

This is to certify that the
thesis entitled
**THE PREPARATION AND ANALYSIS OF
AN ISOMORPHOUS PLATINUM DERIVATIVE CRYSTAL OF
BOVINE PROTHROMBIN FRAGMENT 1**
presented by

Timothy J. Rydel

has been accepted towards fulfillment
of the requirements for

M.S. degree in Chemistry



Major professor

Date November 10, 1983



RETURNING MATERIALS:
Place in book drop to
remove this checkout from
your record. FINES will
be charged if book is
returned after the date
stamped below.

JUN 25 '07 EB /

THE PREPARATION AND ANALYSIS OF
AN ISOMORPHOUS PLATINUM DERIVATIVE CRYSTAL OF
BOVINE PROTHROMBIN FRAGMENT 1

By

Timothy J. Rydel

A THESIS

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

MASTER OF SCIENCE

Department of Chemistry

1983

ABSTRACT

THE PREPARATION AND ANALYSIS OF AN ISOMORPHOUS PLATINUM DERIVATIVE CRYSTAL OF BOVINE PROTHROMBIN FRAGMENT 1

By

Timothy J. Rydel

An isomorphous heavy atom derivative crystal of bovine prothrombin fragment 1 was prepared by soaking fragment 1 crystals in a crystal storing solution made saturated with $K_2Pt(C_5H_5N)_2Cl_2$. Three-dimensional intensity data were collected to 3.5 Å resolution. The intensity data were reduced to structure factor amplitudes and a difference Patterson map was calculated using native fragment 1 and the scaled platinum derivative structure factor amplitudes. The difference Patterson map was consistent with one platinum binding site per molecule, and the fractional coordinates of the platinum position in the unit cell of the crystal are:
 $\bar{x} = 0.089 \pm 0.001$, $\bar{y} = 0.369 \pm 0.001$, and $\bar{z} = 0.101 \pm 0.001$.

In loving memory of my Grandfather,
Joseph Casimir Rydel

ACKNOWLEDGEMENTS

I wish to extend my deepest gratitude and heart-felt thanks to Dr. Alexander Tulinsky for his genuine interest, enthusiasm and encouragement throughout the course of this work. It is a great pleasure for me to be a member of his laboratory.

To Jim Gaier I extend my sincerest thanks and appreciation for taking an interest in my work and for selflessly helping me over innumerable stumbling blocks. He and Dr. Tulinsky have really instilled in me an enthusiasm for protein crystallography.

I would also like to thank Dr. Chang Park, who also works on the fragment 1 problem in this laboratory, for his interest and guidance in my work, and Dr. James P. Fillers, Dr. Barb Karcher, and Dr. Ewa Skrzypczak-Jankun for their assistance, discussions, and encouragement.

I am grateful to Dr. Gary L. Nelsestuen of the University of Minnesota-St. Paul for supplying us with generous amounts of lyophilized fragment 1 and lyophilized deglycosylated fragment 1 for protein crystallization experiments.

I am also lucky to have had so many supportive family members and friends behind me in this effort. To them I

extend my love and thanks. I especially want to thank my parents, my brother and sisters and my Grandma who have been behind me always and who have my unending affection and gratitude.

To my friend Margaret Lynch I extend sincere thanks for kindly and expertