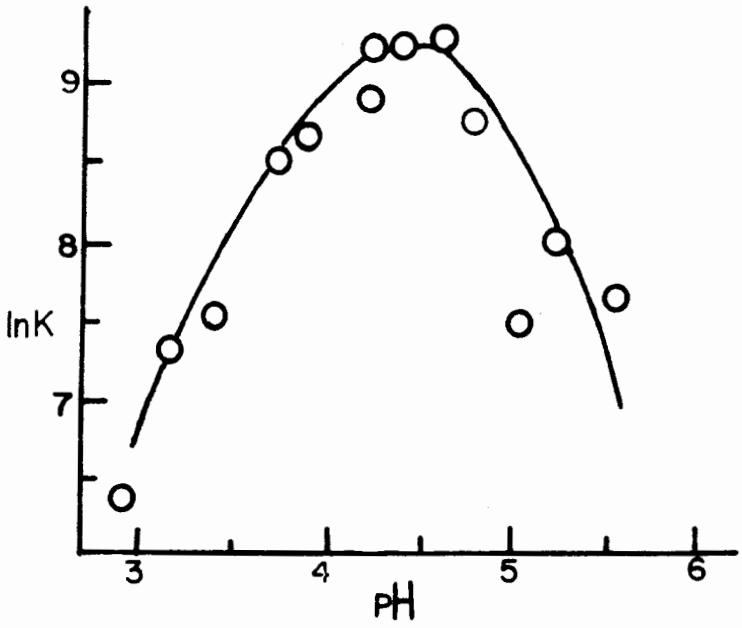


Figure 22. Dependence of the natural logarithm of the dimerization equilibrium constant on pH. Conditions: 0.01M NaCl, 0.01M acetate buffer, 25°C; maximum dimer formation occurs at pH 4.4 (reference 166).

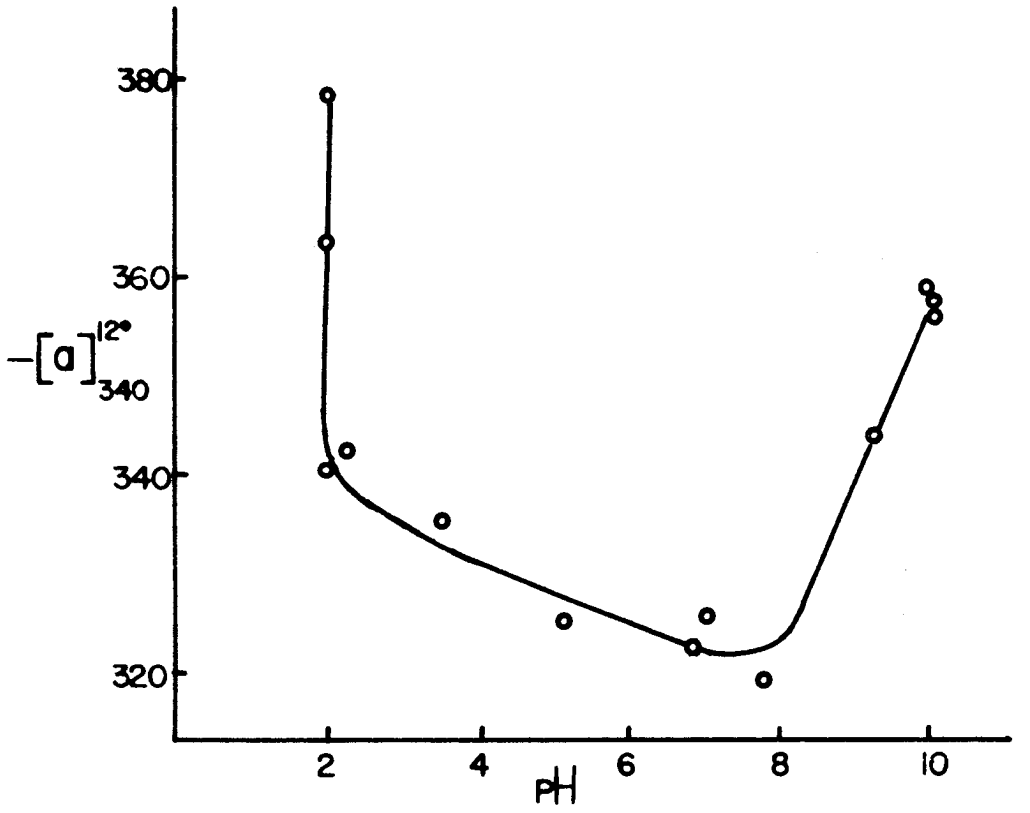


Figure 23. The pH dependence of the specific rotation of α-CHT at 340 nm and 12°C (reference 167).

of charge on its surface, embedded in a "sea" of ions and water molecules. Figure 24 shows such a schematic representation when the protein molecule is positively charged in ammonium sulphate solution. The NH_4^+ and SO_4^{2-} ions shown in Figure 24 are not of particular importance, but $(NH_4)_2SO_4$ frequently is used to crystallize proteins. Clearly a uniform charge distribution as shown in Figure 24 is not a realistic model but it will serve our purpose. We can consider at least two layers of counter ions surrounding the protein with layer 1 being more "ordered" than layer 2 . The boundaries of these two layers are not well defined due to the mobility of the different counter ions and water molecules. That an "ordering" of solute ions does exist can be seen from the low order diffraction patterns of proteins, whose intensity depends on the concentration and kind of salt used during crystallization. "Salt difference Fourier maps" employing low order diffraction data at different salt concentrations can reveal the shape of protein molecules or subunits and their relative positions[168]. α-CHT contains 43 ionizable residues, including the terminal α-carboxyl and α-amino groups of the three peptide chains. Table XI is a compilation of these groups; the numbers which follow the abbreviated name of an amino acid correspond to the position of this residue in the primary structure of the enzyme (Figure 17); circled numbers give the number of times the particular residue in the state shown appears in the sequence of α-CHT. The amount and algebraic sign of charge on the surface of the molecule will depend on the ionization state of these residues. Table XII gives the pK's of the most important amino acids along with their chemical formulas and their common shorthand notation[169]. However, the pK_α's of these amino acids might be quite different in a protein for the

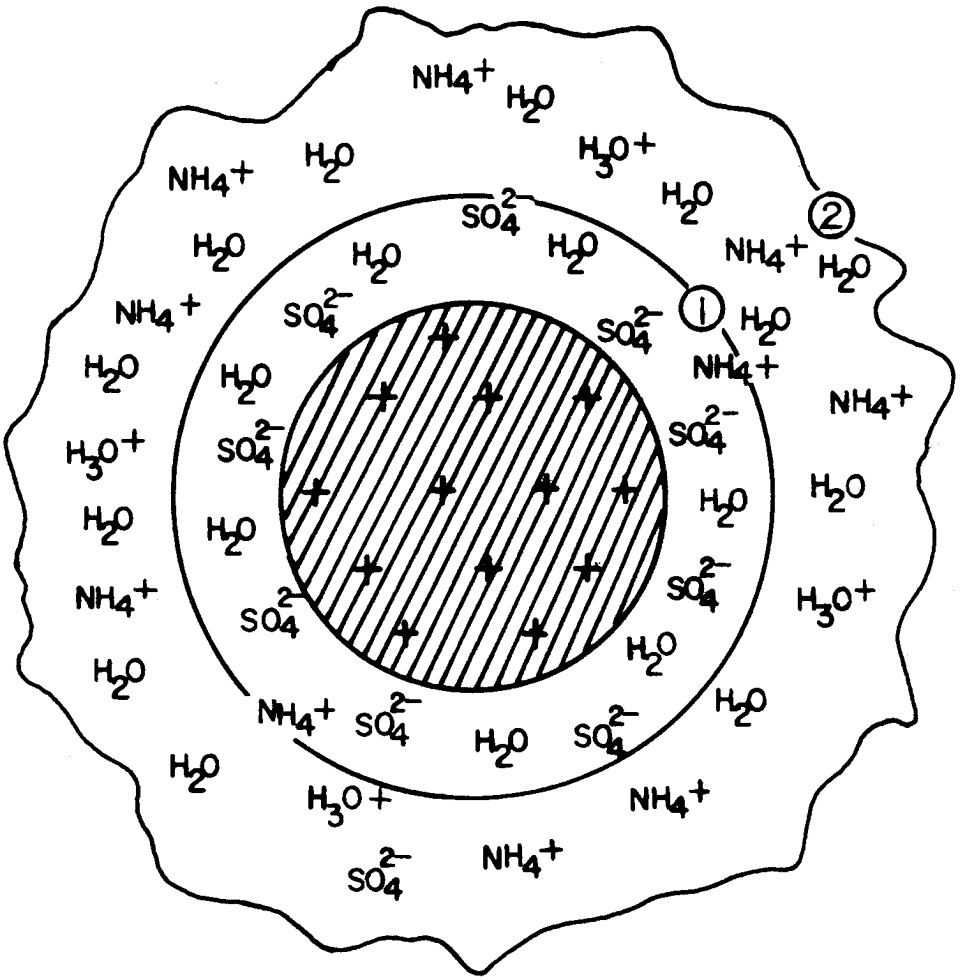


Figure 24. Model of a globular protein molecule, based on the assumption that charge is uniformly distributed over its surface.

TABLE XI. Ionizable Residues in α-CHT.

Basic Residues

LYS	:	36, 79, 82, 84, 87, 90, 93, 109, 169, 170, 175, 177, 202, 203	14
ARG	:	145, 154, 230	3
HIS	:	40, 57	2
CYS-NH_2	:	1	1
ILE-NH_2	:	16	1
ALA-NH_2	:	149	1

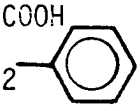
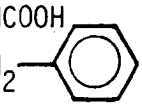
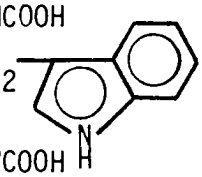
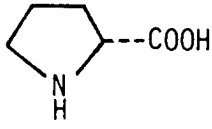
Acidic Residues

ASP	:	35, 64, 72, 95, 102, 128, 129, 153, 178	9
GLU	:	20, 21, 49, 70, 78	5
TYR-OH	:	94, 146, 171, 228	4
LEU-COOH	:	13	1
ASN-COOH	:	245	1
TYR-COOH	:	146	1

TABLE XII. Amino Acids that Commonly Occur in Proteins (reference 169).

Name	Formula	Values of pK_α
Glycine(GLY)	H_2NCH_2COOH	2.34; 9.6
Alanine(ALA)	$H_2NCHCOOH$ $\vert$ CH_3	2.35; 9.70
Valine(VAL)	$H_2NCHCOOH$ $\vert$ $CH(CH_3)_2$	2.32; 9.62
Leucine(LEU)	$H_2NCHCOOH$ $\vert$ $CH_2CH(CH_3)_2$	2.36; 9.60
Isoleucine(ILE)	$H_2NCHCOOH$ $\vert$ $CH(CH_3)CH_2CH_3$	2.36; 9.68
Serine(SER)	$H_2NCHCOOH$ $\vert$ CH_2OH	2.21; 9.15
Threonine(THR)	$H_2NCHCOOH$ $\vert$ $CH(OH)CH_3)$	2.63; 10.43
Aspartic Acid(ASP)	$H_2NCHCOOH$ $\vert$ CH_2COOH	2.09 (α-carboxyl); 3.86 (β-carboxyl); 9.82
Asparagine(ASN)	$H_2NCHCOOH$ $\vert$ CH_2CONH_2	2.02; 8.8
Glutamic Acid(GLU)	$H_2NCHCOOH$ $\vert$ CH_2CH_2COOH	2.19 (α-carboxyl); 4.25 (γ-carboxyl); 9.82

TABLE XII. (Continued)

Name	Formula	Values of $pK_{\alpha-}$
Glutamine(GLN)	$H_2NCHCOOH$ $CH_2CH_2CONH_2$	2.17; 9.13
Lysine(LYS)	$H_2NCHCOOH$ $(CH_2)_4NH_2$	2.18; 8.95 (α-amino) 10.53 (ε-amino).
Histidine(HIS)	$H_2NCHCOOH$ CH_2	1.82; 6.0 (imidazole); 9.17
Arginine(ARG)	$H_2NCHCOOH$ $(CH_2)_3NH(C{=}NH)NH_2$	2.17; 9.04 (α-amino); 12.48 (guanidino)
Phenylalanine(PHE)	$H_2NCHCOOH$ CH_2 	1.83; 9.13
Tyrosine(TYR)	$H_2NCHCOOH$ CH_2 	2.20; 9.11 (α-amino) 10.07 (phenolic hydroxyl)
Tryptophan(TRP)	$H_2NCHCOOH$ CH_2 	2.38; 9.39
Cysteine(CYSH)	$H_2NCHCOOH$ CH_2SH	1.71; 8.33 (sulfhydryl); 10.78 (α-amino)
Methionine(MET)	$H_2NCHCOOH$ $(CH_2)_2\text{-}S\text{-}CH_3$	2.28; 9.21
Proline(PRO)		1.99; 10.60

following reasons[170]:

(a) Changes in the electrostatic environment associated with the formation of the peptide bond.

(b) Electrostatic interactions of dissociable groups with neighboring atoms in the protein which may alter the acidity or basicity.

(c) Combination with ions other than hydrogen ions which may change the electrostatic condition of the protein and affect the hydrogen ion equilibrium.

Therefore the same amino acids in different proteins might have fairly different pK values. The dependence of the pK of a particular residue upon the pH of the solvent is given approximately by the following equation[171]:

$$pH=p(K_{int})_j+\log \frac{r_j}{N_j-r_j} -2\frac{ZW}{2.303} \qquad (98)$$

where

$(K_{int})_j$: intrinsic dissociation constant of the amino acid j

N_j : total number of groups of class j

r_j : the number of amino acids belonging to class j and which have been dissociated at a given pH

Z : the average net charge on the molecule

W : an electrostatic factor defined as

$$W=\frac{Ne^2}{2\varepsilon RT}\left(\frac{1}{b} - \frac{\kappa}{1+\kappa a}\right) \qquad (99)$$

where

N : Avogadro's number

e : electronic charge

ε : dielectric constant of the solvent

R : gas constant

T : absolute temperature

b : the radius of the spherical protein

a : the radius of exclusion

κ : quantity defined in Debye-Hückel theory as

$$\kappa=\left(\frac{8\pi e^2 N}{1000kT}\mu\right)^{\frac{1}{2}}$$

where

μ : ionic strength

k : Boltzmann's constant

The derivation of the above equation is based on the assumption that a protein molecule can be represented as a spherical conductor[171].

Although the above discussion shows how difficult it is to carry out quantitative calculations concerning solvent medium effects on ionization equilibria of proteins, the total charge residing on α-CHT can be estimated in different pH regions by assuming that the pK's of Table XII are $pK_{int}\pm2$. At low pH, both acidic and basic groups of α-CHT are protonated, hence the molecule bears its maximum net positive charge. As the pH increases the prototropic groups dissociate in the order of decreasing acidity, the carboxyl groups from about 2 to 5, the imidazolium ion from about 5 to 7, followed by ammonium and phenolic hydroxyl groups and finally by the guanidinium ion. When the simplified idea, that the net charge on the macroion depends <u>only</u> on the ionizable groups is applied to α-CHT, one obtains with the aid of Table XI and XII a charge distribution as a function of pH; the pH ranges and the corresponding total algebraic charge are given in Table XIII. Figure 25 is a histogram representing the results compiled in this table. Due to the uncertainties in the pK's of the amino acid residues in the protein, Figure 25 or parts of it may be shifted to the left or right, while

TABLE XIII. Charge Distribution of α-CHT as a Function of pH

pH range	Description of the ionization state of different residues	Total algebraic charge(Q) on α-CHT
1-2	acidic amino acids neutral, basic amino acids positively charged	+22
2-3	all terminal carboxyl groups loose their protons and become negatively charged	+19
3-5	all acidic amino acids loose their proton and become negatively charged	+5
5-7	the two histidine residues (40, 57) become neutral	+3
7-9	CYS-NH_3^+1, ILE-NH_3^+16 and ALA-NH_3^+149 loose their proton	0
9-10.5	all four tyrosines loose their hydroxyl hydrogen	-4
10.5-12	lysines are deprotonated	-16
12-14	Arginines are deprotonated	-19

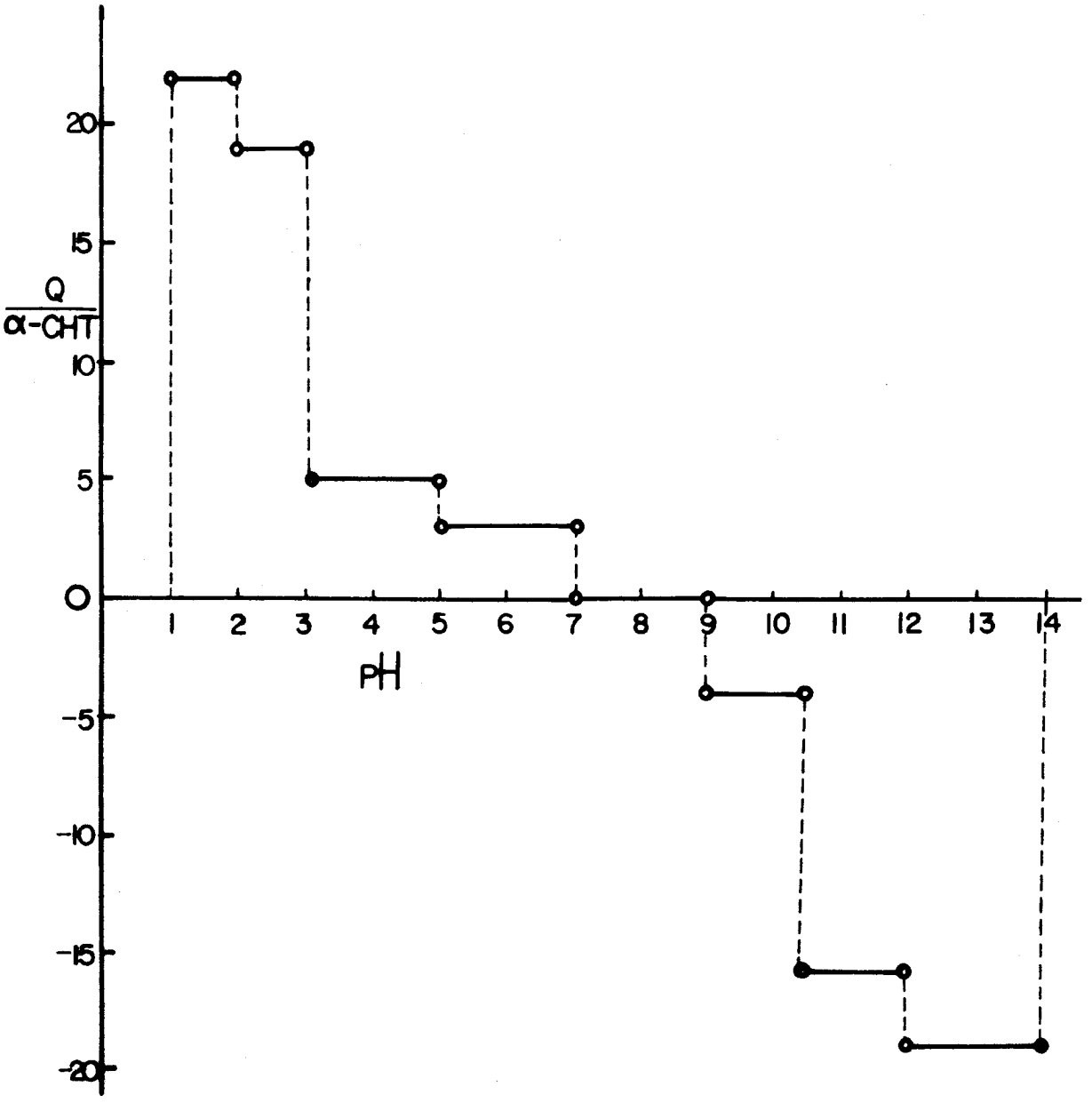


Figure 25. Histogram of total algebraic charge (Q) on α-CHT molecule as a function of pH.

the lengths of the various segments vary. However, it is clear that abrupt changes in the amount of charge residing on the molecule can occur as the pH varies from the one extreme to the other, and consequently changes in the molecular structure might occur as well[154,172,173].

Vandlen and Tulinsky[154] discuss the changes in the structure of crystalline α-CHT by changing the pH from 3.6 (native) to approximately 5.4. Although they found some significant changes, the change of 2 pH units from native α-CHT apparently was not large enough to deprotonate some of the more biochemically important residues. Because of this and because α-CHT possesses maximal activity around pH 8, we decided to study the structure of the enzyme around pH 8.5 using X-ray crystallographic methods[172]. To extend the picture of the behavior of crystalline α-CHT with hydrogen ion concentration, the structure of the enzyme was also studied at lower pH (pH∿2.7).

VIII. EXPERIMENTAL

1. Crystal Preparation

Three times crystallized, salt free, lyophilized α-CHT was obtained from Worthington Biochemical Corporation and used in all studies without further purification or characterization.

Crystals were obtained using the method described by Northrup, et. al.,[173]. α-CHT was dissolved in distilled water adjusted to pH∿3.5 with H_2SO_4 at an approximate concentration of 22 mg of protein per ml of solution. A 75% saturated $(NH_4)_2SO_4$ solution was added to the protein solution until incipient turbidity, then a small amount of water adjusted to pH∿3.5 with H_2SO_4 was added to just clear the turbidity. The final solution contains about 5 mg of protein per ml of solution, has a pH of ∿4.2 and is about 54% saturated in $(NH_4)_2SO_4$. Large, diamond shaped plates up to 1.5×0.70×.05 mm grow from such a solution in three to four weeks at room temperature. The solution from which the crystals are grown was replaced by a 75% $(NH_4)_2SO_4$ solution and the tubes of crystals were stored in a water bath maintained at 15°C. Crystals thus obtained contain approximately 40% mother liquor, and are twinned along the $\vec{c}^*$ axis[145,149,174]. Crystals stored in 75% $(NH_4)_2SO_4$ adjusted to pH 3.6 correspond to "native" α-CHT crystals.

The pH of the soaking solution above the crystals of α-CHT was changed very slowly, by adding aliquots of 0.05 to 0.1 ml (1 to 2 drops) of 75% saturated $(NH_4)_2SO_4$ solution adjusted to pH∿8.7 with concentrated

NH_4OH. The rate of change in pH was ∿0.1 pH unit per 20-24 hours. After a period of about one month, the initial soaking solution was substituted with a corresponding solution of 75% $(NH_4)_2SO_4$ at pH∿8.7. This rather slow process of changing the pH was dictated by the fact that native crystals of α-CHT are extremely sensitive to proton concentration: a sudden change in pH from 3.6 to 5.0 or higher causes severe cracking of crystals, rendering them unsuitable for X-ray crystallographic work. When the pH was changed more slowly, the crystals started to crack around pH∿5-6 but not as severely as before. At pH∿8.7, most of the crystals developed a set of fine crack lines perpendicular to the $\vec{b}$ axis of the crystals. It seems that a stress develops along the $\vec{b}$ axis as the pH is increased, and the resultant strain is manifested by the cracks almost perpendicular to this axis. Increasing the pH further to about 9.0 cracks the larger crystals into finely divided pieces.

The pH of the soaking solution was measured at room temperature to ±0.1 of a pH unit with a Leeds and Northrup 7411 temperature compensated pH meter, equipped with a 117202 Combined Ag/AgCl Electrode. The pH meter was calibrated with Mallinckrodt standard Buffer solutions. Due to the very high ionic strength of the soaking $(NH_4)_2SO_4$ solution (∿4M), the liquid junction potential is considerably different from that of the more dilute standard buffer solutions and the pH values of the former will be affected to this extent. However, the relative pH values of the various soaking solutions remain indicative and since the 75% ammonium sulfate is not expected to change the liquid junction potential by more than 30 millivolts, the true pH values of these solutions are probably in error by less than 0.5 of a unit[175].

Using similar techniques, but 8M H_2SO_4 instead of concentrated NH_4OH, the pH of native α-CHT crystals was lowered from 3.5 to 2.7. The crystals remained essentially uncracked around 2.7, but attempts to lower the pH further, resulted in the development of severe cracks parallel to the $\vec{b}$ axis, rendering the crystals unsuitable for X-ray work.

Although the high and low pH measured are 8.7 and 2.7 respectively, in what follows, we will refer to these as "8.3" and "2.5". The latter pH values correspond to the midpoint for the pH range for which the structure of the enzyme apparently remains stable[172]. The 8.3 value is the midpoint of a 8.0-8.7 pH range, and 2.5 is the midpoint of a 2-3 pH range; the same holds for the 5.4 value which is the midpoint of a 5-6 pH range. A more complete discussion concerning small conformational changes of α-CHT as a function of the pH ranges has recently been given by Mavridis, et. al.,[172].

The 75% saturated $(NH_4)_2SO_4$ solution exhibits good buffering properties with respect to NH_4OH from pH 8 to 9 but poorer buffering properties in the 2.5 region. Solutions at both pH values standing over crystals for long periods of time (months) did not show any appreciable change in pH.

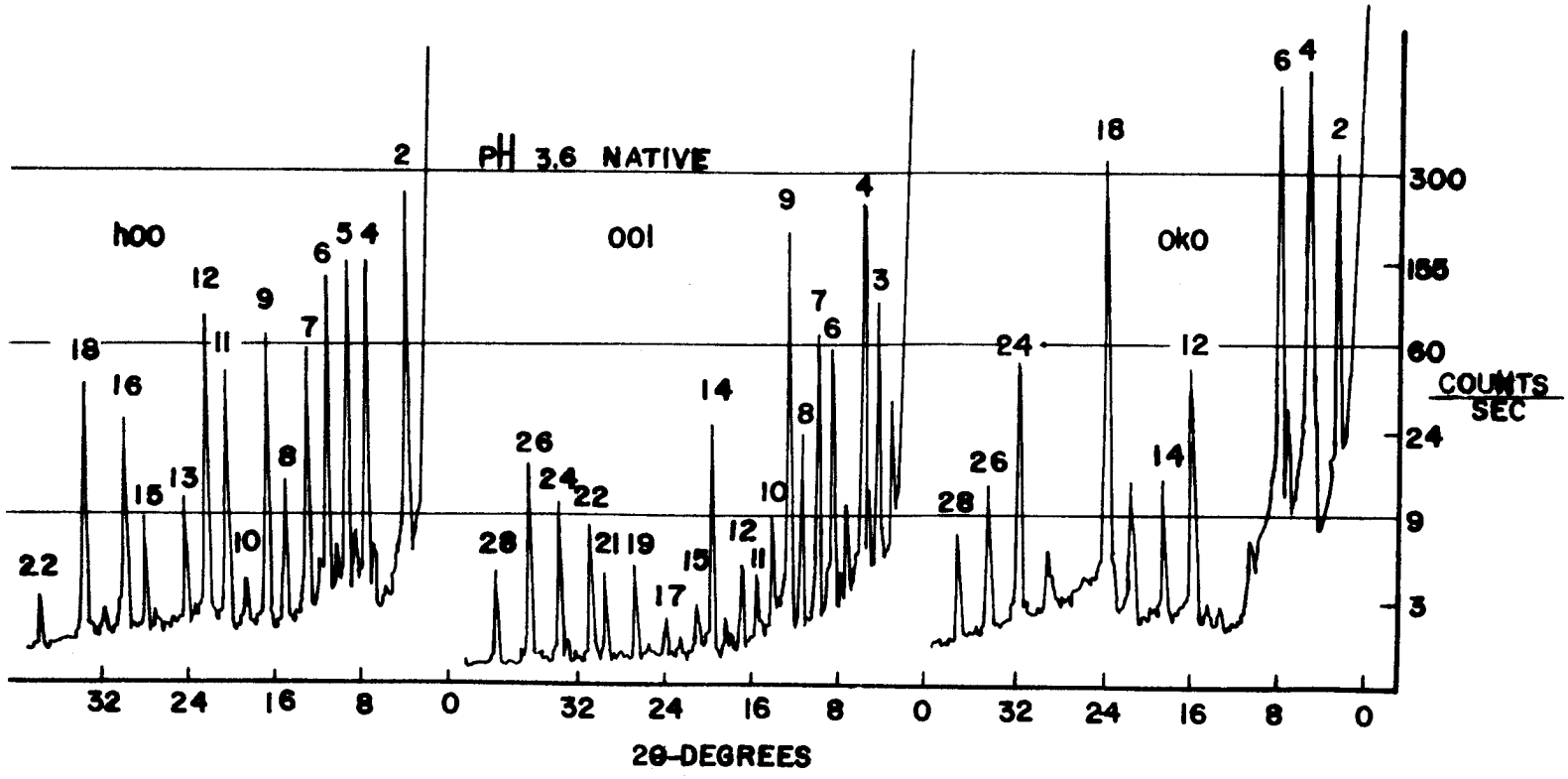
The pH 8.3 crystals used for intensity data collection were cracked, although less severely damaged crystals were always selected for the X-ray work. Comparison of precession photographs ((hk0), (0kl) zones) of slightly cracked crystals at pH 8.3 and uncracked native crystals showed only that the pattern of the crystals at higher pH decreases faster with scattering angles than that of the latter. In addition, the ω-spreads of selected reflections of native crystals, measured from

background to background, are generally 0.2-0.3°, while those of pH changed crystals appear to be slightly larger (0.3-0.4°). Thus no particular problems were experienced in collecting reliable intensity data from the slightly cracked crystals.

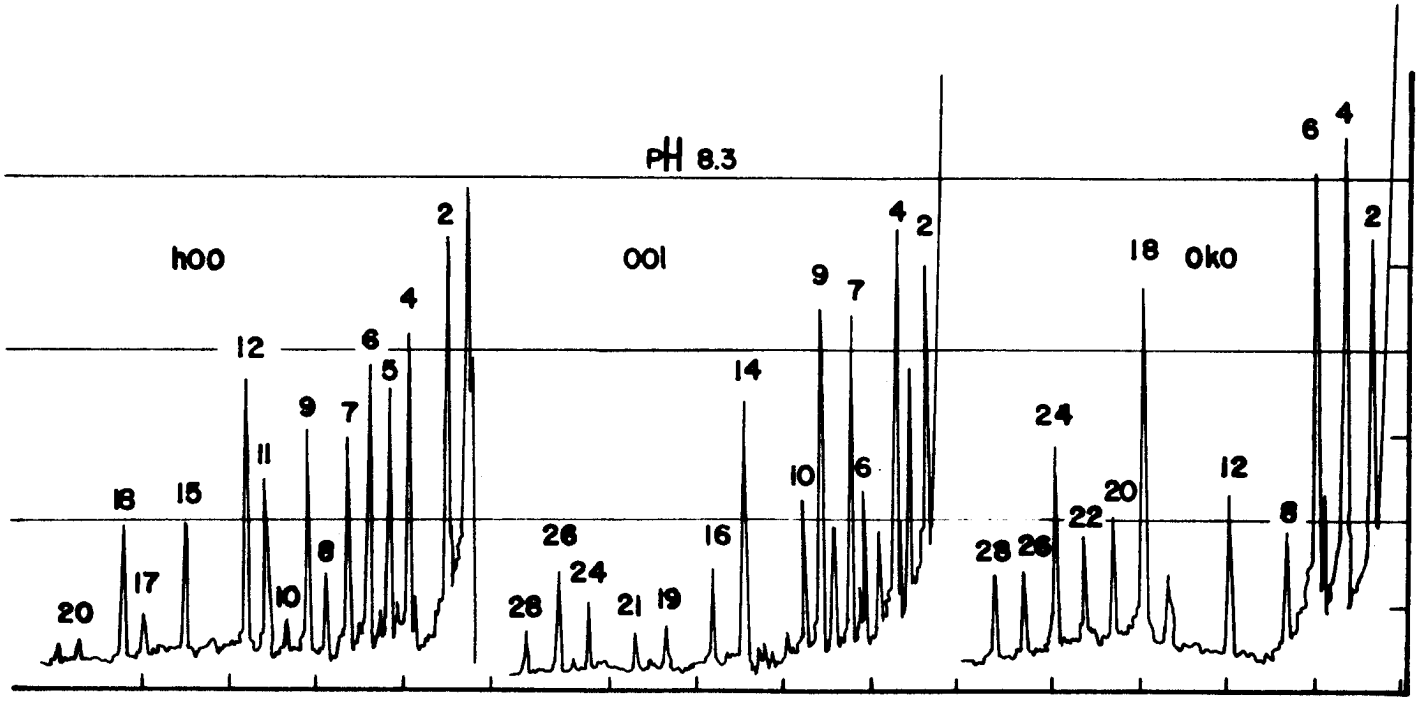
As the pH of the soaking solution was gradually changed, the diffraction patterns along the principal reciprocal axes $\vec{a}^*$, $\vec{b}^*$ and $\vec{c}^*$ were recorded and compared with those of the native enzyme. Around pH 8.0, some dramatic changes in the pattern were observed which remain essentially unaltered up to approximately pH 9.0, where the crystals start to crack so severely that they disintegrate. Figure 26 shows the diffraction patterns along three axial directions of native α-CHT crystals, α-CHT crystals at pH 8.3, and pH 2.5. The diffracted intensities are plotted on a logarithmic scale as a function of the scattering angle 2Θ. The heights are different because the three crystals were of different size; the relative changes in the intensities are the ones of importance. Changes in the low order reflections (1st and 2nd orders) are not indicative of structural transformations and they reflect changes due to variations in the electron density of the mother liquor. The largest number of significant changes are those occurring in the diffraction pattern of the $\vec{c}^*$ axis (Figure 26 (a,b)). The reflections (00,15), (00,17) and (00,22) are missing at pH 8.3, while the (00,6), (00,7), (00,8) and (00,16) reflections are undergoing severe intensity changes. The $\vec{b}^*$ diffraction pattern is the least sensitive with respect to the pH change.

The intensity changes of the low pH crystals are minor compared with those of the high pH, and again the most significant changes occur in the $\vec{c}^*$ diffraction pattern, while the changes along $\vec{b}^*$ and $\vec{c}^*$ are less so.

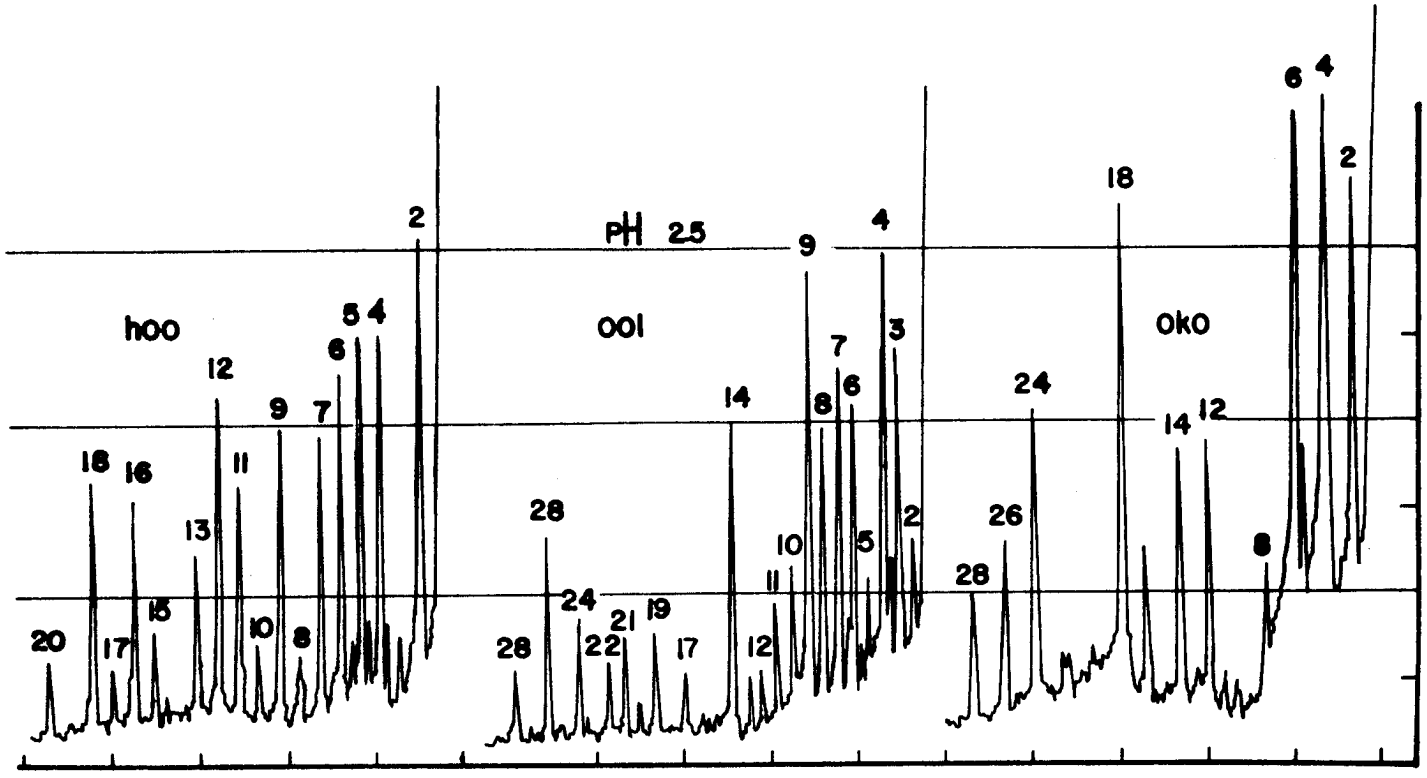
Figure 26. Axial scans along $\vec{a}^*$, $\vec{b}^*$ and $\vec{c}^*$ directions for native (a), pH 8.3 (b) and pH 2.5 (c) α-CHT crystals. The log of the diffracted intensity is plotted as a function of 2θ. The order of diffraction is indicated above each peak.


pH 3.6 NATIVE
h00
001
0k0
COUNTS/SEC
300
155
60
24
9
3
2θ-DEGREES
32
24
16
8
0

(a)


pH 8.3
h00
001
0k0
20
18
17
15
12
11
10
9
8
7
6
5
4
2
28
26
24
21
19
16
14
10
9
7
6
4
2
28
26
24
22
20
18
12
8
6
4
2

(b)


pH 25
h00
001
0k0
20
18
17
16
15
13
12
11
10
9
8
7
6
5
4
2
28
24
22
21
19
17
14
12
11
10
9
8
7
6
5
4
3
2
28
26
24
18
14
12
8
6
4
2

(c)

2. Intensity Data Collection and Reduction.

Crystals of α-CHT at both pH values were mounted in glass capillaries[176] and the unique $\vec{b}$ axis was fixed parallel to the azimuthal Φ axis of a four circle diffractometer. The crystals are held to the wall of the capillary by the surface tension of a small drop of mother liquor. Although the crystals adhere to the capillary wall firmly, nevertheless, they were allowed to stabilize positionally for at least 15 hours prior to X-ray measurements.

There are approximately 10,500 symmetry independent reflections within one quarter of the limiting sphere defined to 2.8A resolution for α-CHT. Excluding unobserved reflections, and a few low order reflections with d-spacings greater than 15A, phases are known for 9,044 reflections[149], of which approximately 6,300 have a figure-of-merit[177] greater than 0.7. The intensities of these reflections were measured, processed and used to calculate the difference electron density Fourier maps.

The high and low pH α-CHT intensity data were collected with the same automatic diffractometer, counting technique and software as that described elsewhere[78,149].

Two sets of three dimensional intensity data at 2.8A resolution of two independent preparations of pH 8.3 crystals were collected and converted to difference maps. The second independent experiment was carried out because the fall off of the intensities of the original set with X-ray exposure was large (about 40% for the 80 hours of exposure necessary to collect the intensity data). The second set of intensity data was collected 5 months later employing a hardware renovated X-ray diffractometer which featured the repositioning of the balanced Ni/Co filter pair to between the X-ray source and the crystal, and the

intercession of a 60 cm He-tunnel between the crystal and the scintillation counter detector. The fall off of this data set was only about 15% for 80 hours of X-ray exposure. The two independent difference electron density maps proved to be identical in all principal features and practically so for most of the lesser ones. Before the three dimensional data collection, the diffraction pattern along the principal reciprocal axes was examined and compared to that of the native enzyme; these axial patterns were also repeated after the completion of the data collection and compared with the former in order to ascertain any changes in the diffraction pattern which might arise from radiation damage. In both experiments, at high and low pH, these patterns remained essentially unaltered. In addition, intensity data at 6A resolution ($2\theta<15^{\circ}$, 131 reflections) were collected for the (h0l) zone before and after the three dimensional data collection. The purpose of this was two fold: first, to provide a way of scaling the intensity data of the derivative with respect to the native α-CHT data and second, to check the reliability of the correction to be applied to the intensity data to correct for radiation damage.

The following corrections were applied to the measured intensities:

(a) Lack of balance

An empirical correction measured on each crystal for the non-balance of the Ni/Co system of filters[149].

(b) Twinning

Crystals of α-CHT are twinned along the $\vec{c}^*$ direction and therefore the (0kl) reflections are comprised of the crystal (larger) and its twin (smaller). In order to correct for this effect the twin ratio R_t

was calculated by measuring the intensity of the (600) reflection for the crystal and its twin,

$$R_t = \frac{I_{crystal}}{I_{twin}}, \qquad (100)$$

with

$$I_{obs} = I_{crystal} + I_{twin}, \qquad (101)$$

where I_{obs} is the intensity of (0kl) reflections.
Combining (100) and (101), we obtain

$$I_{crystal}(0kl) = \frac{R_t}{R_t+1} I_{obs}(0kl). \qquad (102)$$

Another class of reflections systematically affected by twinning is that of (±2 kl) for l>9-10, since reflections 0.10-0.15° apart in Φ are unresolved under our data collecting instrumental conditions[149]. However, since there is no simple way to correct these reflections, this class was processed without applying any correction.

(c) Absorption

The method used has been described in the first part of this work[81]. Plots of I_{max}/I as a function of ϕ are shown in Figures 27 and 28 for the high and low pH crystals respectively. From Figures 27 and 28, it can be seen that the absorption curves for the different layers are fairly symmetrical but show substantial absorption. All four of the high pH and the three absorption curves at low pH were used to correct the absorption of the pH 8.3 and 2.5 crystals of α-CHT respectively.

(d) Decay

The intensities were corrected for radiation damage, which is a major source of error in protein crystallography. The effect is quite

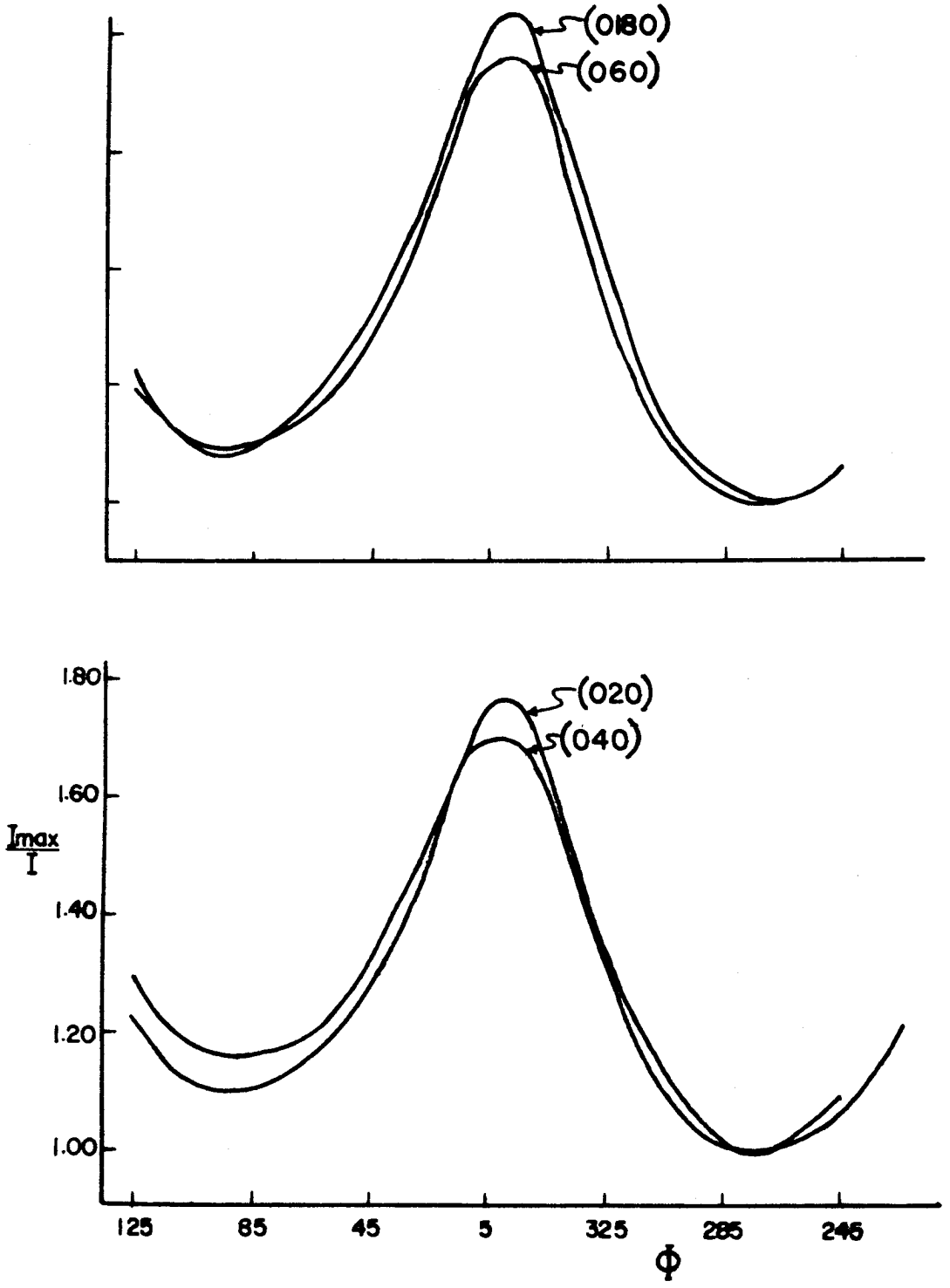


Figure 27. Absorption curves for the high pH crystals of α-CHT as a function of ϕ.

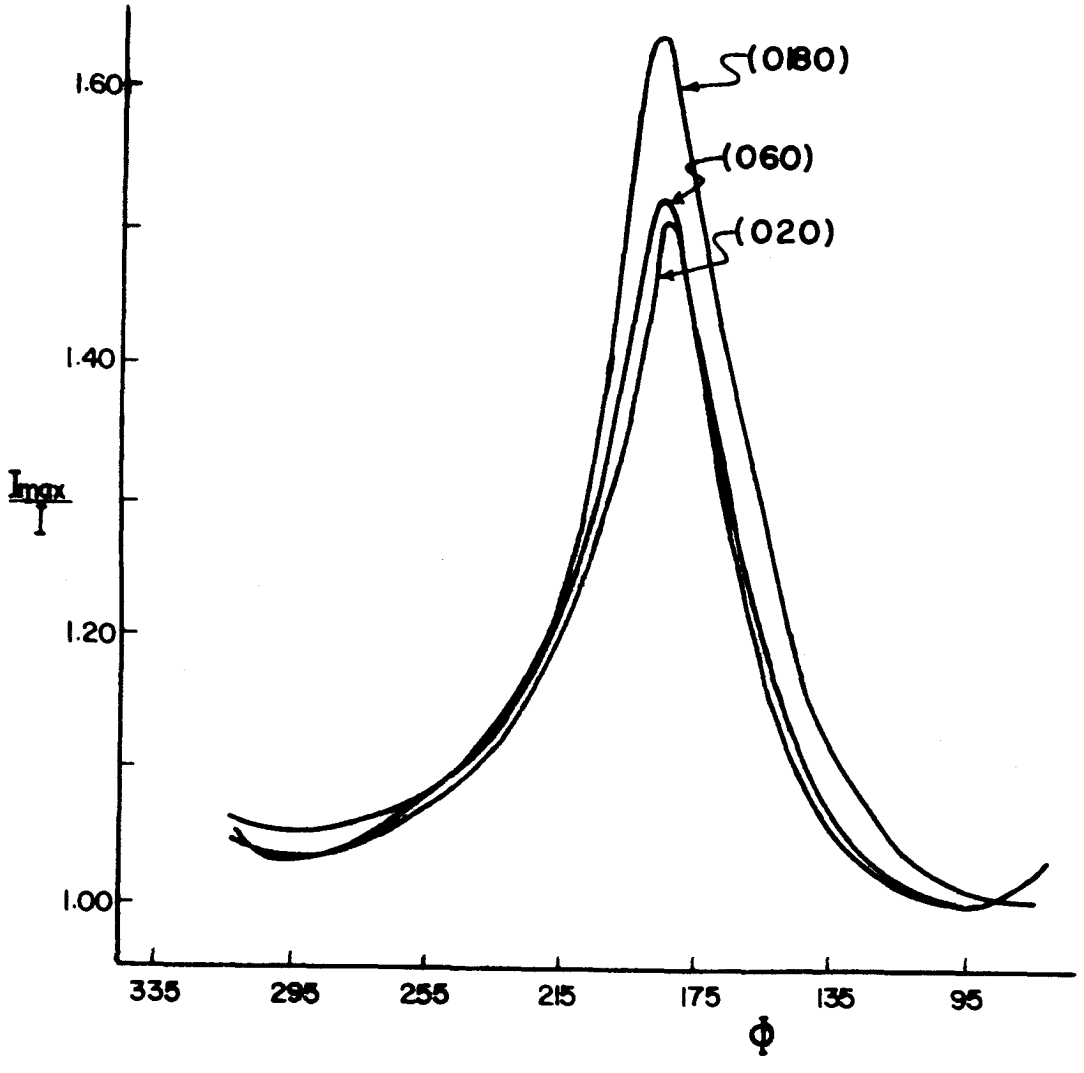


Figure 28. Absorption curves for the low pH crystals of α-CHT as a function of ϕ.

complicated and not well understood[178]. The decay was monitored during the data collection by measuring periodically the intensities of monitored reflections which had a 2Θ value of about 25°. These intensities decrease with X-ray exposure, (for a typical decay curve, see reference 78). The rate of decrease in intensity of the monitored reflections determined the decay slope which was used to correct the intensity of reflections as follows: Decay=N×slope+intercept, where N is the serial number of the reflection during the data collection. The decay slope of reflections with $2\Theta<15^{\circ}$ is taken to be half of those with $2\Theta\sim25^{\circ}$[149].

Finally the intensities were corrected for Lorentz and polarization factors and thus converted to structure amplitudes.

3. Data Processing

Scale factors for the high and low pH crystals were obtained by comparison of the peak heights of electron density peaks of the (h0l) projection (6A resolution) with those of the native enzyme. The average of the ratios of the peak heights of the native enzyme to the peak heights of high and low pH crystals were taken as the scale factors. These ratios are tabulated in Table XIV, for the high pH crystals, peaks 2, 5, 9 and 10 were excluded from the estimation of the scale factor, and peak 11 was excluded for the low pH crystals. Except for the peaks that were excluded (apparently these were influenced by the pH change) the ratios are fairly constant, particularly in the low pH case, once again indicating that the structural changes which have occurred in lowering the pH from 3.6 to 2.7 are not large. The scale factors for high and low pH crystals were 1.50(σ=±0.07) and 1.95(σ=±0.06), respectively.

TABLE XIV. Scaling of High and Low pH Crystals of α-CHT With Respect to Native (h0l) Projection.

Peak Number	Native/pH 8.3	Native/pH 2.5
1	1.59	1.92
2	1.13←	1.92
3	1.53	1.94
4	1.46	2.00
5	1.30←	1.88
6	1.50	1.97
7	1.40	1.97
8	1.42	2.04
9	1.33←	1.99
10	1.31←	2.00
11	1.63	1.67←
12	1.49	2.00
13	1.48	1.81
14	1.49	2.06
15	1.55	1.94

To verify the scaling scheme and other corrections that were applied to the intensities, the scaled $|F|^2$s of the high and low pH crystals were used to plot radial intensity distributions, $<|F|^2>$ vs. $<2\Theta>$ and these were compared with the corresponding distribution of the native enzyme. The distribution of $<|F|^2>$ over the 2Θ range of the data should not change appreciably with the pH induced changes. Certain changes can be tolerated and corrected with a function of the form

$$\mathcal{S}(\Theta)=S\exp\frac{B\sin^2\Theta}{\lambda^2} \qquad (103)$$

applied to the amplitudes $|F|$, where S is a scale factor and the exponential is a "temperature factor".

The constants S and B that were applied to the amplitudes $|F|$ for high and low pH crystals are given below

pH	S	$B(A^2)$
8.3	0.950	6.00
2.5	0.865	3.50 .

The corrected radial distribution curves for high and low pH crystals of α-CHT, along with the native radial distribution are shown in Figure 29.

The final sets of structure amplitudes, which were used in the Fourier difference electron density syntheses, contained 6185 and 6186 independent reflections for the high and low pH crystals, respectively.

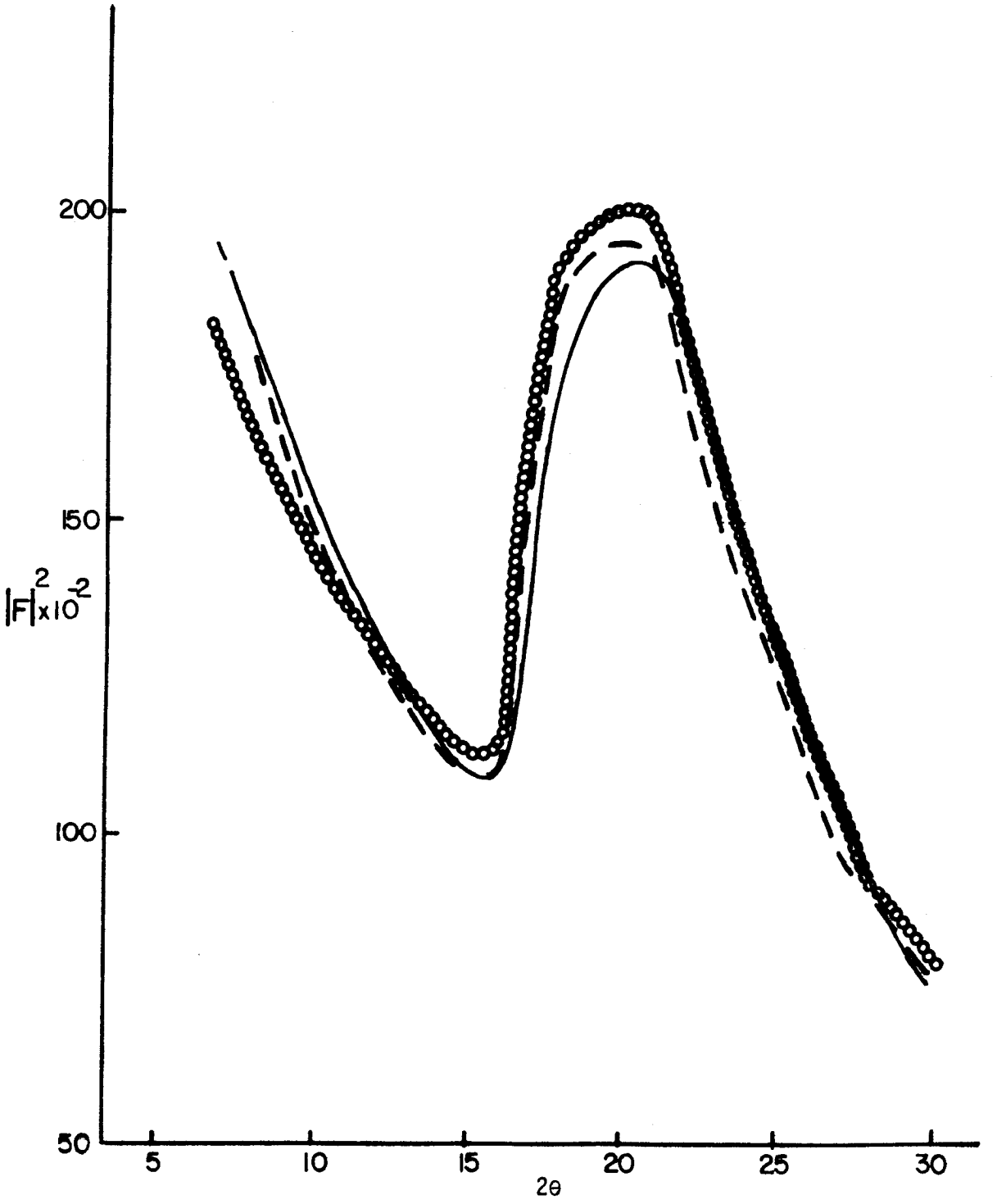


Figure 29. Radial distribution plots of native (——), pH 8.3 (---) and pH 2.5 (oooo) crystals of α-CHT.

IX. DIFFERENCE FOURIER METHOD

The difference Fourier method has proved to be a very valuable tool in protein crystallography. The power of the method lies in the fact that the known phases of a protein structure can be used to determine the structure of a protein derivative, thus circumventing the difficult problem of determining the phases of the latter. The assumption is made that the protein phases are the same as those of the protein derivative and the differences between the two structures can be obtained simply from the experimentally measured structure amplitudes of the derivative. This is a paradoxical situation because the phases depend on the structure amplitudes, and differences in structure amplitudes imply different phase angles. However, the method works well in practice because protein phases do not change much and because the resolution of proteins seldom exceeds 2.0A (in our case 2.8A). For the results to be reliable, it is essential that the average differences in the structure amplitudes be small, and the unit cell parameters of the derivative be similar to those of native crystals[179,180]. Changes in the dimensions of the unit cell can be an index of the quality and reliability of the difference maps.

Designating $|F_P(hkl)|\exp i\alpha^b(hkl)$ the structure factor of the native protein, where $\alpha^b(hkl)$ is the "best" phase[177] and $|F_D(hkl)|$ the structure amplitude of the derivative, a "best" Fourier difference map is computed

with coefficients

$$\{|F_D(hkl)|-|F_P(hkl)|\}\exp i\alpha^b(hkl)=\Delta F(hkl)\exp i\alpha^b(hkl) \ . \qquad (104)$$

The difference electron density will be given by

$$\Delta\rho=\frac{1}{V_P}\sum_h\sum_k\sum_l m(hkl)\Delta F(hkl)\exp\{2\pi i(hx+ky+lz)+i\alpha^b(hkl)\} \ , \qquad (105)$$

where m(hkl) is a statistical weighting factor representing the reliability in the phase of the (hkl) reflection (referred as figure-of-merit), V_P is the volume of the unit cell of the protein and x, y, z are fractional position coordinates in the unit cell.

The mean-square error in the difference electron density is given by the following expression according to Henderson and Moffat[181]

$$<(\Delta\rho)^2>=\frac{1}{V_P}\sum_h\sum_k\sum_l\{(\Delta F(hkl))^2+(\Delta F(hkl))^2(1-m^2(hkl))+\delta^2(hkl)\}, \qquad (106)$$

where the first term $\{\Delta F(hkl)\}^2$ is due to errors caused by the approximation $m\Delta F\exp i\alpha^b$ instead of the true but <u>unobservable</u> structure factors (F_D-F_P); the contribution of the second term is due to errors in the phases of $\alpha^b(hkl)$ of the parent structure, and the third term is due to experimental errors in the $\Delta F(hkl)$ values. Equation (106) does not take into account errors due to differences in lattice parameters.

The root mean square (r.m.s.) errors of the difference maps at high and low pH α-CHT crystals using equation (106) but not taking into account δ are 0.022 and 0.015 eA^{-3}, respectively. Differences smaller than seven times the r.m.s. errors (0.15 and 0.11 eA^{-3}) were not considered in the interpretation of these maps.

Table XV summarizes the unit cell parameters of α-CHT at four different pH values, 2.5, 3.6 (native), 5.4 and 8.3; the cell parameters of pH 8.3 are averaged values of two crystals. The length of the $\vec{b}$ axis, as can be seen from Table XV, changes by about 0.7A (1.4%) from the native crystals to pH 8.3 crystals. However, most of this elongation (∿0.6A) had already taken place at pH 5.4, and since the difference map at this hydrogen ion concentration had interpretable changes[153,154], it is safe to assume that the structural changes observed in the difference electron density map of α-CHT at pH 8.3 are also real and not the result of serious deviations from isomorphism.

The 2.8A resolution electron density map of native α-CHT has been synthesized using 8500 coefficients of the type

$$m(hkl)|F_p(hkl)|\exp i\alpha^b(hkl) ,$$

where the figure-of-merit, m, was greater than 0.30 [149]. With the aid of a "Richard's optical box"[182], a Kendrew model of the native α-CHT was built and fitted to the electron density[153].

The difference electron density maps were computed with coefficients

$$m(hkl)|\Delta F(hkl)|\exp i\alpha^b(hkl) ,$$

where m(hkl) was greater than 0.7, contoured and traced onto transparent cellulose acetate sheets, which in turn were superimposed on the plexiglass sheets displaying the native electron density map. The difference density contours were drawn beginning at 0.15 and 0.10 eA^{-3} for the high and low pH difference maps respectively, in intervals of 0.05 eA^{-3}. The structural differences caused by the change of hydrogen ion concentration were then interpreted with the aid of the model of native α-CHT and the optical box comparator.

TABLE XV. Lattice Parameters of α-CHT Crystals at Four Different pH Values.

pH	$\lvert\vec{a}\rvert$(A)	$\lvert\vec{b}\rvert$(A)	$\lvert\vec{c}\rvert$(A)	β(deg)
2.5	49.30	67.43	65.97	102.20
3.6(native)	49.24(7)	67.20(10)	65.94(6)	101.79(6)
5.4	49.13(5)	67.83(7)	65.81(7)	101.92(6)
8.3	49.30	67.87	65.84	102.00

X. RESULTS AND DISCUSSION

General Observations

The appearance of the high pH difference electron density map (in contrast to the low pH map) is very complicated. Series of negative troughs, intertwined with positive peaks create the impression of a futuristic tableau. Notwithstanding such complexity, a closer examination reveals the most salient structural changes that have occurred in the molecule with the pH change. Due to the nature of changes and the marginal resolution of the native electron density map, detailed motions of components of individual groups of the enzyme cannot be given and no attempt was made to "calculate" the magnitude of such movements; the latter was obtained whenever it was possible by simply trying to "fit" the native model with respect to the difference density taking into consideration the magnitude of the difference peaks.

In the pH 5.4 work[154] most of the changes were concentrated on the exterior of the molecule, especially around residues whose side chains were polar. The same type of changes, but more pronounced, were observed in the high pH work, but in addition some other rather dramatic changes occurred in the interior of the dimeric enzyme, which relate to its functionality and activity[172]. A characteristic of the high pH difference map is a general lack of local two fold symmetry. This phenomenon is related to the fact that the two molecules of α-CHT are crystallographically independent and therefore not exactly alike[150].

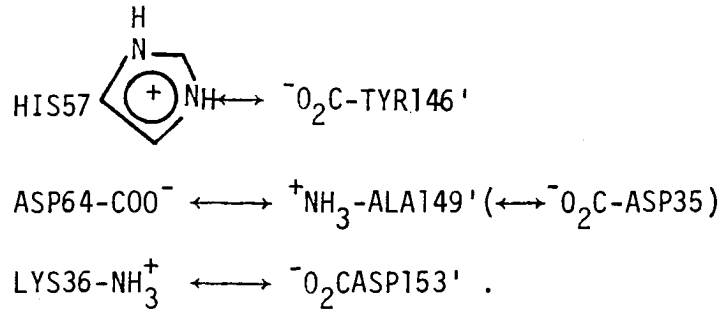
The low pH difference map does not contain many or very large changes, which was expected from a number of other observations previously mentioned.

A. High pH (8.3) Structural Changes of Crystalline α-CHT.

1. Dimeric Interactions

Under acidic conditions two molecules of α-CHT interact with each other in a geometrically complicated way to form a dimer, simply and schematically outlined in Figure 19. Essentially two types of forces, coupled with the gross complementarity of the monomeric molecules, are responsible for the phenomenon: electrostatic interactions, which are pH dependent and whose importance varies with the local dielectric constant, and weak van der Waals interactions which can be considered as a first approximation pH independent.

The dimeric interfacial electrostatic interactions between the two molecules are as follows[154]



Solution studies based on sedimentation equilibrium methods by Aune and Timasheff[166] show a pH dependence of the dimerization of α-CHT and they suggest this is due mainly to the ionization behavior of the imidazolium of HIS57 $\longleftrightarrow$ carboxylate of TYR146' electrostatic interaction. The suggestion is consistent with the crystal structure of the molecule

because HIS57 is near TYR146'. That TYR146 is important in the dimerization process, is shown by the inability of the molecule to associate if it is removed from the enzyme[183] or if it participates in a peptide linkage as in δ-chymotrypsin[184].

As the acidity of the solution is decreased the imidazolium ring of HIS57 deprotonates and the Coulombic interaction with the already ionized carboxylate group of TYR146' is disrupted; the result of this is a movement of the carboxylate terminal of TYR146'. The exact pK of HIS57 in α-CHT is somewhat uncertain; Cruickshank and Kaplan[164] give a value of 6.8, but more recent results obtained with C^{13} Fourier Transform NMR techniques[185] suggest a pK_{α} value much less than 6.8 and closer to 5. However, these results were not obtained for HIS57 of α-CHT, but for the single HIS residue of α-lytic protease, whose environment is supposedly similar to that of α-CHT.

The difference electron density Fourier maps of α-CHT at various pH values show negative density located on the native density of the terminal carboxylate of TYR146' and TYR146[154,186], indicating movement of the carboxylate group away from HIS57 as the latter deprotonates. Table XVI gives the peak heights of the negative difference densities of the TYR146' carboxylate at different hydrogen ion concentrations. These values clearly indicate a disturbance of the carboxylate ion just above pH 5. The peak height increases with a gradual deprotonation of the imidazolium ion until it probably reaches a maximum of $0.51eA^{-3}$ in the pH 8.3 range, where the imidazole ring is neutral. A pK value of approximately 6 is consistent with the values of Table XVI.

Two large positive difference peaks of about the same height ($\sim 0.30eA^{-3}$) are on either side of the $0.51eA^{-3}$ negative difference

TABLE XVI. Negative Difference Densities Associated With the Movement of the Terminal Carboxylate Group (TYR146') of α-CHT as a Function of pH.

pH	State of the Enzyme	Peak Height(eA^{-3})	Ref.
4.6	Oxidized α-CHT (with CCl_3-SO_2-Cl)	∿0.0	186
∿5.2	Oxidized α-CHT (with CCl_3-SO_2-Cl)	-0.24	186
5.4	Pure α-CHT	-0.20	154
7.3	Pure α-CHT	-0.33	186
8.3	Pure α-CHT	-0.51	(this work, 172)

density, making the interpretation of the movement of TYR146' ambiguous. Vandlen and Tulinsky[154] observed the same negative difference density at pH 5.4, although at about half the height (Table XVI), because the imidazole ring was not fully deprotonated. On the other hand they also observed only one positive difference peak so that interpretation of the movement of the carboxylate of TYR146' was straightforward. Thus the direction of movement of the carboxylate terminal of TYR146' is away from the imidazole ring of HIS57 and this is effected by a counter-clockwise rotation around the NH-C_α bond which corresponds to a translation of approximately 2A (when a rotation around a bond is discussed, the bond is taken as the rotation axis and the sense of the axis is taken from the first to the second atom of the bond). Figure 30 illustrates the motion, the rotation axis being essentially perpendicular to the local two fold axis. The phenol side chain of TYR146' remains virtually undisturbed as the carboxylate group rotates; a small negative density located on the phenyl group could indicate a slight rotation around C_β-C_γ bond coupled with the rotation of the carboxylate ion.

The difference Fourier map is capable of indicating if a small segment of the protein (like the carboxylate of TYR146') has moved. However, the question of "why it is moving" always suffers from the subjectiveness of the interpreter. In this case a possible explanation is that the carbonyl group of the peptide bond of HIS57, which makes a close approach (∿3A from the $-COO^-$ of TYR146') and since it carries a "partial negative charge", it might repel the negatively charged carboxyl group in the absence of a protonated imidazole of HIS57.

There is a two fold equivalent negative difference density which corresponds to a structural change of TYR146, but no corresponding

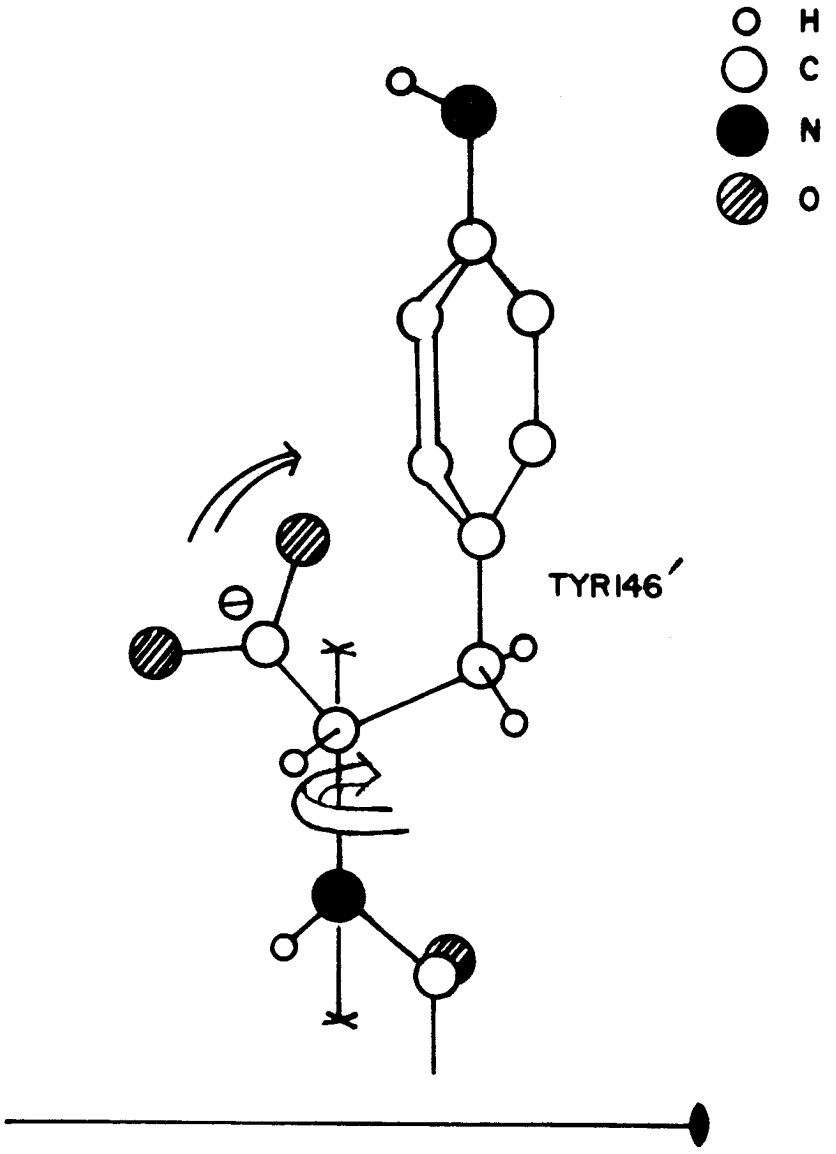


Figure 30. Rotation of terminal carboxylate group of TYR146' around the NH-C_α bond. The local two fold axis shown is perpendicular to the rotation axis.

positive difference peaks. This is probably because the movement of the carboxylate of TYR146 is further obscured by a large movement of the main chain from TYR146 to LEU143, which is not observed with corresponding TYR146' segment. The coordinates and peak heights of the negative densities of the carboxylate (at pH 8.3) are given below. Hereafter, coordinates of the main difference density peaks expressed in A units will be given, along with the peak heights in electrons A^{-3}, and the deviation from a two fold symmetry operation, $|\delta|$; a positive (+), or negative (-) sign will indicate if it is a positive or negative difference density respectively. The coordinates will be referred to a Cartesian system utilizing the position of the local two fold axis as one of the Cartesian axes and the $\vec{b}$ and $\vec{c}$ crystallographic directions as the other axes[151,187]; since the origin of coordinates will be on the local two fold axis (along y=z=0), the directions and deviations from local two fold symmetry can be obtained by inspection.

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
TYR146	16.0	8.2	5.0	-0.35
TYR146'	16.0	-8.0	-4.8	-0.51
$\|\delta\|$	0.0	0.2	0.2	0.16

From the foregoing it can be seen that the two fold axis is practically exact with respect to the coordinates but not to the peak heights; the complete lack of local two fold symmetry for the positive difference peaks of TYR146' and TYR146 is shown clearly in Figure 31 which also shows a general lack of local two fold related changes between the two molecules of α-CHT (I and I') with a change in pH to 8.7. In addition,

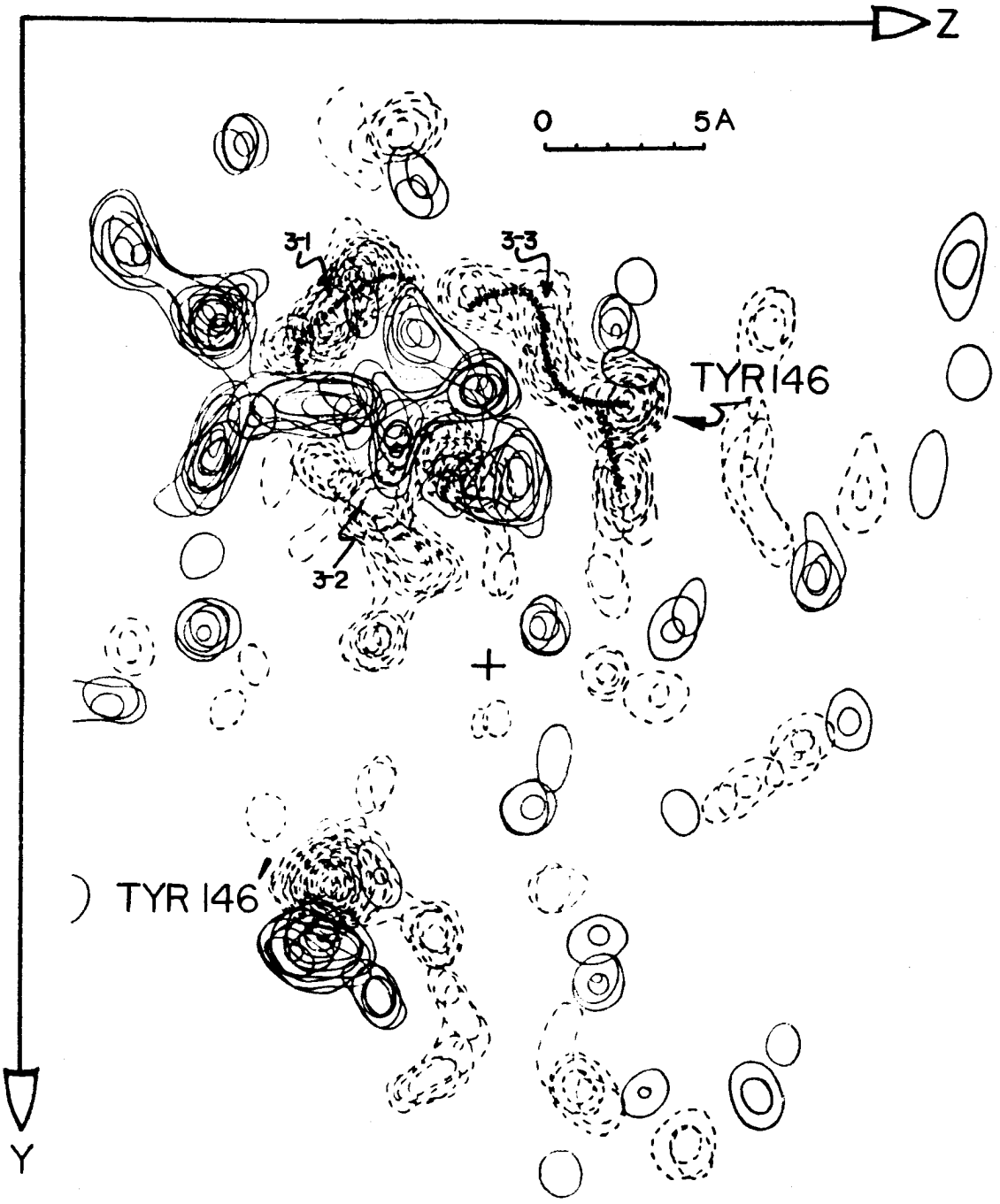


Figure 31. Difference electron density between pH 8.3 and 3.6 crystals of α-CHT in vicinity of active site viewed down the local two fold axis; contours at $0.05eA^{-3}$ beginning with $0.15eA^{-3}$; solid contours positive, broken negative; thickness shown ∿6A; 3-1, 3-2, 3-3 are section numbers where the difference densities are discussed; position of local two fold axis indicated by cross.

Figure 31 shows the restricted motion of the TYR146' chain as compared to that of the TYR146 segment, where the movement of the carboxylate group does not end abruptly in the peptide bond of TYR146-ARG145, but goes further through the chain of the tetrapeptide segment TYR146-ARG145-THR144-LEU143.

HIS57 does not show any difference electron density at pH 8.3, implying no positional changes; the position might be so critical for the action of the enzyme, that little flexibility exists in this region of the molecule.

A negative difference density near ASP64 and ALA149' and the local two fold equivalent have the following coordinates:

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
ALA149	2.8	-10.8	-4.2	-0.27
ALA149'	2.9	11.2	4.4	-0.27
$\|\delta\|$	0.1	0.4	0.2	0.00

Tulinsky and Wright[187] in a SO_4^{2-}-SeO_4^{2-} exchange experiment, determined the positions of localized sulfate ions in crystalline α-CHT at pH 3.6. They found localized sulfate ions close to ALA149 and ALA149'; the coordinates and electron density peak heights of these are given below:

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
SO_4^{2-}149	3.3	-12.8	-4.2	+0.71
SO_4^{2-}149'	2.4	15.4	5.0	+0.73
$\|\delta\|$	0.9	3.4	0.8	+0.02

For a fairly mobile and chemically unbound ion the two sets of foregoing coordinates can be considered to be related. The discrepancy in the peak heights, 0.27 vs. 0.7 eA^{-3} is due to the phase angle assumption employed in the synthesis of the difference Fourier map, as discussed by Luzzati[188], and due to the fact that only reflections with figure-of-merit m>0.7 were used in the difference calculation while m>0.3 was employed to obtain the electron density. Thus the sulfates located near ALA149 and ALA149' appear to move from their previous positions upon changing the pH from 3.6 to 8.7; this is probably related to the deprotonation of the quaternary nitrogen of N-terminal of ALA149.

2. Histidine 40

There are only two histidine residues in α-CHT; HIS57 is in the active site and its presence is required for enzymatic activity, and HIS40 is located close to the surface of the molecule and apparently does not participate in any enzymatic function, but presumably is associated with the quenching of the fluorescence of chymotrypsin[143,189].

According to model studies by Katchalski and co-workers[190], fluorescence quenching of tryptophan occurs in conjunction with histidine type of compounds at low pH. The explanation is that of the formation of a "charge transfer complex" involving the imidazolium ion (electron deficient) with the electron rich tryptophan. As the pH increases, the imidazole ring deprotonates, and the complex is disrupted leading to an increase in fluorescence. Hess and co-workers[143] investigated the role of pH in the quenching of fluorescence of δ-chymotrypsin which is essentially structurally identical with α-CHT. They found that the fluorescence decreases with decreasing pH as a group of the enzyme with

pK∿6 ionizes; their results are summarized in Figure 32. Blocking imidazolium ion formation of HIS57 with the specific reagent TPCK(L-1-tosylamido-2-phenylethylchloromethyl ketone) the change in fluorescence in the 4-7 pH region remains constant. Based on Katchalsky's work, these authors suggested that HIS40 could be responsible for the quenching of fluorescence; at low pH HIS can form an intramolecular charge transfer complex with a tryptophan which then separates as the pH is increased above 6. No explanations are offered why the fluorescence falls off after pH 7.6 and the fact that around pH 10 the absolute quantum yield is almost the same as at pH 4.

In native (crystalline) α-CHT (pH∿3.6), TRP141 is approximately 5-6A away from the ring of HIS40, a distance too large for the formation of a charge transfer complex. However, it is known that tryptophan fluorescence of proteins is extremely sensitive to the microenvironment of tryptophans[191] (for instance to the local dielectric constant), so that small conformational changes can conceivably create large variations in the quantum yields of fluorescence spectra.

The difference Fourier map at high pH shows a small movement of the imidazole ring of HIS40. A prevailing positive difference density is observed in the vicinity of HIS40 as well as HIS40'. The coordinates are given below:

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
HIS40	5.8	-0.9	-6.9	+0.28
HIS40'	5.2	1.7	6.7	+0.28
$\|\delta\|$	0.6	0.8	0.2	0.00

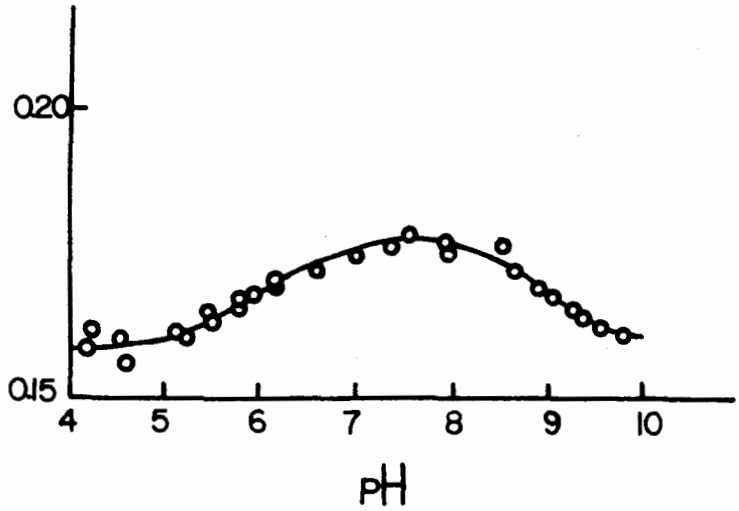


Figure 32. The quenching of the fluorescence of δ-chymotrypsin as a function of pH, measured at 340 nm and 25°C. (Reference 143).

A negative difference density ($-0.20eA^{-3}$) is located at the edge of the native density corresponding to the imidazole ring of HIS40 and its two fold equivalent. Therefore, a rotation of the imidazole ring around the C_α-C_β bond is suggested in such a way as to bring the deprotonated ring closer to the hydroxyl group of SER32, with the possible formation of a hydrogen bond with an unprotonated nitrogen of the imidazole ring and the hydroxyl group (-OH) of SER32. Figure 33 illustrates the above motion of the imidazole, the axis of rotation being perpendicular to the local two fold axis. A small negative difference density ($\sim -0.2eA^{-3}$) on SER32 shows a disturbance of this residue, probably caused by the approach of HIS40. The reason for the movement of HIS40 might be the disruption of the electrostatic interaction between the positively charged imidazole ring of HIS40, and the negatively charged carboxylate of ASP72, upon deprotonation of HIS40. The distance between the carboxylate of ASP72 and the imidazole ring of HIS40 is approximately 4A. Clearly the deprotonation and movement of HIS40 can change the microenvironment of TRP141 drastically and a resultant enhancement of fluorescence can be observed with the increase of pH.

TRP172 is well resolved in the native map of α-CHT; the difference electron density map shows a movement of this residue which might be related with the high pH quenching of chymotrypsin (Figure 32). The following coordinates correspond to negative difference densities on the upper and lower parts of the native density which has been assigned to the benzo ring of the indole of TRP172

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
TRP172	28.3	-4.1	-6.3	-0.27
	29.3	0.9	3.9	-0.18 .

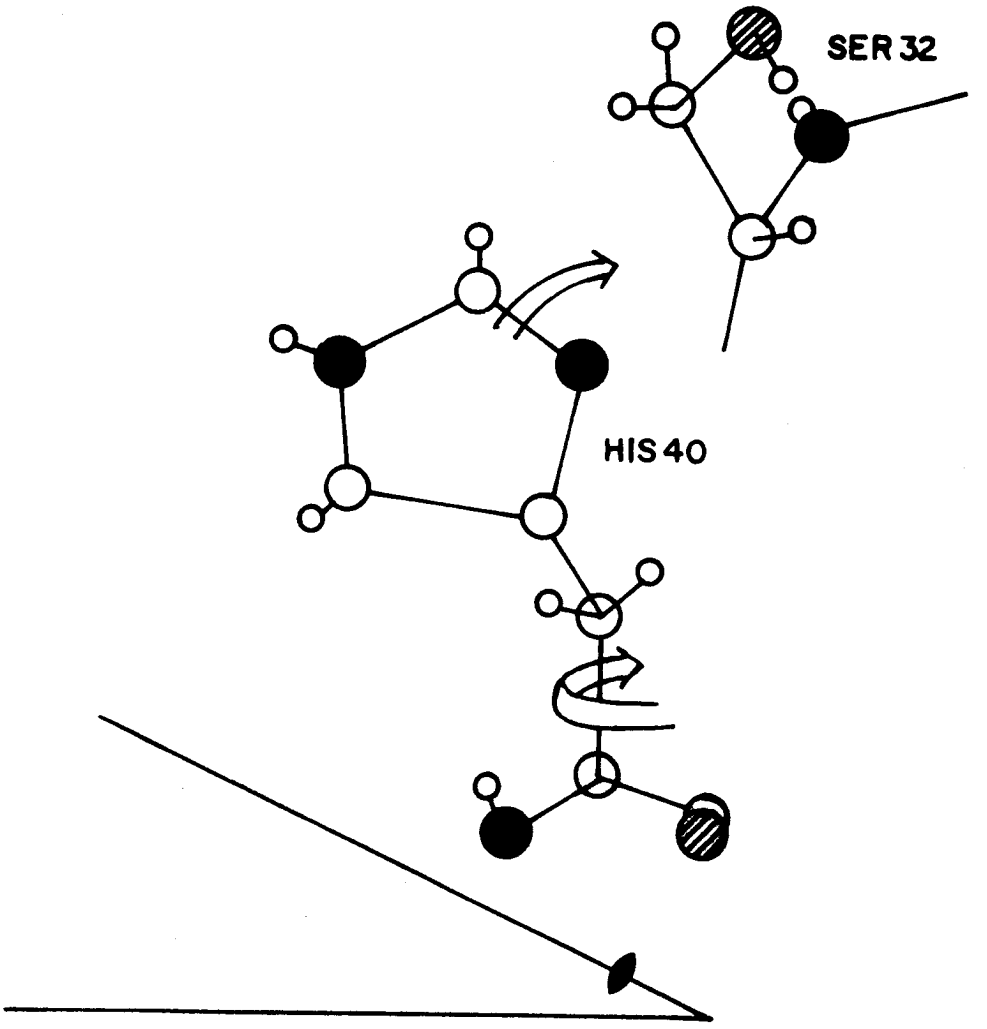


Figure 33. Movement of the imidazole ring of HIS40 by rotation around the C_α-C_β bond; the rotation axis is approximately perpendicular to the local two fold axis.

A positive difference density exists with the following coordinates:

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
TRP172	28.6	-1.5	-7.2	+0.20
	28.6	0.0	-5.7	+0.21

and there is no two fold equivalent for either. In the native enzyme, the plane of TRP172 is inclined approximately 45° with respect to the xz crystallographic plane and its distance from the disulfide bond of CYS168 is approximately 5A. The motion implied by the forementioned difference densities is such as to make this angle smaller and to make the TRP172 plane more horizontal to xz plane. A pair of negative and positive difference densities $0.25eA^{-3}$ and $0.16eA^{-3}$, respectively, are located on the very well defined native density of the -S-S- bond of CYS168-182 which indicate a slight translation of this group consistent with the approaching motion of TRP172. Such an interaction of TRP172 with the disulfide bond of CYS168-182 could be the cause of a radiationless transition instead of fluorescence. Due to the high conformational complexity of the enzyme, the cause of the movement of these groups is unclear, but it is conceivable that motions of different groups can occur which are not directly related to pH effects.

3. The Active Site Region

The basic "ingredients" of the active site are HIS57, SER195 and ASP102; if one of these residues is removed or chemically altered α-CHT is not operational[157-159,192-194]. No structural changes are observed in the above residues upon changing the pH from 3.6 to 8.7; however, large structural changes associated with other residues implicated indirectly

to the activity of the enzyme and located close to the active site are observed in the difference Fourier synthesis. We subdivide this section into three parts:

3-1, deals with the internal ion pair ILE16⟷ASP194.

3-2, concerns the chain between residues ASP194-CYS191-220.

3-3, deals with the region LEU143-TYR146.

Figure 31 shows a part of the spectacular and asymmetrical changes which occur with change in pH to 8.7. The numbers 3-1, 3-2 and 3-3 in Figure 31 indicate the negative difference densities associated with structural changes of the chains between ILE16-VAL17, ASP194-CYS191-220 and LEU143-TYR146, respectively. The positive difference densities accompanying the negative cannot be marked as clearly. In general, the positive difference densities are not as well defined, and many times difficulty is encountered in assigning the appropriate positive difference density which corresponds to a given negative counterpart.

3-1. Isoleucine16 Valine17 Dipeptide.

The importance of the ILE16⟷ASP194 ion pair has been mentioned repeatedly in this work; Figure 34 gives a schematic representation of the enzyme folding around this salt bridge in a view perpendicular to the local two fold axis. It is a well documented fact that the activity of the enzyme falls off rapidly after pH 8, and this has been attributed to the deprotonation of the terminal $-NH_3^+$ of ILE16 and the disruption of the Coulombic attractive interaction with the carboxylate of ASP194. Vandlen and Tulinsky[154] did not observe any changes in the ILE16 region of the enzyme at pH 5.4 because the deprotonation of the $-NH_3^+$ of ILE16 had not yet taken place at the lower pH (preliminary results at pH~7.3[186]

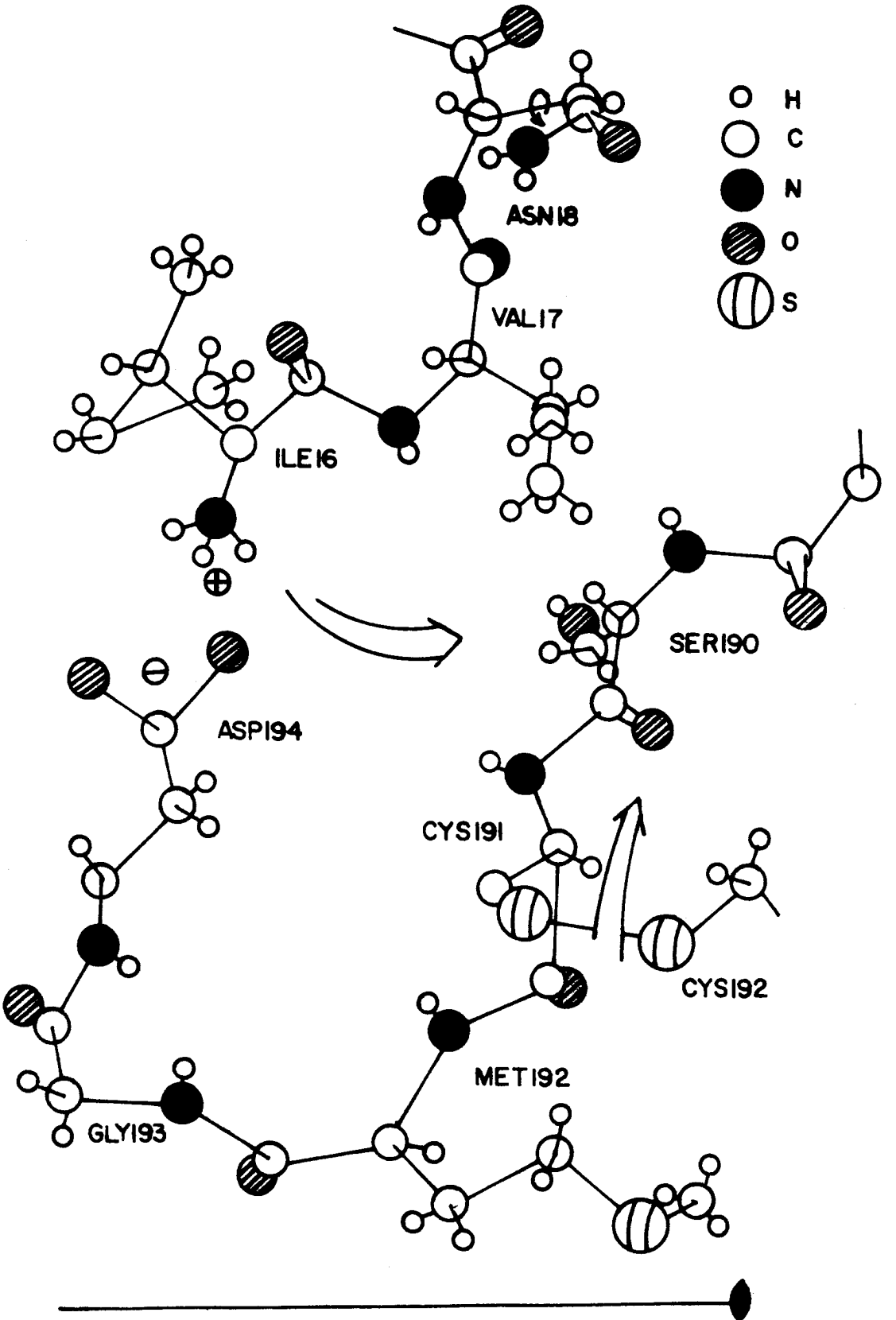


Figure 34. Schematic representation of α-CHT folding around the internal ion pair ILE16⟷ASP194; view perpendicular to the local two fold axis.

show that the structure of the enzyme also remains essentially unaltered in the ILE16 region at this pH, although there are small difference densities $\sim 0.15 eA^{-3}$ with coordinates corresponding with the coordinates of the 8.3 difference maps). We hoped to see structural modifications around ILE16 by raising the pH above 8.0 which could be related to the inactivation process of α-CHT. The fact that we have examined the phenomenon in the crystalline state should deter but slightly the significance of the correlation of the present results with those observed in solution studies.

The difference Fourier map revealed a massive negative density located on the native density, starting at the amino terminal of ILE16, and ending just before the amide bond of ASN18-GLY19; there are 23 atoms (hydrogen atoms are not included) embedded in this negative difference density. Table XVII gives the coordinates as a function of the x-coordinates of the negative density associated with the movement of the ILE16-VAL17 dipeptide and part of the ASN18 residue. The first four entries of Table XVII correspond to a simple rotation of the side chain of ASN18 around the C_α-C_β bond (see Figure 34).

There is a complete lack of local two fold symmetry in the changes in this region of the molecule (see Figure 31), which indicates that the ILE16' chain of the other molecule (I') is not influenced by the pH change. This lack of symmetry implies that the pK_α of the $-NH_3^+$ terminal of ILE16 is different in the two molecules probably due to the slightly different environments in the independent molecules[150]. The negative density in Table XVII has a thickness of ∿9.0A along the x-direction. A perspective view of this density is given in Figure 35, which shows schematically the manner in which this density moves along the x and y

directions (the z-coordinate, as can be seen from Table XVII, is essentially constant).

The question as to the direction in which this branch is moving is not easy to answer because the positive difference density is not as neatly distributed, but is rather scattered, in pieces, and is not a continuous counterpart of the negative density. Nevertheless, it is clear that the amino terminal of ILE16 approaches SER190 (see Figure 34); at the same time, the isobutyl side chain of ILE16 rotates around the C_α-C_β bond in a such a way as to remove itself from the path of the $-NH_2$ group. This rotation brings the isobutyl side chain close to the carbonyl of the peptide bond between residues TRP141-GLY142 and makes the approach of the amino terminal towards the -OH group of SER190 feasible. The coordinates of the positive difference density which correspond to the movement of the $-\underset{\underset{O}{\|}}{C}-CH-NH_2$ part of the ILE16 are:

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
14.3	-6.6	-9.2	+0.26
13.6	-7.6	-8.8	+0.28
13.0	-8.3	-8.7	+0.28
12.4	-8.3	-8.6	+0.25 .

Coordinates of positive difference densities which correspond to the peptide bond of ILE16-VAL17 are given below

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
16.2	-10.1	-9.0	+0.22
16.9	-10.4	-9.1	+0.32
17.5	-10.4	-9.2	+0.33
18.2	-10.7	-9.4	+0.28

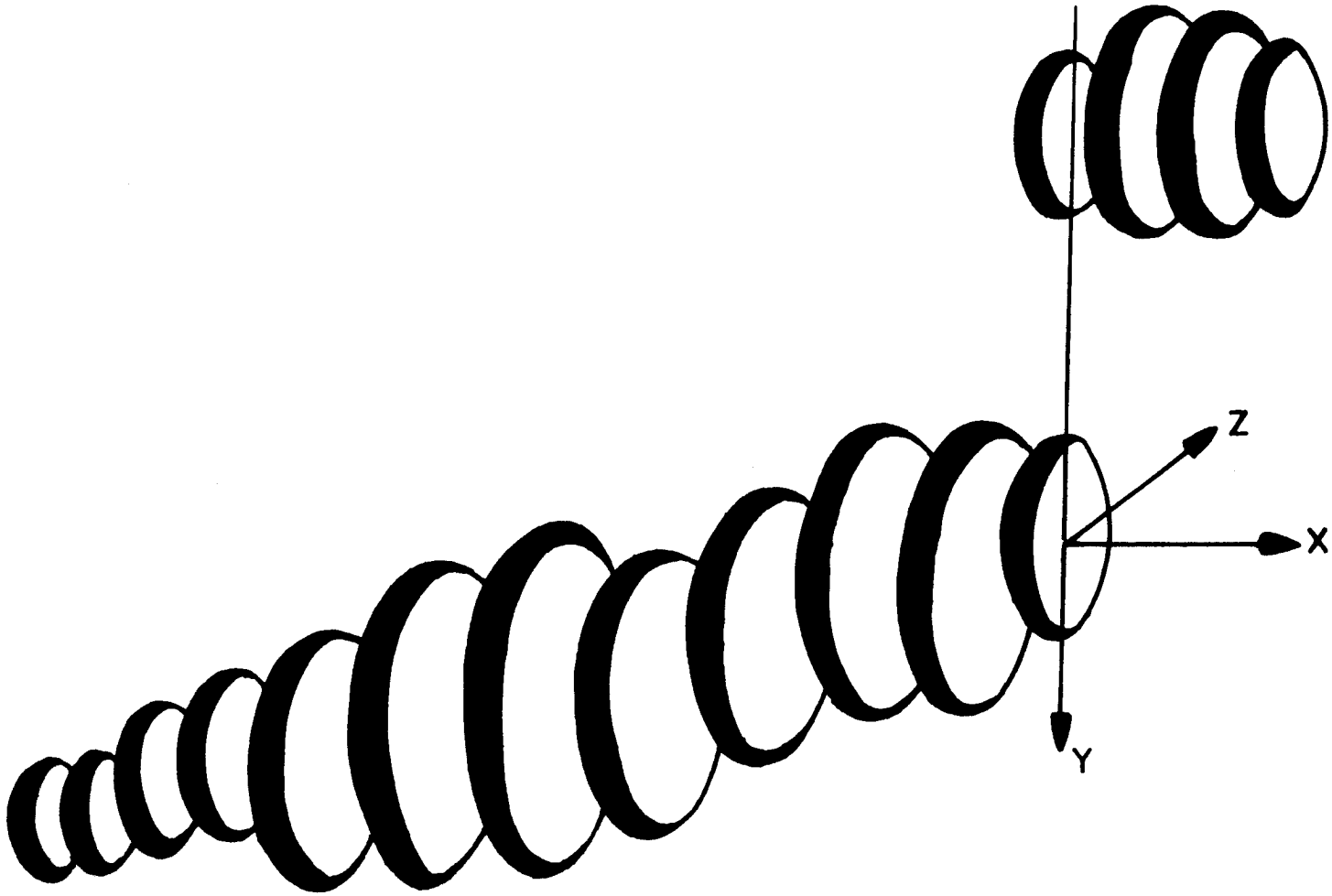


Figure 35. Negative difference electron density layers perpendicular to the plane of the paper (yz plane); they correspond to Table XVII. Their relative size is approximately proportional to peak heights.

TABLE XVII. Coordinates of Difference Density in ILE16 Region

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
20.1	-16.9	-3.4	-0.20
19.5	-16.9	-3.4	-0.27
18.9	-16.9	-3.4	-0.27
18.3	-16.9	-3.4	-0.16
18.3	-12.4	-4.3	-0.25
17.6	-12.1	-4.2	-0.34
17.0	-12.1	-4.4	-0.33
16.3	-11.5	-4.9	-0.30
15.6	-10.8	-5.4	-0.33
14.9	-10.8	-5.5	-0.37
14.2	-10.4	-5.6	-0.36
13.6	-10.2	-5.8	-0.31
12.9	-10.2	-5.7	-0.22
12.3	-9.7	-5.8	-0.20
11.6	-9.5	-5.9	-0.18
11.0	-9.5	-6.0	-0.18

In summary, it can be said that upon deprotonation (complete or partial) of the $-NH_3^+$ terminal of ILE16, the dipeptide ILE16-VAL17 moves toward SER190; in particular, under favorable stereochemical conditions, the $-NH_2$ terminal can form a hydrogen bond with the -OH group of SER190. There are many degrees of freedom in the tripeptide ILE16-VAL17-ASN18 so that almost any kind of motion can be assigned relatively easily to correspond to the difference map.

The motion of the ILE16 chain does not seem to influence the stereochemistry of the active center of the enzyme. However, the carboxylate group of ASP194 is apparently in an energetically unfavorable condition, due to the deprotonation, since it remains charged and buried in an environment of low dielectric constant; its movement might influence the activity of α-CHT.

3-2. Aspartic 194.

The disruption of the ion pair between ILE16 and ASP194 by the deprotonation of ILE16 triggers the movement of the dipeptide ILE16-VAL17 and also causes structural changes indirectly to the tetrapeptide ASP194-GLY193-MET192-CYS191-220. The Fourier difference density map shows an impressive negative density located exactly on a very well defined native density, beginning with the carboxylate of ASP194 and continuing along the main chain until CYS191-SER190. Table XVIII gives the coordinates and peak heights of this density which extends for about ∿10A and is associated to the movement of the forementioned tetrapeptide (folding shown schematically in Figure 34). There is also a complete lack of local two fold symmetry, molecule I' again being passive to the changes occurring in molecule I.

TABLE XVIII. Coordinates of Difference Peaks in ASP194-SER190 Region

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
19.5	-5.4	-0.7	-0.21
18.8	-5.7	-0.6	-0.29
18.2	-5.7	-0.7	-0.31
17.5	-6.3	-1.1	-0.30
16.9	-6.6	-1.5	-0.32
16.2	-7.0	-2.4	-0.29
16.2	-4.7	-3.7	-0.27
15.6	-4.7	-3.8	-0.30
14.9	-3.5	-2.9	-0.27
14.3	-3.1	-2.5	-0.29
13.6	-2.8	-2.9	-0.26
13.0	-2.8	-3.4	-0.22
12.3	-2.8	-6.3	-0.33
11.7	-2.8	-6.4	-0.38
11.0	-2.8	-6.6	-0.29

One clear indication from this density is that the $-S-S-CH_2-NH-\underset{\underset{O}{\|}}{C}$ segment of CYS191-220 moves in the y-direction by approximately 2A (see Figure 34). The coordinates of positive difference density corresponding to this motion are as follows

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
19.5	-8.6	-0.7	+0.28
18.9	-8.6	-0.6	+0.28
18.3	-8.6	-0.5	+0.22
19.5	-6.3	-3.3	+0.25
18.9	-5.7	-3.1	+0.28
18.3	-5.1	-3.5	+0.23 .

The carboxylate group of ASP194 rotates in a counterclockwise direction around the C_α-C_β bond (see Figure 34) and it repositions itself in a direction opposite to the direction of the movement of the terminal $-NH_2$ of ILE16 (toward the peptide bond of GLY140-TRP141). The positive difference density which corresponds to the above motion of the carboxylate end of ASP194 has the following coordinates.

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
16.2	-6.3	-9.0	+0.23
15.6	-6.3	-9.2	+0.27
14.9	-6.3	-9.2	+0.27

It has been speculated elsewhere by Blow[147] that the deprotonation of the $-NH_3^+$ of the ILE16 will cause the ASP194 to seek an alternative in

a more polar environment, for instance, that of HIS40. There are no such indications here; the ASP194 moves in that general direction but not nearly close to HIS40.

A very significant negative difference density ($0.38eA^{-3}$), with coordinates given in Table XVIII, second row from the bottom, appears on the native density assigned to the peptide bond ASP194-GLY193. It is accompanied by a positive difference density with the following coordinates

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
12.4	-0.25	-8.0	+0.26

This pair of difference densities suggest a movement of a segment of the chain of the dipeptide ASP194-GLY193 towards the -OH group of SER195. Since SER195 is one of the three catalytic residues of the active site, an interaction with or even the partial blocking of this crucial hydroxyl group can conceivably greatly influence the activity of the enzyme.

3-3. Tyrosine146 Chain.

The difference Fourier map shows a very pronounced negative density located on native density of the tetrapeptide TYR146-ARG145-THR144-LEU143, which begins at TYR146 and follows the main chain, shaped as a U, and terminates at the side chain of LEU143 (Figure 36). There is no difference density on the side chain of ARG145 probably due to the lack of density in the native structure in that segment of chain. Except for the local two fold equivalent of negative difference density of the carboxylate group of TYR146 (discussed in section 1), there is no two fold counterpart for the negative density just described. The lack of local two fold symmetry is a characteristic of all three chains discussed up to

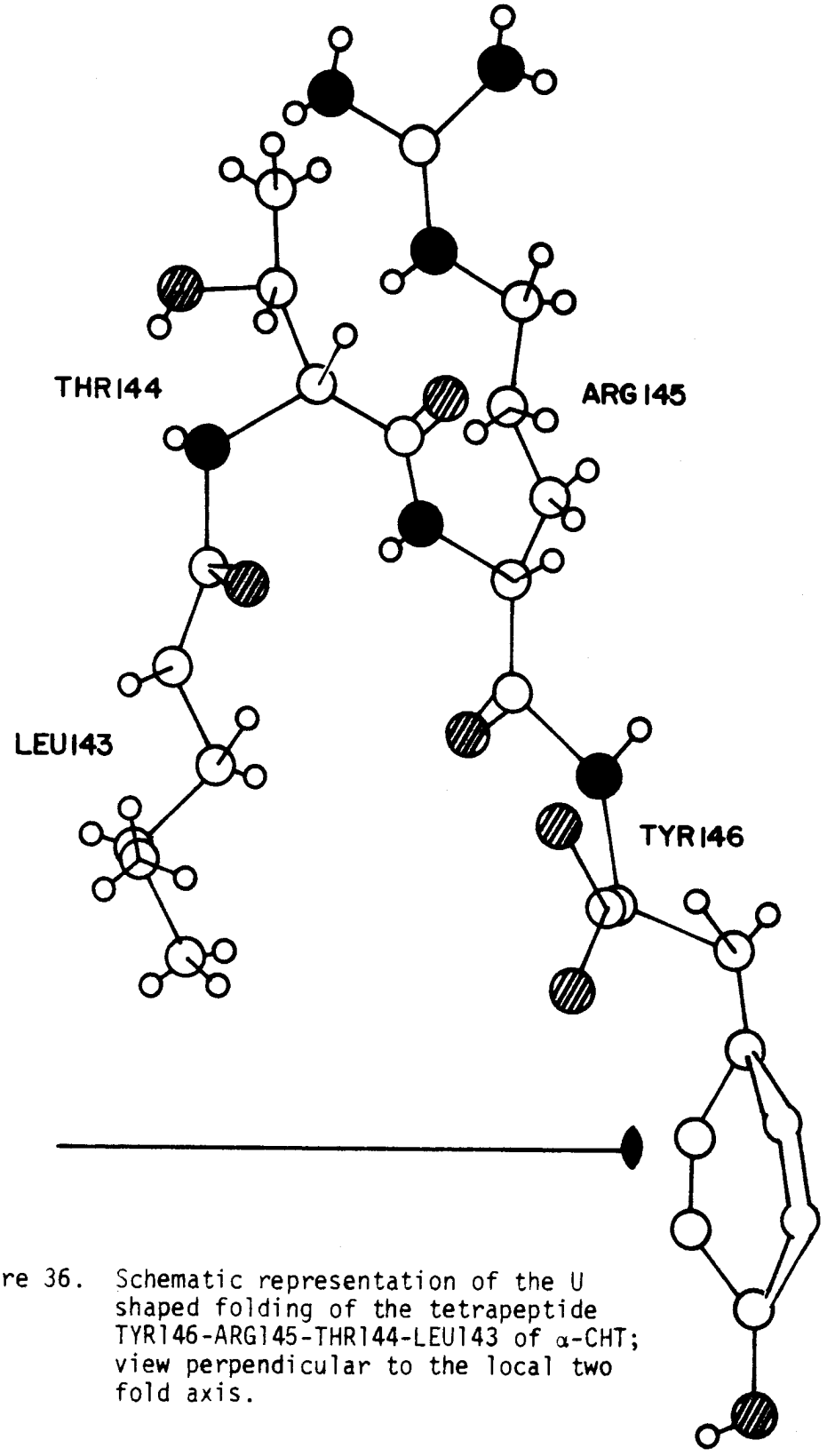


Figure 36. Schematic representation of the U shaped folding of the tetrapeptide TYR146-ARG145-THR144-LEU143 of α-CHT; view perpendicular to the local two fold axis.

now which reorient themselves with increasing pH. This is not the general rule however, as most of the isolated structural changes which occur with the increase in pH show relatively good local two fold symmetry. The foregoing indicates that the two crystallographically independent molecules of α-CHT, can sustain approximately equivalent small structural changes (after all they are chemically identical) but no extensive structural changes can occur independently (due to their close contacts), which shows that the presence of one molecule influences the other.

The coordinates of the negative density starting from the peptide bond of TYR146-ARG145 are given in Table XIX, from which it can be seen that the peak heights of some are particularly high (up to $0.40eA^{-3}$). Table XX gives the coordinates of the positive difference densities which correspond to those of Table XIX. As before the positive difference densities are less extensive and lower in peak height.

The side chain of THR144 is embedded in particularly large negative density of up to $0.35\text{-}0.40eA^{-3}$ with a positive counterpart of $0.25\text{-}0.3eA^{-3}$. This density gradient implies a movement of the THR144 side chain toward the side chain of ASN140. The motion can be achieved by a rotation around the C_α-C_β bond and C_α-CO bond of THR144, which are superimposed on a small translational motion of the whole chain. The side chain of LEU143 does not seem to move severely, although it is near a small gradient of difference density. The part of the chain starting from the α-carbon of TYR146 to the α-carbon of THR144 moves approximately parallel to itself and towards the direction of the -S-S- bridge of CYS191-200. The overall movement of this hairpin shaped tetrapeptide can be described as a

TABLE XIX. Coordinates of Negative Difference Peaks in TYR146 Tetrapeptide Region.

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
8.4	-8.3	5.0	-0.21
8.4	-9.9	1.4	-0.26
7.8	-9.2	-3.3	-0.22
7.1	-8.9	1.9	-0.18
6.5	-9.2	-1.3	-0.22
5.8	-8.6	-2.4	-0.34
5.2	-8.3	-2.6	-0.40
4.5	-8.6	-2.2	-0.36
3.9	-9.2	-1.8	-0.25
8.4	-11.8	-0.4	-0.25
7.8	-11.8	-1.0	-0.34
7.1	-11.8	-1.2	-0.38
6.5	-11.4	-1.3	-0.34
5.8	-11.8	-1.7	-0.27
5.2	-11.8	-2.1	-0.24
4.5	-12.4	-2.7	-0.26
3.9	-12.4	-3.1	-0.25

TABLE XX. Coordinates of Positive Difference Peaks in TYR146 Tetrapeptide Region.

x(A)	y(A)	z(A)	Peak Height(eA^{-3})
16.2	6.0	1.1	+0.21
15.6	5.7	1.3	+0.27
14.9	5.7	0.4	+0.24
14.3	5.4	-1.3	+0.22
16.2	10.2	-1.6	+0.18
15.6	10.4	-2.5	+0.27
12.3	9.9	0.4	+0.27
11.7	8.9	0.5	+0.30
11.0	8.3	0.5	+0.25
11.7	7.0	2.8	+0.23
11.0	7.6	2.9	+0.22

stretching of the two stems in opposite directions accompanied by a translation toward one another.

The cause of this motion can probably be attributed to the deprotonation of HIS57' and therefore the disruption of the ion pair HIS57' ⟷TYR146. The fact that the motion is not confined at the carboxylate terminal of TYR146 as in the case of TYR146' is apparently due to the influence of the movements of the two neighboring chains starting with ILE16 and ASP194. In fact, the part which contains TYR146 and ARG145 moves toward the direction of the forementioned chain.

4. Surface Interactions

In this section we will be concerned with surface changes of α-CHT dimer caused by the change of pH from 3.6 to 8.7. One result of this change in the hydrogen ion concentration of the solvent is that the electrical properties of the surface change drastically. Although the histogram shown on Figure 9 is based on naive assumptions, nonetheless, it shows correct trends; the difference in the surface charge can be as high as 10 units from pH 3.6 to 8.7 due to the deprotonation of basic moieties of the protein. In addition to the direct effects that this change in charge can create, subsequent redistributions of the surrounding ions (Figure 24) can cause changes in the local physical constants of the solvent (usually water), with almost unpredictable effects for the structure of the protein. The Fourier difference density map has revealed a variety of isolated troughs and peaks located mainly on the exterior of the molecule, usually close to polar residues and with peak heights ranging from 0.15 to 0.35eA^{-3}. Although a surface activity of this kind is expected, its interpretation is difficult in many of the cases. Thus,

a positive difference peak located close to a polar group can be interpreted as:

(a) Solvent molecules like water or clusters of water, anions or cations that have been "attached" to the group under consideration.

(b) A movement of the particular group and subsequent occupation of its previous position by the surrounding solvent molecules, so the negative density indication is suppressed.

(c) An "ordering" effect which makes the native electron density more pronounced in the derivative structure.

(d) Possibly a Fourier artifact for the smaller peaks.

The observations of this section are condensed into a table (Table XXI) which gives a serial number of the peak, coordinates in A with respect to the origin previously defined, peak height in eA^{-3}, deviations from local two fold symmetry, and finally, comments concerning the peak(s) under consideration.

The residues THR135, SER186 and SER159 are located on the "top" ($y \sim 0$) of our model of the α-CHT molecule I and form an approximate isosceles triangle with sides (THR135)-(SER186), (SER186)-(SER159) of about 9.5A and a smaller side (SER159)-(THR135) of $\sim$6.0A. Molecule II' of another dimeric molecule resides above molecule I, 4 to 6A away along the y direction (Figure 37). Figure 37 gives a view of the geometrical arrangement described viewed down the y axis (an xz projection). The circled negative and positive signs indicate the difference densities 11', 14', 16' and 18 previously discussed. As the pH changes from 3.6 to 8.7, an expansion of $\sim$0.7A occurs in the dimension of the $\vec{b}$ axis. Since the $\vec{b}$ axis has already increased considerably in the pH 5.4 region and some of the above densities were also observed at that pH (though

TABLE XXI. Surface Interactions

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})	Comments
1	-3.9	10.0	-15.4	+0.37	Peak 1 is located on the surface of the dimer and between the main chains of LYS84, SER109. A hydrogen bond formed between the -NH of the amide bond of SER190 and the >C=O of the amide of LYS84 will be disrupted if the cause of this density is an ion or a cluster of water molecules; there is no accompanying negative difference density.
*1'	-3.9	-7.9	16.7	+0.20	
	(0.0)	(2.1)	(1.3)	(0.17)	
2	-2.6	-7.3	5.1	+0.31	Difference density 2 is "surrounded" by the main chain of LYS36-ASP35-GLN34. It is of significant height and the local two fold relationship of 2 and 2' is remarkably good; there is no negative counterpart.
2'	-2.6	7.4	-4.6	+0.28	
	(0.0)	(0.1)	(0.5)	(0.03)	
3	-1.3	-3.5	-9.1	+0.20	Peak 3 is close to the carboxylate group of ASP72 which can form an ion pair with HIS40. Upon deprotonation of the latter, the electrostatic interaction is disrupted causing a small movement of the carboxylate group, away from HIS40 and toward the solvent. It could also be a solvent effect because there is no accompanying negative difference density. It was also present in the 5.4 work[154] but to a lesser extent.
3'	-1.3	3.6	11.4	+0.18	
	(0.0)	(0.1)	(2.3)	(0.02)	
4	1.0	-2.2	-6.9	-0.22	Difference density 4 is located close to HIS40; the deprotonation of the imidazolium ring of this residue changes the local environment, and this negative density could suggest the loss of a H_2O molecule or an ion although no native density superimposes with the negative difference density. There is no positive counterpart.
4'	0.65	2.3	6.9	-0.22	
	(0.35)	(0.1)	(0.0)	(0.00)	

*Prime numbers correspond to molecule I'.

TABLE XXI. (Continued)

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})	Comments
5	2.0	18.6	1.9	-0.22	Peak 5 is located on the native density corresponding to the $-NH_3^+$ of LYS36. That the side chain of a LYS can move is not surprising; there is no two fold equivalent for 5 and this is reasonable for a LYS residue and also because LYS36 does not show two fold symmetry in the native density[154]
6	6.0	-6.6	2.7	-0.26	Peak 6 is located at the $-COO^-$ terminal of ASN245, at the end of the only true piece of α-helix in α-CHT; There is no positive counterpart.
6'	6.5	7.7	-3.4	-0.27	
	(0.5)	(1.1)	(0.7)	(0.01)	
7	6.5	-16.3	-8.3	+0.30	The positive difference density 7 is very extensive going through 8 layers (or 5.2A) in the x direction (coordinates given here correspond to highest peak located approximately in the middle), whereas 7' goes through only 3 layers (∿2A). The corresponding negative difference density has coordinates numbered 8 and 8'.
7'	6.5	15.0	11.1	+0.21	
	(0.0)	(1.3)	(2.8)	(0.09)	
8	6.5	-18.2	-10.3	-0.20	The negative difference density 8 is also fairly extensive (7 layers in x or 4.5A) and is located on the native density which has been assigned to the residue GLU21. This particular region of α-CHT is fairly complicated due to two reasons: (a) it is an interface between dimeric molecules (see Figure 20) where U indicates the contact, and (b) this region is near the position of the uranyl binding site of the UO_2^{2+} derivative used in the phase determination[149]. The UO_2^{2+} was located between the carboxylate groups of GLU21 in molecule I and ASP'153 in molecule II'.
8'	9.7	18.9	10.3	-0.20	
	(3.2)	(0.7)	(0.0)	(0.00)	

TABLE XXI. (Continued)

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})	Comments
					The difference densities 7, 8 and their two fold equivalents suggest a movement of GLU21 toward GLN156 and ARG154 in such a way as to move away from molecule II'.
9	-10.4	-16.3	14.6	-0.26	This peak is found on the native density assigned to amide side chain of GLN157 which is also close to the uranyl region; a positive difference density close to the above negative implies a rotation of the amide terminal around the C_α-C_β bond. There is no local two fold equivalent.
10	22.8	21.5	-17.6	+0.25	This is a fairly extensive difference density; it is located on the native density which corresponds to the side chain of LYS93 which is exceptionally well defined. The difference density covers the α, β and γ carbons of the side chain. There is no two fold equivalent and it implies an increase in ordering of the side chain.

{A series of difference densities were observed on the interface between dimeric molecules I and II' (see Figure 20) which correspond mainly to molecule II'. They have the same y coordinate and they are close to residues with hydroxyl groups. Their coordinates are given below starting with the number 11, and ending with the number 16.}

11'	21.8	-20.1	-4.0	-0.24	Peak 11' is located on the -OH group of SER186' (molecule II'); it goes through five layers in x direction (∿3.3A) and there is no positive counterpart or two fold equivalent.

TABLE XXI. (Continued)

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})	Comments
12'	22.7	-20.1	-9.5	-0.27	This peak is located at the edge of PRO161' (molecule II'), and is not very extended; there is no two fold equivalent density, but it has a positive counterpart with the following coordinates (peak 13').
13'	22.7	-20.7	-6.6	+0.28	Peak 13' is located on the opposite edge of PRO161' from that of negative density 12' and is toward the neighboring SER186'. A "tipping" motion of PRO161' ring is implied.
14'	25.0	-20.1	-14.1	-0.28	Peak 14' is on the α-carbon of THR135' (molecule II'); there is no local two fold equivalent or corresponding positive difference density.
15'	25.3	-20.1	-3.8	-0.25	This peak is located on the peptide bond of ALA185'-SER186' (molecule II'). There is no local two fold equivalent or corresponding positive density.
16'	22.4	-20.1	-7.1	+0.24	It is located on the peptide bond of LEU160'-SER159' (molecule II') approximately 6A from SER186.
16	22.4	18.3	6.6	+0.18	
	(0.0)	(1.8)	(0.5)	(0.06)	
17'	27.4	-19.4	-10.0	-0.30	The negative difference density 17' corresponds to the -OH group of THR135' (molecule II'); there is a corresponding positive difference density with the following coordinates (peak 18).
17	27.3	21.0	9.2	-0.21	
	(0.01)	(1.6)	(0.8)	(0.09)	

TABLE XXI. (Continued)

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})	Comments
18	28.3	-15.6	-11.4	+0.28	Difference density 18 is close (∿3A) to the hydroxyl group of SER186; it is fairly extensive (6 layers in x direction or ∿4A), and does not have a local two fold equivalent. From difference peaks 17' and 18 a movement of THR135' toward SER186 is implied.

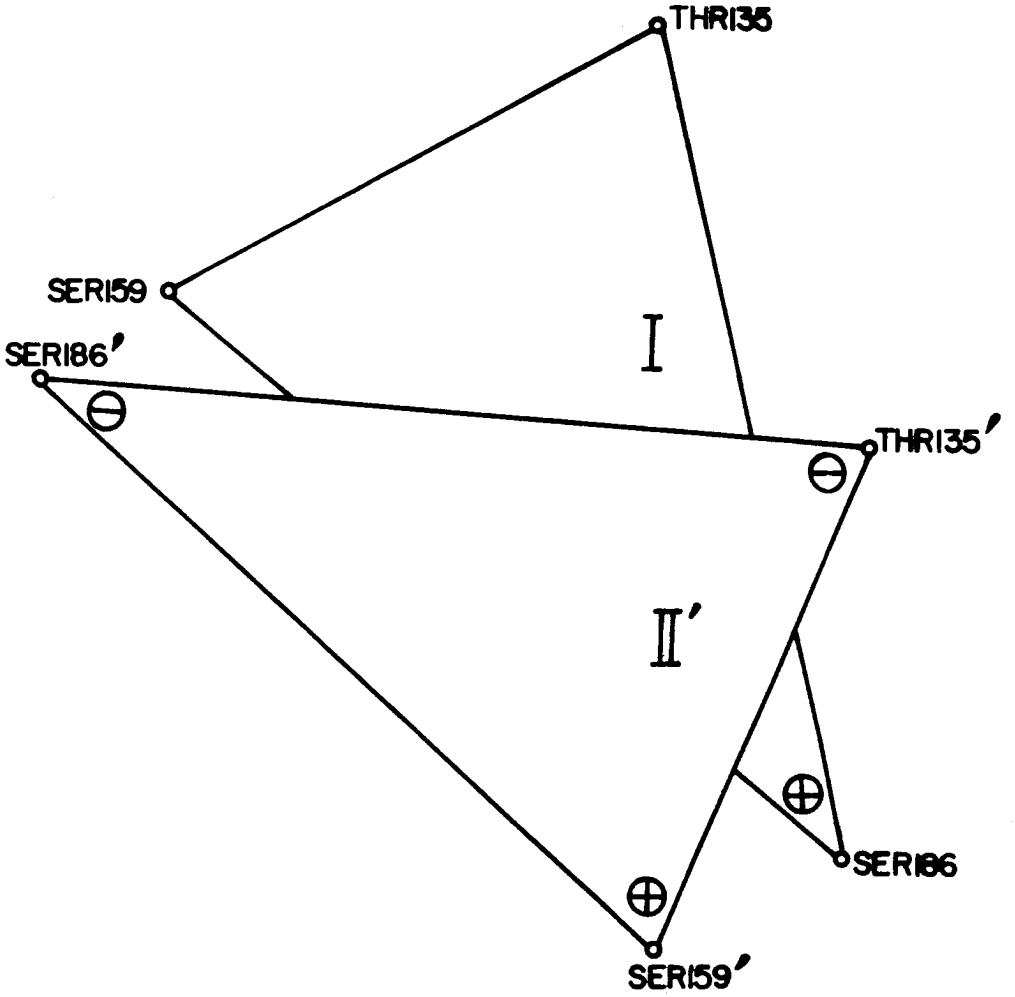


Figure 37. Geometrical arrangement of the residues THR135, SER159, SER186 and the corresponding residues on molecule II' of another dimer. The circled positive and negative signs indicate positive and negative difference densities respectively.

at considerably smaller peak heights)[154], the change of the interactions, of these groups might be related to the change in the axial length.

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
19'	29.3	6.5	9.2	-0.27

Difference density 19' is on the native density of the well defined PRO225' (molecule I'). There is no two fold equivalent and a corresponding positive difference density exists with coordinates:

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
20'	29.9	3.9	10.9	+0.27

A movement of PRO225' is implied toward the sulfur-sulfur (-S-S-) bond of CYS168-182'. A negative difference density ($\sim$0.2eA^{-3}) on the -S-S- bond is probably due to the approach of PRO225'.

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
21'	34.2	10.0	8.1	-0.29
22'	32.0	10.6	8.1	+0.26

Difference densities (21'), (22') correspond to the phenolic group of TYR171' (where the phenyl is perpendicular to the local two fold axis). The above pair implies a rotation around the C_α-C_β bond in such a way as to make the aromatic plane parallel to the two fold axis. There is no two fold equivalent difference density. This is consistent with the fact that this region also displays lack of two fold symmetry in the electron density of the native enzyme[150].

In conclusion, the following can be summarized:

(a) Substantial changes in electron density are observed on the surface of the dimeric enzyme, which are the result of a change of pH.

(b) Virtually all the difference densities are located close to polar groups ($-COO^-$, $-NH_3^+$ or -OH).

(c) About half of the difference densities have a local two fold equivalents, an observation which strengthens the assumption that isomorphous conditions are maintained and that the difference electron densities we are observing are very real.

B. Low pH (2.5) Structural Changes of Crystalline α-CHT

The high pH (8.3) structure of crystalline α-CHT was discussed in the previous sections; for completeness and continuity, the effects of higher hydrogen ion concentration on the native enzyme were also studied. Due to the fact that a change of pH from 3.6 to 2.7 cannot greatly influence the electrical properties of the protein molecule, large changes were not expected in the structure of the molecule. It would certainly be interesting to study the effects of larger hydrogen ion concentration (pH∿1) on crystals of α-CHT, because under such conditions all the carboxylate groups would be protonated, the electrical properties of the molecule would change drastically, and crucial ion pairs like the ILE16⟷ASP194, or HIS57⟷TYR146' would be disrupted. However, this was not possible because crystals of α-CHT crumble to a finely divided solid at pH 1.0. That the structural changes at pH 2.5 region would be small was also suggested by the unit cell parameters of the low pH crystals which, except for the β-angle, are essentially the same as those of the native protein (Table XV), and also from the axial intensity

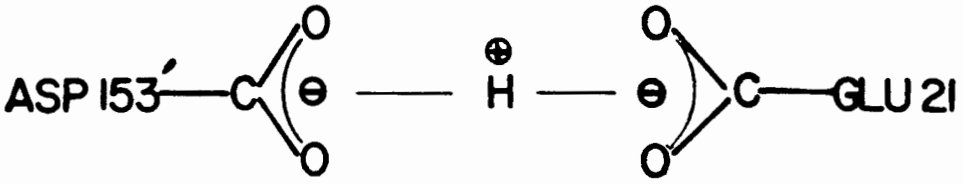
distributions (Figure 26) which did not show large changes. Nevertheless, a set of intensity data was collected and a difference Fourier synthesis was calculated between the structures at pH 3.6 and 2.5. As expected, the structural changes were generally small and localized, with difference densities being of the order of $0.2eA^{-3}$. These changes did not correspond to any internal rearrangements of the molecule, but rather, they tended to be associated with structural changes near or in the aqueous environment around the surface of the molecule. The fact that in many cases the difference density had a local two fold equivalent counterpart gives additional significance to these smaller observations.

The largest (positive) difference density was found to be close to the carboxyl group of ASP153', in the uranyl binding region between the two molecules (see section 4 of III). Its coordinates are:

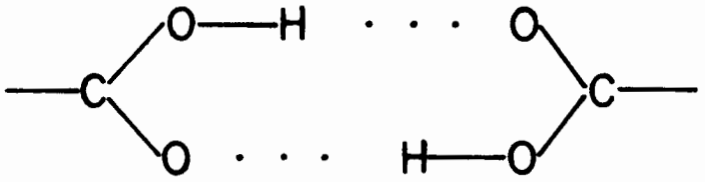
Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
1	2.6	19.5	26.3	+0.32

This peak is fairly extensive, extending 3.3A in the x-direction; it has no two fold equivalent because of the uniqueness of the site and essentially no negative counterpart. As discussed previously, the uranyl site is fairly complicated, especially with respect to detailed interpretation. With this in mind, the following are possibilities that can account for the observed difference density.

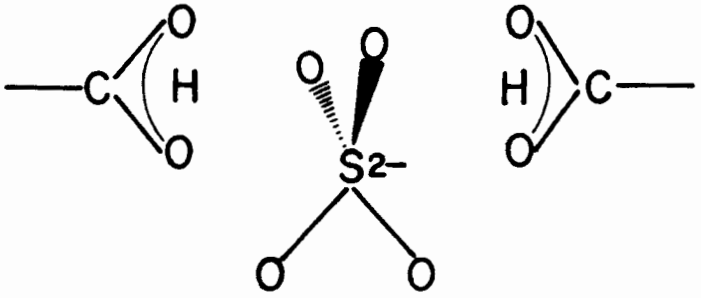
The carboxylate group of ASP153' faces the carboxylate of GLU21; the energy of this system decreases when a proton is in the negative potential between these two groups (Figure 38a) (similar to the "one electron bond"). This probably represents the situation at pH 3.6;



(a)



(b)



(c)

Figure 38. Possible schemes of interactions between the inter-dimer carboxyl groups of ASP153' and GLU21 residues.

as the pH is decreased, another proton can be accomodated and the energy of the system is additionally lowered with the formation of a pair of hydrogen bonds (as in Figure 38b). This effect would lead to an increased order and could be responsible for the observed positive difference peak. On the other hand, from the magnitude of the density ($0.32eA^{-3}$), it could also correspond to a partially ordered sulfate ion (Figure 38c), or one of lower occupancy. This would be plausible if one considers that the net positive surface charge of the molecule is increased slightly by the lowering of the pH, and a counter ion would be attracted more easily. A SO_4^{2-}-SeO_4^{2-} exchange experiment at pH 2.5 could verify this possibility. As the pH is increased to 5.4, changes also occur in this region[154], suggesting the ionization of the lone proton of the pH 3.6 configuration (Figure 38a).

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
2	2.6	-11.2	-5.3	+0.25
2'	2.6	12.0	5.0	+0.22
$\|\delta\|$	(0.0)	(0.8)	(0.3)	(0.03)

The positions of localized sulfate ions in crystals of α-CHT have been determined in an exchange experiment with selenate ions by Tulinsky and Wright[187]. The coordinates of sulfate 149 and its two fold equivalent, which are near amino terminal ALA149, are given below[187]:

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
SO_4^{2-}149	3.3	-12.8	-4.2	-0.71
SO_4^{2-}149'	2.4	15.4	+5.0	-0.73
$\|\delta\|$	(0.9)	(3.4)	(0.8)	(0.02)

From these, it can be seen that the positive difference density at the peaks of the pH 2.5 structure is close to sulfate 149 but 2' is approximately 3.5A away from sulfate 149' and is located between ALA149' and ASP64 (Figure 39). While 2 and 2' have excellent two fold symmetry, the same is not true of sulfates 149 and 149'. A consistent way of interpreting these observations is to suggest an increase in occupancy of sulfate ions in this region of the α-CHT dimer. Although this is not an interesting result from a biochemical point of view, nonetheless, it illustrates the power of the difference method under favorable circumstances. The increase in density of the SO_4^{2-} is probably related to the protonation of ASP64 which helps to neutralize the charge of terminal ALA149'; a negative density of about $-0.17eA^{-3}$ located on the carboxyl group of ASP64 supports this suggestion.

The coordinates of an extended negative difference density (∿5.2A in x-direction), located between the main chains of the peptides ALA55-THR54-VAL53-VAL52 and THR104-LEU105-LEU106 are given below:

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
	8.4	7.4	-19.1	-0.21
	9.1	7.4	-19.2	-0.18
	9.8	7.4	-19.1	-0.22
	10.4	7.2	-18.9	-0.21
3	11.1	6.8	-18.4	-0.19
	11.7	6.3	-17.8	-0.19
	12.4	6.3	-17.2	-0.20
	13.0	6.3	-17.1	-0.17

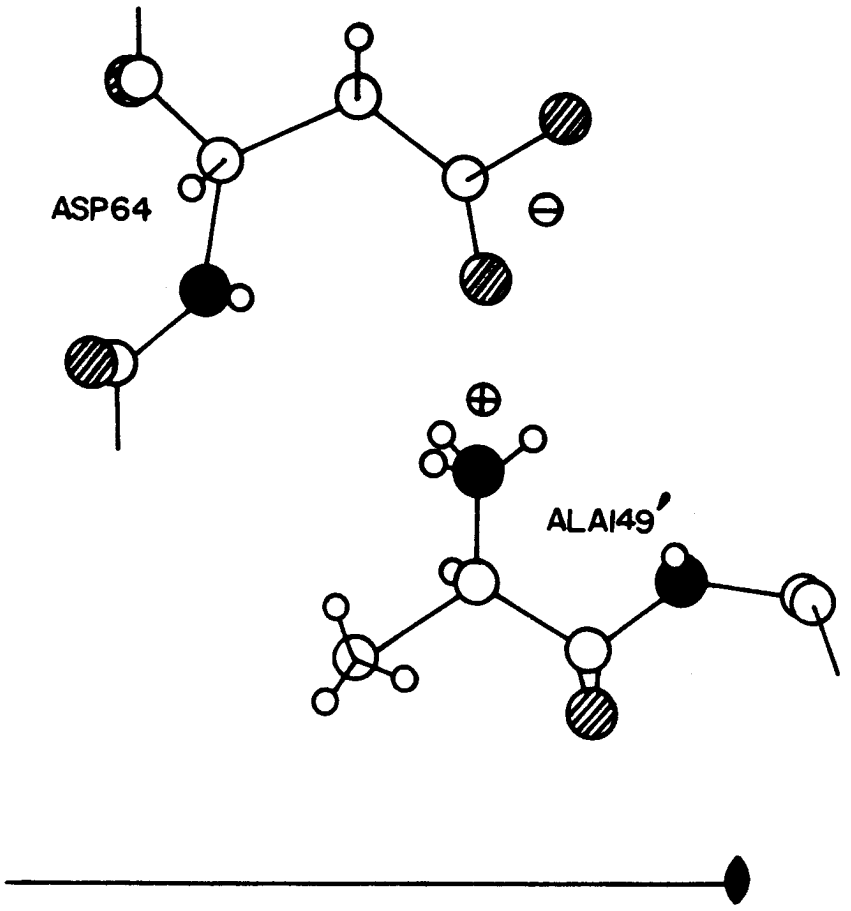


Figure 39. Schematic representation of the residues ASP64 and ALA149' of α-CHT; view perpendicular to the local two fold axis.

Figure 40. Geometrical arrangement of a part of α-CHT molecule showing the manner of formation of an extended pleated sheet.

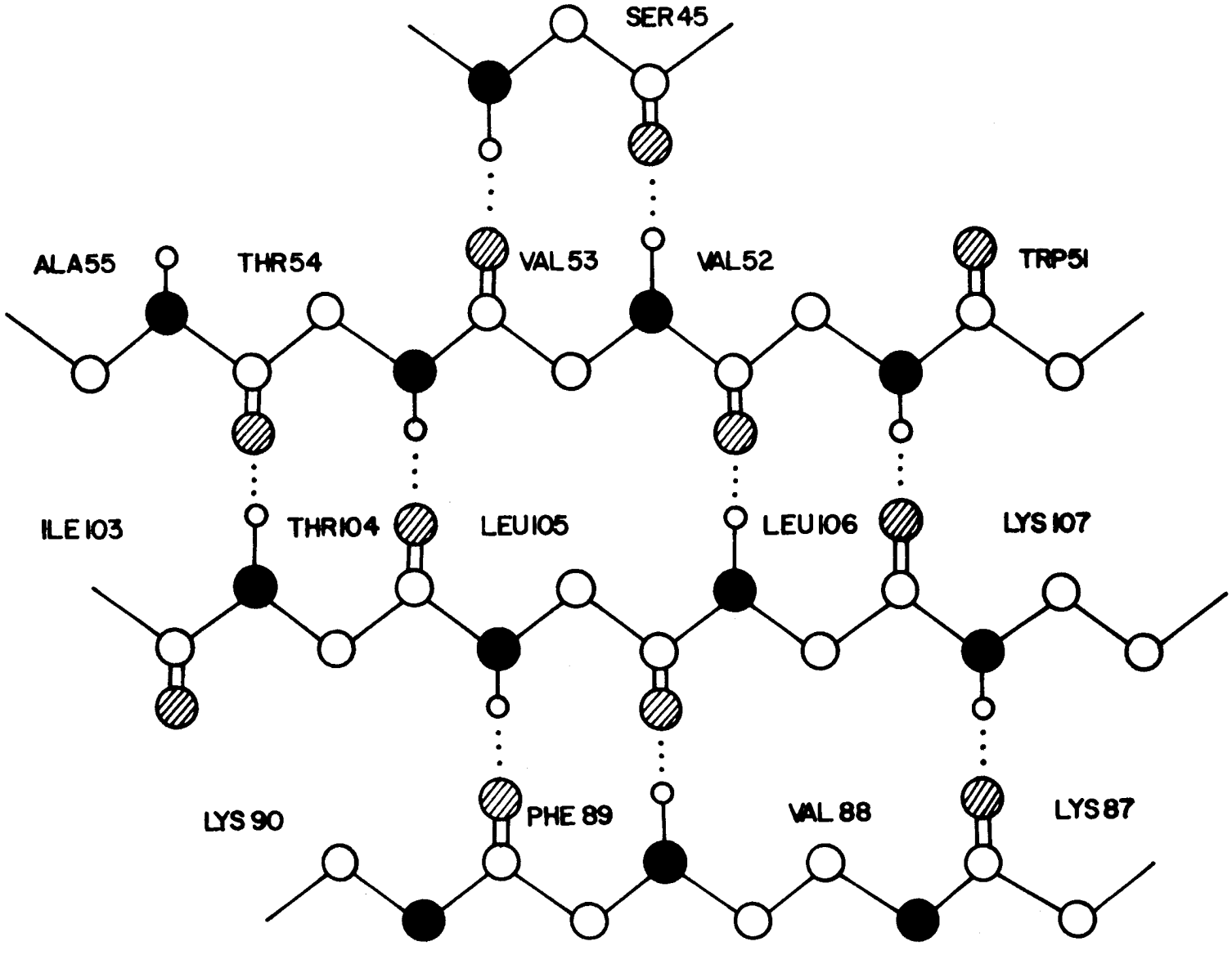

SER 45
ALA 55
THR 54
VAL 53
VAL 52
TRP 51
ILE 103
THR 104
LEU 105
LEU 106
LYS 107
LYS 90
PHE 89
VAL 88
LYS 87

Figure 40 illustrates the geometrical arrangement of the above peptides and the way they form an imperfect extended pleated sheet with neighboring residues. There is no two fold equivalent or any positive counterpart to this density. The native density is very well defined in this region of the molecule due to the ordering effect caused by the net of 9 hydrogen bonds to form this distorted pleated sheet (Figure 40). The possibility of breaking some hydrogen bonds due to the pH change around the amide bonds could account for this negative density.

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
4	16.9	-12.4	-9.9	+0.17
4'	16.9	11.1	10.2	+0.22
$\|\delta\|$	(0.0)	(1.3)	(0.3)	(0.05)

The densities (4) and (4') are located between the side chains of ALA108 and ILE16. Even speculative explanation cannot be offered to rationalize the above density and its two fold equivalent.

Another sulfate ion and its local two fold equivalent which have been unambiguously located with the exchange experiments[187] are those of 217 and 217' close to SER217 and SER217' respectively; the coordinates of these SO_4^{2-} are

	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
SO_4^{2-}217	28.2	4.8	+0.2	1.17
SO_4^{2-}217'	27.9	-2.6	-0.3	0.88
	(0.3)	(2.2)	(0.1)	(0.29)

A negative difference density is located on sulfate SO_4^{2-}217' with a positive counterpart approximately 3A away; their coordinates are the following:

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
5'	30.2	-2.2	0.0	-0.18
6	31.5	-1.5	-2.6	+0.20

It can be seen that the coordinates of 5' compare well with the coordinates of the sulfate ion close to SER217'. Thus, a partial migration of the latter to a new position given by coordinates 6 might be occurring (approximately 3-4A). Taking the ratio of the peak height of the difference multiplied by two (Luzzatti's correction for approximate phasing[188]), to the peak height of the sulfate in the native density, 2x(0.18)/0.88, the occupancy of the sulfate is apparently reduced by 40-50%.

Peak Number	x(A)	y(A)	z(A)	Peak Height(eA^{-3})
7	26.7	-0.9	-3.4	+0.17
7'	26.7	1.0	4.1	+0.12
	(0.0)	(0.1)	(0.7)	(0.05)

The positive difference density 7 is located close to the SER217 residue in the interface region of the dimer molecule. The difference Fourier shows many difference peaks around SER217 and SER217' ($\sim$0.1-0.15eA^{-3}) without two fold equivalents which most likely can be attributed to solvent effect. This region of the enzyme also lacks two fold symmetry in the native structure.

XI. FINAL COMMENTS

α-CHT is one of the most thoroughly studied enzymes. The vast biochemical literature concerning its properties has now even been increased by crystallographic studies, and by the coordinates for almost every atom in its sequence. Nevertheless, the main question of "how the enzyme works" seems far from being answered. α-CHT is the end product of a series of tryptic and chymotryptic cleavages starting from an inactive precursor, chymotrypsinogen (see Figure 18). The two other forms of chymotrypsin, namely π and δ derived along the activation route as well as the γ form, are also active, and their three dimensional structures are similar with the α-form of the enzyme and with each other[195-199]. A comparison of the structure of the native enzyme with the structure of chymotrypsinogen (hereafter denoted as CHTG) would be expected to give answers as to why CHTG is inactive, while α-CHT (and all other forms of the enzyme) is active. Such a comparison has been carried out by Kraut and his collaborators[200] and by Wright[201-202] and the results are somewhat startling: α-CHT and CHTG have very similar three dimensional structures, particularly with respect to the catalytic constituents of the active site. Table XXII gives all residues for which the α-carbon atoms differ in position by more than 3.6A between the CHTG and α-CHT structures[200]. The reason this Table is shown is because most of the differences observed upon activation occur at similar

TABLE XXII. List of Residues for Which α-Carbon Atoms Differ in Position by More than 3.6A Between the CHTG and α-CHT Structures (From Reference 200).

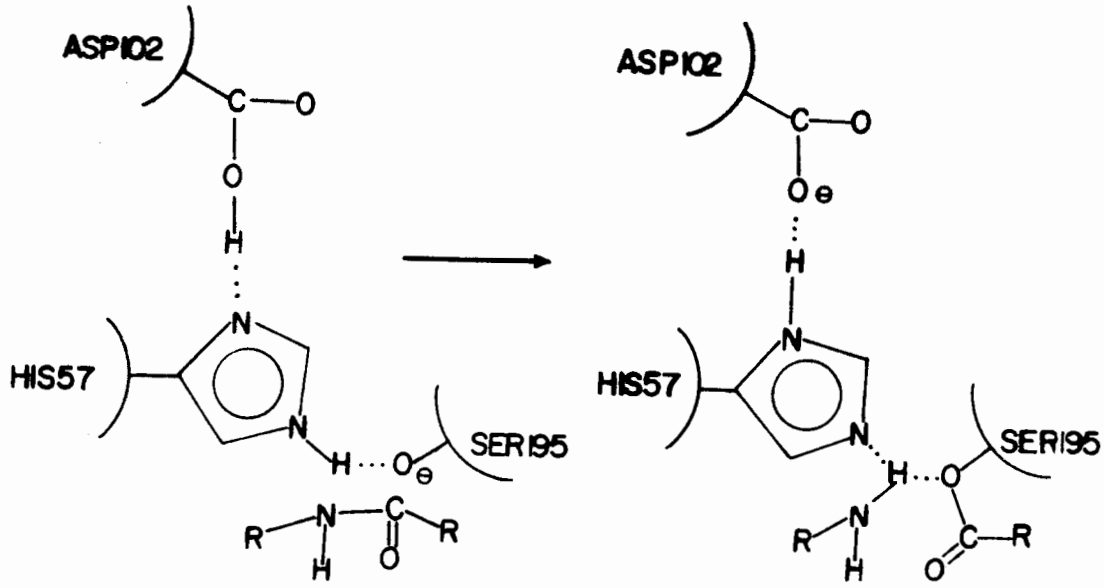
Residue	Displacement(A)
GLY7	4.8
ILE16	11.3
VAL17	6.6
THR37	4.0
GLY38	6.6
ASP72*	5.6
GLY73*	9.6
GLY74*	9.1
SER75*	6.2
SER76*	10.1
SER77*	5.6
THR144	5.9
ARG145	8.7
TYR146	4.6
ALA149	4.7
MET192	8.4
GLY193	6.6

*These residues are ill-defined in the Cambridge structure[148] and show variability between the monomeric units in the MSU structure[150]; therefore, the differences might not be associated with the activation process.

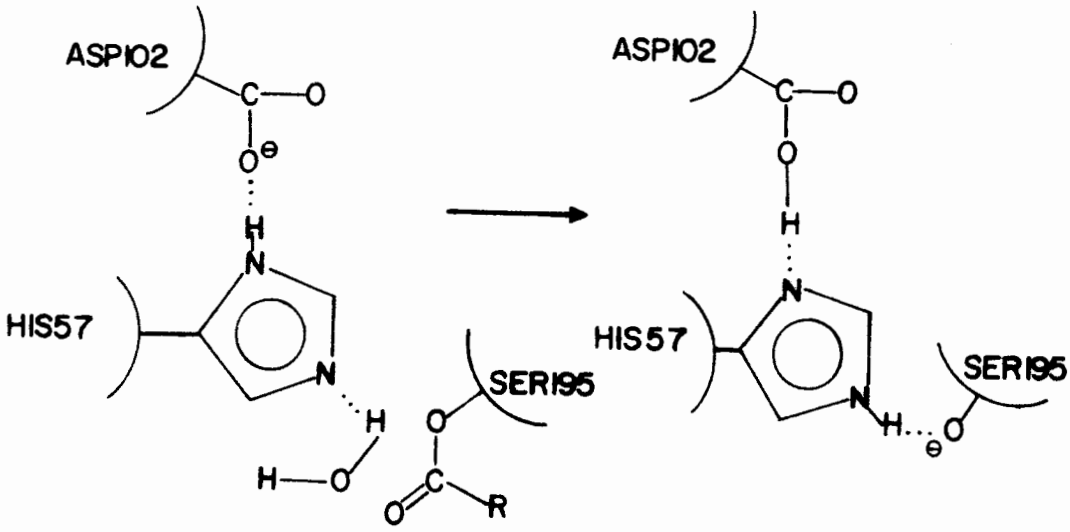
residues to structural changes observed in changing the pH of α-CHT from 3.6 to 8.7. The segments ILE16-VAL17, MET192-GLY193, THR144-ARG145-TYR146, given in Table XXII, were also found to undergo severe conformational changes as the pH was raised from 3.6 to 8.7 (sections 3-1, 3-2, 3-3, respectively). In addition, the high pH structure of α-CHT also shows that the catalytic residues of the active site remain intact similarly to CHTG.

The foregoing indicates that a qualitative correlation exists between the high pH changes of α-CHT and the activation process of CHTG; however, it still does not answer why the latter is inactive. Figure 41 shows the most plausible mechanism of α-CHT hydrolysis of an amide substrate consistent with many known facts[136,202-204]. The crux of the mechanism is the system of hydrogen bonds between ASP102, HIS57 and SER195, which have been referred as the "charge relay system"[136]. Presumably, without the ease of proton transfer, the enzyme would not function and this implies that the geometrical arrangement of the catalytic residues HIS57, SER195 and ASP102 with respect to each other and with respect to the substrate must be very important.

Freer et. al.,[200] suggest that the absence of the specificity pocket (for aromatic side chains) in the structure of CHTG is the reason for its lack of activity. On the other hand, Wright[202] attributes great significance not only to the specificity pocket but also to the ability of the substrate to form a hydrogen bond between the carbonyl group of the substrate and the amide group of the GLY193. He therefore suggests in addition that CHTG is not active because GLY193 is further away and not in an appropriate position for the formation of this crucial stabilizing bond.



(a)



(b)

Figure 41. Possible mechanism of action for α-CHT; (a) acylation step; (b) deacylation step (reference 136).

A severe reorientation of GLY193 was observed in the high pH structure of α-CHT (section IV, 3-2) that could prevent the formation of a hydrogen bond with a substrate or even block the hydroxyl group of SER195 and in this way render the enzyme inactive. To this extent, our results are in agreement with Wright's suggestion and also with the implication of conformational changes around the specificity pocket.

The fact that GLY193 in CHTG cannot form "the hydrogen bond" with a substrate appears to be rather insignificant. What is probably more important is the ability of the proton transfer (charge relay system), in conjunction with an exact geometry of the region in which the active site resides. What is probably involved is a quantum mechanical tunneling phenomenon, where the hydrogen ions can be thought of as wave packets instead of classical particles. The problem of the motion of a proton in a hydrogen bond is to a first approximation a one particle problem involving a fixed outer potential. Each electron pair attracts the proton (for instance, oxygen of SER195, nitrogen of HIS57) and this may be represented by single potential wells having a minimum corresponding to the classical equilibrium position[205]. Since there are two electron pairs involved, the total effect can be represented by an asymmetric double-well potential (Figure 42). The real situation is much more complicated because the potentials (V_1, V_2, V_3) are three dimensional; nonetheless, the principle is the same. In order to be able to use a potential of more or less arbitrary shape, the JWKB (Jeffreys, Wentzel, Kramers, Brillouin)[206-209] approximation will be employed in what follows.

The one dimensional, one particle time independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x)}{\partial x^2}+V(x)\Psi(x)=E\Psi(x). \qquad (107)$$

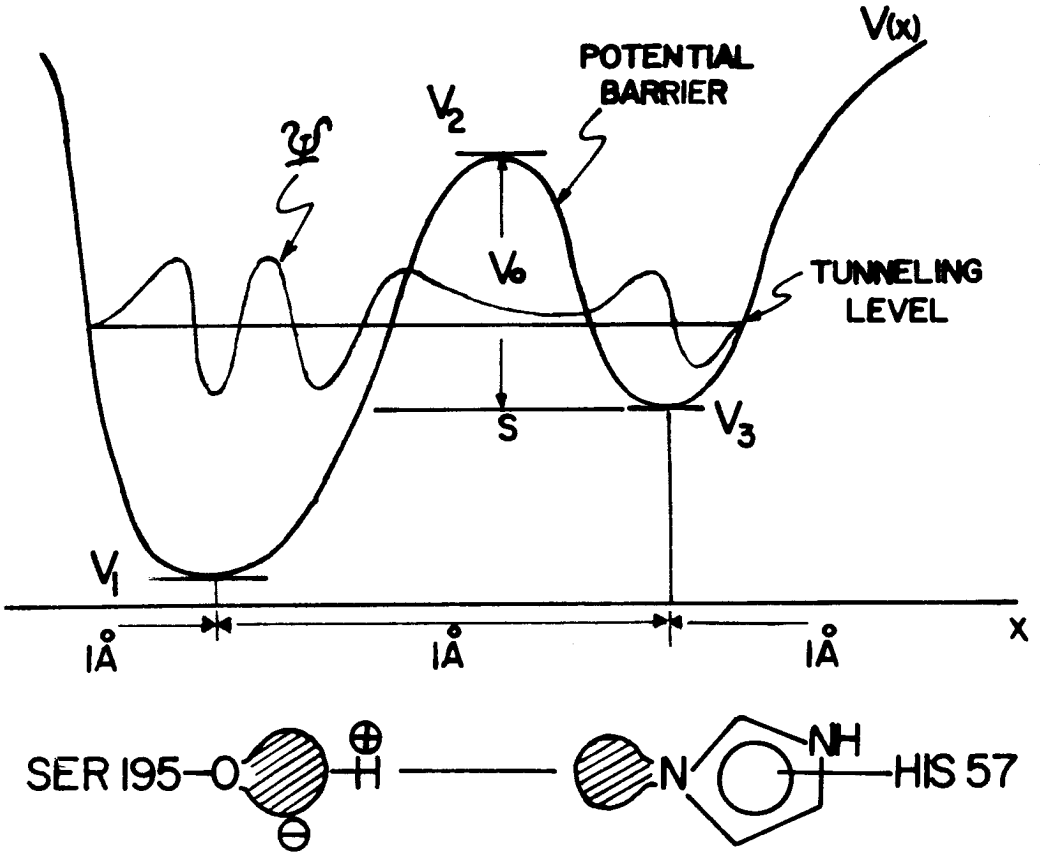


Figure 42. A one dimensional, double well potential between two pairs of electrons ∿3A apart; here these pairs are residing on the imidazole nitrogen of HIS57 and on the oxygen of the hydroxyl group of SER195 (reference 205).

The non-normalized JWKB solution is given by

$$\Psi(x) \simeq \frac{1}{\sqrt{K(x)}} \exp(\pm i \int^{x} K(x)dx), \tag{108}$$

$$K(x) = \{\frac{2m}{h^2}(E-V(x))\}^{1/2}, \; E>V(x), \tag{109}$$

$$K(x) = -i\{\frac{2m}{h^2}(V(x)-E)\}^{1/2} = -iK'(x), \; E<V(x), \tag{110}$$

(see Merzbacher reference 210, or Bohm reference 211).

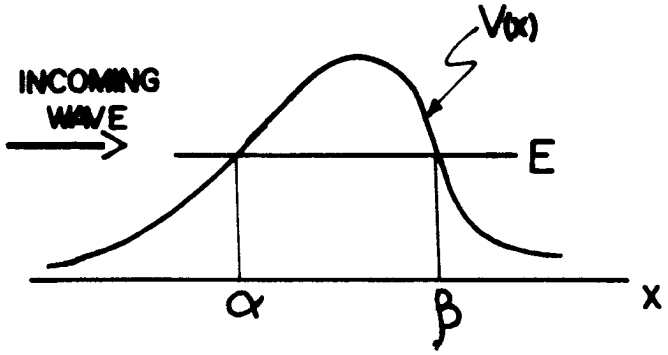
In order for solution (108) to be valid the condition

$$|\frac{dK(x)}{dx}| << |K^2(x)| \tag{111}$$

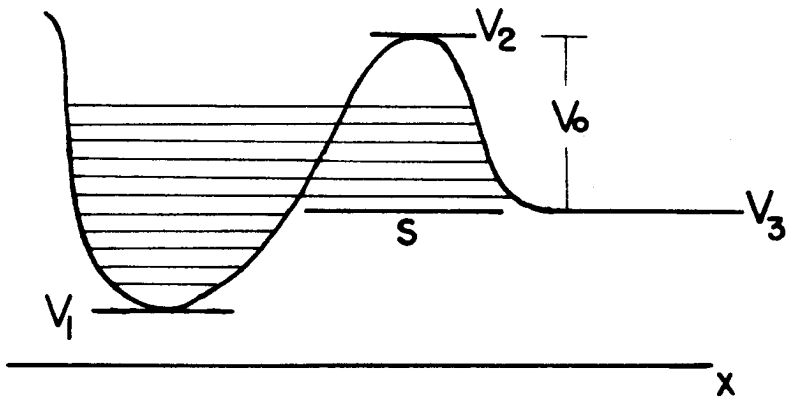
must be fulfilled.

The solution (108) can be used to calculate the transmission coefficient for a potential barrier upon which particles are incident from the left with insufficient energy to pass to the other side classically (Figure 43a). We obtain for the three regions of the potential of of Figure 43a, the following equations as solutions of the Schrödinger equation:

$$\Psi(x) = \begin{cases} \frac{A}{\sqrt{K(x)}} \exp(-i \int_{\alpha}^{x} K(x)dx) + \frac{B}{\sqrt{K(x)}} \exp(-i \int_{\alpha}^{x} K(x)dx), & x<\alpha \\ \frac{C}{\sqrt{K(x)}} \exp(-i \int_{\alpha}^{x} K'(x)dx) + \frac{D}{\sqrt{K'(x)}} \exp(\int_{\alpha}^{x} K'(x)dx), & \alpha<x<\beta \\ \frac{F}{\sqrt{K(x)}} \exp(i \int_{\beta}^{x} K(x)dx) + \frac{G}{\sqrt{K(x)}} \exp(-i \int_{\beta}^{x} K(x)dx), & \beta<x \end{cases} \tag{112}$$



(a)



(b)

Figure 43. (a) A one dimensional potential barrier; incoming wave from the left; α and β classical turning points (at those points the JWKB approximation breaks down).

(b) A double well potential with the right hand well spread out, so there is no "back tunneling" (reference 205).

Using the connection equations for the turning points α and β, the following matrix equation for the constant coefficients of equations (112) are obtained

$$\begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2\theta+\frac{1}{2\theta} & i(2\theta-\frac{1}{2\theta}) \\ -i(2\theta-\frac{1}{2\theta}) & 2\theta+\frac{1}{2\theta} \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix} \tag{113}$$

where

$$\theta=\exp\left(\int_\alpha^\beta K(x)dx\right). \tag{114}$$

The transmission coefficient g is defined as[210]

$$g=\frac{|\Psi_{trans}|^2\, v_{trans}}{|\Psi_{inc}|^2\, v_{inc}}=\frac{|\Psi_{trans}\sqrt{K_{trans}}|^2}{|\Psi_{inc}\sqrt{K_{inc}}|^2}=\frac{|F|^2}{|A|^2},$$

(because $p=K\hbar \rightarrow v=K\hbar/m$), and assuming that G=0 (no wave incident from the right) we obtain

$$g=\frac{4}{(2\theta+\frac{1}{2\theta})^2} \tag{115}$$

For a high and broad barrier, $\theta>>1$, therefore

$$g=\frac{1}{\theta^2}=\exp\left(-2\int_\alpha^\beta K(x)dx\right). \tag{116}$$

The next step is to consider a double well potential of the form shown in Figure 43b, where the right hand well has been spread out so that there is no back-tunneling. In addition, we assume that the potential on the left hand side is parabolic; a classical particle will then be a harmonic oscillator with the energy $(n+1/2)h\nu$, with n=0,1,2,... .

If we further assume that the potential barrier in Figure 43b is of parabolic nature, the integral (116) can be calculated, and according to Löwdin (205)

$$g=\exp\{-\frac{\pi^2}{h}\kappa S\sqrt{2mV_0}\} \tag{117}$$

where κ is defined by the equation

$$E=V_2-\kappa V_0 \tag{118}$$

Thus, $\kappa=0$ and $\kappa=1$ correspond to the top and the bottom of the tunneling barrier respectively. The <u>tunneling rate</u> can be defined as the number of times per second (ν) a particle hits the "wall" of the barrier times the transmission coefficient g;

$$C=\nu g. \tag{119}$$

The frequency ν can be estimated from the uncertainty principle, $\Delta p\Delta x\sim\hbar$; or using b instead of Δx (b is the distance between the two classical turning points for the ground state), and p instead of Δp, we obtain

$$b^2p^2\sim\hbar^2. \tag{120}$$

But $E=\frac{p^2}{2m}\sim\frac{\hbar^2}{2mb^2}=\frac{h\nu_0}{2}$, therefore $\nu_0\sim\frac{h}{4\pi^2b^2m}$. (121)

Using atomic units ($e=1$, $\hbar=1$, $m_e=1$), the transmission coefficient g (equation (117)) for a particle with a mass of a proton becomes

$$g=10^{-14.994\kappa S\sqrt{V_0}}, \tag{122}$$

where V_0 is expressed in electron volts (eV).

The lengths of hydrogen bonds are ~3A (the distance between the hydroxyl group of SER195 and the imidazole ring of HIS57 is approximately

3A). The x axis of Figure 42 can be divided into three equal parts of 1A each; S can be considered as 0.5-1.0A depending on the potential. The frequency ν_0 can be calculated from (121) using a value of $b \sim 0.3$A and a value of $\sim 10^{13}$ sec^{-1} is obtained. According to (119), (122) and the value of ν_0 just obtained, the tunneling rate for a proton is given by the expression:

$$C_0 = \nu_0 g = 10^{13-14.994\kappa S\sqrt{V_0}} \ . \qquad (123)$$

Using this expression with $\kappa=1$ (bottom of the barrier) for a range of values $S\sqrt{V_0}=0\text{-}2eV^{205}$, we obtain a range for tunneling time τ ($\tau = \frac{1}{C}$) from 10^{-13} sec to 10^{17} sec (or 3×10^9 years) respectively.

It is clear that in the above discussion we were dealing with orders of magnitudes due to the different approximations and simplifications; in addition, the effect of temperature on the tunneling was not considered and this also leads to startling influences on the tunneling rate[205]. Nevertheless, it is evident that the results are extremely sensitive with respect to small modifications of the potential energy curves, and it is this dependence which makes the tunneling phenomenon particularly interesting.

In conclusion, it is suggested here that the importance of the charge relay system in α-CHT comes from its geometry which is very favorable for the tunneling process. Structural changes around the catalytic residues caused by a pH change or temperature or whatever can easily change such a sensitive three dimensional potential, thereby causing enormous decreases in the transmission coefficient g, which in turn, reflects the rate at which the enzyme operates. The speculative character of the foregoing discussion has to be admitted; only very detailed calculations can show if some of the ideas presented are applicable and this is presently beyond the capabilities of modern electronic computers.

REFERENCES

REFERENCES

1. A. Schoenflies, "Krystallsysteme and Krystallstructure", Leipzig (1891).

2. E. von Fedorov, Z. Krist. 20, 25 (1892).

3. W. Friedrich, P. Knipping, and M. Laue, Sitzber. Math.-Physik, Kl. Bayer, Akad. Wiss., 303 (1912).

4. M. Laue, Sitzber. Math.-Physik, Kl. Bayer Akad. Wiss., 363 (1912).

5. W. Friedrich. P. Knipping, and M. Laue, Ann. Physik 41, 971 (1913).

6. W. H. Bragg, and W. L. Bragg, Proc. Roy. Soc. (London) A88, 428 (1913); A89, 246 (1913).

7. W. L. Bragg, Proc. Cambridge Phil. Soc. 17, 43 (1913).

8. P. P. Ewald, Ann. Physik 49, 1 (1916); 49, 117 (1916); 54, 519 (1918).

9. L. de Broglie, Thesis, Paris, "Researches sur la theorie des quanta" (1924).

10. C. J. Davisson, and L. J. Germer, Phys. Rev. 30, 705 (1927).

11. "Fifty Years of X-ray Diffraction", edited by P. P. Ewald, N. V. A. Oosthoek's Vitgeversmaatschappij Utrecht, the Netherlands (1962).

12. W. H. Zachariasen, "Theory of X-ray Diffraction in Crystals", John Wiley and Sons, Inc., New York (1945).

13. W. L. Bragg, "The Crystalline State, a General Survey", G. Bell and Sons, Ltd., London (1933).

14. R. W. James, "The Optical Principles of the Diffraction of X-rays", G. Bell and Sons, Ltd., London (1948).

15. H. Lipson, and W. Cochran, "The Determination of Crystal Structures", G. Bell and Sons, Ltd., London (1953).

16. M. J. Buerger, "Crystal Structure Analysis", John Wiley and Sons, Inc., New York-London (1960).

17. B. E. Warren, "X-ray Diffraction", Addison-Wesley Publishing Co. (1969).

18. A. J. C. Wilson, "Elements of X-ray Crystallography", Addison-Wesley Publishing Co. (1970).

19. M. M. Woolfson, "Introduction to X-ray Crystallography", Cambridge University Press (1970).

20. J. W. Jeffery, "Methods in X-ray Crystallography", Academic Press, London and New York (1971).

21. R. W. James, Solid State Physics **15**, 53 (1964).

22. R. V. Churchill, "Fourier Series and Boundary Value Problems", McGraw-Hill Book Co. (1961).

23. J. W. Gibbs, and E. B. Wilson, "Vector Analysis", Yale University Press, New Haven, Conn. (1902).

24. P. P. Ewald, Z. Krist. **56**, 129 (1921).

25. J. C. Slater, "Symmetry and Energy Bands in Crystals", Dover Publications, Inc. New York (1972).

26. P.-O. Löwdin, Advances in Quantum Chemistry **3**, 324 (1967).

27. J. M. Ziman, "Principles of the Theory of Solids", Cambridge at the University Press (1972).

28. A. L. Patterson, Phys. Rev. **46**, 372 (1934).

29. A. L. Patterson, Zeit für Krist. **90**, 517 (1935).

30. M. J. Buerger, "Vector Space", John Wiley and Sons, Inc., New York (1959).

31. F. Bloch, Z. Phys. **74**, 195 (1932).

32. D. W. J. Cruickshank, D. E. Pilling, A. Bujosa, F. M. Lovell, and M. R. Truter, "Computing Methods in the Phase Problem", Oxford, Pergamon Press (1961).

33. R. L. C. Pilgrim, J. Chem. Educ. **51**, 316 (1974).

34. J. E. Falk "Porphyrins and Metalloporphyrins", Elsevier Publishing Co., Amsterdam (1964).

35. "The Chemical and Physical Behavior of Porphyrin Compounds and Related Structures". Edited by A. D. Adler, Published by the New York Academy of Sciences, ANYAA9 206 1-761 (1973).

36. W. Z. Küster, Z. Physiol. Chem. **82**, 463 (1912).

37. R. Lemberg, and J. W. Legge, "Haematin Compounds and Bile Pigments", Interscience, New York (1949).

38. S. J. Silvers and A. Tulinsky, J. Am. Chem. Soc. 89, 3331 (1967).

39. B. M. L. Chen, and A. Tulinsky, J. Am. Chem. Soc. 94, 4144 (1972).

40. P. W. Codding, and A. Tulinsky, J. Am. Chem. Soc. 94, 4151 (1972).

41. E. D. Becker, R. B. Bredley and C. J. Watson, J. Am. Chem. Soc. 83, 3743 (1961).

42. R. J. Abraham, Mol. Phys. 4, 145 (1961).

43. R. J. Abraham, A. H. Jackson, G. W. Kenner, and D. Warburton, J. Chem. Soc., 853 (1963).

44. W. S. Caughey, and P. K. Iber, J. Org. Chem. 28, 269 (1963).

45. G. D. Dorough, J. R. Miller, and F. M. Huenekens, J. Am. Chem. Soc. 73, 4315 (1951).

46. M. Gouterman, J. Mol. Spect. 6, 138 (1961).

47. M. Gouterman, G. H. Wagniere, and L. C. Snyder, J. Mol. Spect. 11, 108 (1963).

48. C. Weiss, H. Kobayashi and M. Gouterman, J. Mol. Spect. 16, 415 (1965).

49. M. Zerner, and M. Gouterman, Theoret. Chim. Acta 4, 44 (1966).

50. M. Zerner, and M. Gouterman, Inorg. Chem. 5, 1699 (1966).

51. M. Zerner and M. Gouterman, Inorg. Chem. 5, 1707 (1966).

52. M. Zerner, M. Gouterman, and H. Kobayashi, Theoret. Chim. Acta. 6, 363 (1966).

53. M. Zerner, M. Gouterman, Theoret. Chim. Acta 8, 26 (1967).

54. D. Eastwood, and M. Gouterman, J. Mol. Spect. 30, (437) (1969).

55. P. G. Seybold, and M. Gouterman, J. Mol. Spect. 31, 1 (1969).

56. R. L. Ake and M. Gouterman, Theoret. Chim. Acta 15, 20 (1969).

57. L. Edwards, D. H. Dolphin and M. Gouterman, J. Mol. Spect. 35, 90 (1970).

58. L. Edwards, D. H. Dolphin, M. Gouterman, and A. D. Adler, J. Mol. Spec 38, 16 (1971).

59. D. Eastwood, and M. Gouterman, J. Mol. Spect. 35, 359 (1970).

60. M. Gouterman, R. A. Mathies, B. E. Smith, and W. S. Caughey, J. Chem. Phys. 52, 3795 (1970).

61. R. L. Ake, and M. Gouterman, Theoret. Chim. Acta 17, 408 (1970).

62. J. B. Callis, M. Gouterman, Y. M. Jones, and B. H. Henderson, J. Mol. Spect. 39, 410 (1971).

63. A. M. Schaffer and M. Gouterman, Theoret. Chim. Acta 18, 1 (1970).

64. L. Bajema, and M. Gouterman, and C. B. Rose, J. Mol. Spect. 39, 421 (1971).

65. A. J. McHugh, M. Gouterman, and C. Weiss, Jr., Theoret. Chim. Acta 24, 346 (1972).

66. A. M. Schaffer, and M. Gouterman, Theoret. Chim. Acta 25, 62 (1972).

67. J. B. Callis, J. M. Knowles, and M. Gouterman, J. Phys. Chem. 77, 154 (1973).

68. M. Gouterman, F. P. Schwarz, P. D. Smith, and D. Dolphin, J. Chem. P 59, 676 (1973).

69. W. T. Simpson, J. Chem. Phys. 17, 1218 (1949).

70. D. M. Collins, W. R. Scheidt, and J. L. Hoard, J. Am. Chem. Soc. 94, 6689 (1972).

71. D. L. Cullen, and E. F. Meyer, private communication.

72. F. Glockling, "The Chemistry of Germanium", Academic Press-London ar New York (1969).

73. M and J Chemical Co., 372 N. Bluff St., Apt. 116-A, Joliet, Ill. 60435.

74. T.R. Janson, Private Communication.

75. M. J. Buerger, "X-ray Crystallography", John Wiley and Sons, Inc., New York-London-Sydney (1942).

76. N. F. M. Henry, H. Lipson, and W. A. Wooster, "The Interpretation c X-ray Diffraction Photographs", Mac Millan and Company, Ltd., London (1961).

77. U. W. Arndt, and B. T. M. Willis, "Single Crystal Diffractometry", Cambridge University Press (1968).

78. R. L. Vandlen, and A. Tulinsky, Acta Cryst. B27, 437 (1971).

79. P. Kirkpatrick, Rev. Sci. Instr. 10, 186 (1939).

80. H. W. Wyckoff, M. Doscher, D. Tsernoglou, T. Inagami, L. N. Johnso K. D. Hardman, N. M. Allewel, D. M. Kelly, and F. M. Richards, J. Mol. Biol. 27, 563 (1967).

81. C. T. North, D. C. Phillips, and F. S. Mathews, Acta Cryst. A24, 351 (1968).

82. "International Tables for X-ray Crystallography", Vols. I-III, The Kynoch Press, Birmingham, England (1952).

83. W. R. Busing, and H. Levy, Acta Cryst. 10, 180 (1957).

84. J. De Meulenaer, and H. Tompa, Acta Cryst. 19, 1014 (1965).

85. T. C. Furnas, "Single Crystal Orienter Instruction Manual", General Electric Company, Milwaukee (1957).

86. D. Harker, Acta Cryst. 6, 731 (1953).

87. A. J. C. Wilson, Nature 150, 152 (1942).

88. D. Harker, J. Chem. Phys. 4, 381 (1936).

89. M. M. Woolfson, "Direct Methods in Crystallography", Oxford at the Clarendon Press (1961).

90. H. Ott, Z. für Krist. 66, 136 (1928).

91. D. Harker, and J. S. Kasper, Acta Cryst. 1, 70 (1948).

92. H. Hauptman and J. Karle, "Solution of the Phase Problem I. The Centrosymmetric Crystal", A. C. A. Monograph No. 3. Brooklyn: Polycrystal Book Service (1953).

93. D. Sayre, Acta Cryst. 5, 60 (1952).

94. H. A. Hauptman, "Crystal Structure Determination", Plenum Press, New York-London (1972).

95. W. Cochran, and M. M. Woolfson, Acta Cryst. 8, 1 (1955).

96. R. E. Long, Ph.D. Thesis, University of California, Los Angeles, California (1965).

97. C. J. Brown, Acta Cryst. 21, 146 (1966).

98. J. W. Lauher, and J. A. Ibers, J. Am. Chem. Soc. 96, 4447 (1974).

99. F. R. Hopf, T. P. O'Brien, W. R. Scheidt, and D. G. Whitten, private communication.

100. C. K. Johnson, "ORTEP, a Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations", ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tenn. (1965).

101. Unpublished results of this laboratory.

102. D. F. Koenig, Acta Cryst. 18, 663 (1956).

103. I. Moustakali, and A. Tulinsky, J. Am. Chem. Soc. 95, 2066 (1973).

104. D. M. Collins, R. Countryman, and J. L. Hoard, J. Am. Chem. Soc. 94, 2066 (1972).

105. R. C. Pettersen, Acta Cryst. B25, 2527 (1969).

106. D. M. Collins, and J. L. Hoard, J. Am. Chem. Soc. 92, 3761 (1970).

107. M. J. Hamor, T. A. Hamor, and J. L. Hoard, J. Am. Chem. Soc. 86, 1938 (1964).

108. E. B. Fleischer, C. K. Miller, and L. E. Webb, J. Am. Chem. Soc. 86, 2342 (1964).

109. L. J. Radonovich, A. Bloom, and J. L. Hoard, J. Am. Chem. Soc. 94, 2073 (1972).

110. J. W. Lauher, and J. A. Ibers, J. Am. Chem. Soc. 95, 5148 (1973).

111. D. L. Cullen and E. F. Meyer, Abstracts of the Winter Meeting of the American Crystallographic Association, Gainesville, Florida, January 1973.

112. C. L. Christ, and D. Harker, Am. Mineral 27, 219 (1952).

113. R. G. Little, and J. A. Ibers, J. Am. Chem. Soc. 96, 4452 (1974).

114. R. Timkovich, and A. Tulinsky, J. Am. Chem. Soc. 91, 4430 (1969).

115. M. D. Glick, G. H. Cohen, and J. L. Hoard, J. Am. Chem. Soc. 89, 1996 (1967).

116. J. L. Hoard, G. H. Cohen, and M. D. Glick, J. Am. Chem. Soc. 89, 1992 (1967).

117. E. F. Meyer, Acta Cryst. B28, 2162 (1972).

118. J. L. Hoard, M. J. Hamor, T. A. Hamor, and W. S. Caughey, J. Am. Chem. Soc. 87, 2312 (1965).

119. E. B. Fleischer, J. Am. Chem. Soc. 85, 146 (1963).

120. J. L. Hoard, Science 174, 1295 (1971).

121. J. C. Kendrew, Science 139, 1259 (1963); M. F. Perutz, Science 140, 863 (1963).

122. A. Liljas, and M. G. Rossman, Annual Review of Biochemistry 43, 475 (1974).

123. For an idea of the progress in protein crystallography see "Recent Low Resolution Structures", in Titles of Papers, 2nd East Coast Protein Crystallography Workshop, April 21-25 (1974).

124. D. C. Phillips, Advances in Structure Research by Diffraction Methods 2, 75 (1966).

125. J. J. Ladic, Advances in Quantum Chemistry 7, 397 (1973).

126. P.-O. Löwdin, "Electronic Aspects of Biochemistry" p. 167, edited by B. Pullman, Academic Press (1963).

127. Encyclopaedia Britannica 8, 620 (1972).

128. Webster Third New International Dictionary, G and C Merrian Company (1968).

129. D. G. Smyth, W. H. Stein, and S. J. Moore, J. Biol. Chem. 238, 227 (1963).

130. C. C. F. Blake, D. F. Koening, G. A. Mair, A. C. T. North, D. C. Phillips, and V. R. Sarma, Nature 206, 757 (1965).

131. G. Kartha, J. Bello, and D. Harker, Nature 213, 862 (1967).

132. Research Groups at Rockefeller University and Merck Company.

133. B. S. Hartley, "Structures and Activity of Enzymes" edited by T. W. Groodin, J. T. Harris and B. S. Hartley, p. 47, Academic Press, London (1964).

134. B. S. Hartley, and D. L. Kauffman, Biochem. J. 101, 229 (1966).

135. B. Meloun, I. Kluh, V. Kostka, L. Moravek, J. Prusic, J. Vanecek, B. Keil, and F. Sorm, Biochem, Biophys. Acta 130, 543 (1966).

136. D. M. Blow, J. J. Birktoft, and B. S. Hartley, Nature 221, 337 (1969).

137. R. B. Corey, O. Battfay, D. A. Brueckner, and F. G. Mark, Biochim. Biophys. Acta 94, 535 (1965).

138. A. Yapel, M. Han, R. Lumry, A. Rosenberg, and D. F. Shiao, J. Am. Chem. Soc. 88, 2573 (1965).

139. H. Neurath, and G. Schwert, Chem. Rev. 46, 69 (1950).

140. M. L. Bender, and F. T. Kezdy, Ann. Rev. Biochem. 34, 49 (1968).

141. L. Cunningham, "Comprehensive Biochemistry" 16, 96 (1965).

142. G. P. Hess, "The Enzymes", 3rd ed., 3, 213 (1971).

143. G. P. Hess, J. Mc Con, E. Ku, and G. McConkey, Phil. Trans. Roy. Soc. London (B) 257, 89 (1970).

144. B. W. Matthews, P. B. Sigler, R. Henderson, and D. M. Blow, Nature 214, 652 (1967).

145. P. B. Sigler, D. M. Blow, B. M. Matthews, and R. Henderson, J. Mol. Biol. 35, 143 (1968).

146. J. J. Birktoft, D. M. Blow, R. Henderson, and T. A. Steitz, Phil. Trans. Roy. Soc. London B257, 67 (1970).

147. D. M. Blow, "The Enzymes", 3rd ed., 3 185 (1971).

148. J. J. Birktoft, and D. M. Blow, J. Mol. Biol. 68, 187 (1972).

149. A. Tulinsky, N. V. Mani, C. N. Morimoto, and R. L. Vandlen, Acta Crystallogr. B29 1309 (1973).

150. A. Tulinsky, R. L. Vandlen, C. N. Morimoto, N. V. Mani, and L. H. Wright, Biochemistry 12, 4185 (1973).

151. J. J. Birktoft, B. W. Matthews and D. M. Blow, Biochem. Biophys. Res. Comm. 36, 131 (1969).

152. P. B. Sigler, B. A. Jeffrey, B. W. Matthews, and D. M. Blow, J. Mol. Biol. 15, 175 (1966).

153. R. L. Vandlen, Ph.D. Thesis, Michigan State University (1972).

154. R. L. Vandlen, and A. Tulinsky, Biochemistry 12, 4193 (1973).

155. L. Pauling, and R. B. Corey, Proc. Natl. Acad. Sci. U.S. 37, 729 (1951).

156. A. R. Fersht, J. Mol. Biol. 64, 497 (1972).

157. N. K. Shaffer, S. C. May, and W. H. Summerson, J. Biol. Chem. 202, 67 (1953).

158. W. J. Ray, and D. E. Koshland, Brookhaven Symp. Biol. 13, 135 (1960).

159. E. B. Ong, E. Shaw, and G. Schoellman, J. Am. Chem. Soc. 86, 1271 (1964).

160. R. Lumry, and R. Biltonen, "Structure and Stability of Biological Molecules" edited by S. N. Timasheff and G. D. Fasman, p. 65, Marcel Dekker, Inc., New York (1969).

161. P. A. Kollman, and L. C. Allen, Chemical Reviews 72, No. 3, 283 (1972).

162. B. H. Havsteen, and G. P. Hess, Biochem. Biophys. Res. Commun. 14, 313 (1964).

163. B. L. Labouesse, H. L. Oppenheimer, and G. P. Hess, Biochem. Biophys. Res. Commun. 14, 318 (1964).

164. W. H. Cruickshank, and H. Kaplan, Biochem. J. 130, 1125 (1972).

165. M. L. Bender, G. E. Clement, F. J. Kezdy and H. D' A. Heck, J. Am. Chem. Soc. 86, 3680 (1964).

166. K. C. Aune, and S. N. Timasheff, Biochemistry 10, 1609 (1971).

167. B. H. Havsteen, and G. P. Hess, J. Am. Chem. Soc. 85, 791 (1963).

168. R. E. Dickerson, M. L. Kopka, C. L. Borders, J. Varnum, J. E. Weinzierl, and E. Margoliash, J. Mol. Biol. 29, 77 (1967).

169. H. R. Mahler, and E. H. Cordes, "Biological Chemistry" p. 10, Harper International Edition (1966).

170. J. Steinhardt, and E. M. Zaiser in "Advances in Protein Chemistry" 10, 151 (1955).

171. J. Steinhardt, and S. Beychok "The Proteins" 2, 139 (1964).

172. A. Mavridis, A. Tulinsky, and M. N. Liebman, Biochemistry 13, 3661 (1974).

173. J. H. Northrup, M. Kunitz, and R. Herriot, "Crystalline Enzymes", Columbia University Press, New York (1939).

174. J. D. Bernal, I. Fankuchen, and M. F. Perutz, Nature, London 141, 523 (1938).

175. R. M. Bates, "Determination of pH", 2nd Ed., Wiley and Sons, New York (1973).

176. M. V. King, Acta Cryst. 7, 601 (1954).

177. D. M. Blow, and F. H. C. Crick, Acta Cryst. 12, 794 (1971).

178. S. C. Abrahams, Phys. Today 22, No. 8, 30 (1969).

179. L. Stryer, J. C. Kendrew, and H. C. Watson, J. Mol. Biol. 8, 96 (1964).

180. D. Harker, Acta Cryst. 9, 1 (1956).

181. R. Henderson, and J. K. Moffat, Acta Cryst. B27, 1414 (1971).

182. F. M. Richards, J. Mol. Biol. 37, 225 (1968).

183. J. A. Gladner, and H. Neurath, J. Biol. Chem. 206, 911 (1954).

184. K. E. Neet, and S. E. Brydon, Arch. Biochem. Biophys. 223, 136 (1970).

185. M. W. Hunkapiller, S. H. Smallcombe, D. R. Whitaker, and J. H. Richards, J. Biol. Chem. 248, 8306 (1973).

186. Values other than pH 5.4 and 8.3, unpublished results of this laboratory.

187. A. Tulinsky, and L. H. Wright, J. Mol. Biol. 81, 47 (173).

188. V. Luzzatti, Acta Cryst. 6, 142 (1953).

189. J. R. Garel, S. Epely, and B. Labouesse, Biochemistry 13, 3117 (1974).

190. M. Shinitzky, E. Katchalski, V. Grisano, and N. Sharon, Arch. Biochem. Biophys. 116, 332 (1966).

191. S. V. Konev, "Fluorescence and Phosphorescene of Proteins and Nucleic Acids" Plenum Press, New York (1967).

192. K. L. Carraway, P. Spoerl, and D. E. Koshland, J. Mol. Biol. 42, 133 (1969).

193. J. P. Abita, M. Delaage, M. Lazdunski, and J. Savrda, Eur. J. Biochem. 8, 314 (1969).

194. A. R. Fersht, and J. Sperling, J. Mol. Biol. 74, 137 (1973).

195. J. Kraut, H. J. Wright, M. Kellerman, and S. T. Freer, Proc. Nat. Acad. Sci. U.S. 58, 304 (1967).

196. H. T. Wright, J. Kraut, and P. E. Wilcox, J. Mol. Biol. 37, 363 (1968).

197. B. W. Matthews, G. H. Cohen, E. W. Silverton, H. Braxton, and D. R. Davies, J. Mol. Biol. 36, 179 (1968).

198. G. H. Cohen, E. W. Silverton, B. W. Matthews, H. Braxton, and D. R. Davies, J. Mol. Biol. 44, 129 (1969).

199. D. R. Davies, G. H. Cohen, E. W. Silverton, H. P. Braxton, and B. W. Matthews, Acta Cryst. A25, 5182 (1969).

200. S. T. Freer, J. Kraut, J. D. Roberts, H. T. Wright and Ng. H. Xuong, Biochemistry 9, 1997 (1970).

201. H. T. Wright, J. Mol. Biol. 79, 1 (1973).

202. H. T. Wright, J. Mol. Biol. 79, 13 (1973).

203. D. M. Blow, C. S. Wright, D. Kukla, A. Rühlmann, W. Steigemann and R. Huber, J. Mol. Biol. 69, 137 (1972).

204. A. Rühlmann, D. Kukla, P. Schwager, D. Barteis, and R. Huber, J. Mol. Biol. 77, 417 (1973).

205. P.-O. Löwdin, "Advances in Quantum Chemistry" 2, 216 (1965).

206. H. Jeffreys, Proc. London Math. Soc. (2) 23, 428 (1923).

207. G. Wentzel, Z. Physik 38, 518 (1926).

208. H. A. Kramers, Z. Physik 39, 828 (1926).

209. L. Brillouin, Comp. Rend. 183, 24 (1926).

210. E. Merzbacher, "Quantum Mechanics", 2nd Edition, John Wiley and Sons, Inc. (1970).

211. D. Bohm, "Quantum Theory", Prentice Hall, Inc. (1951).

APPENDIX

The Observed and Calculated Structure Factors of Dimethoxyporphinato Ge(IV)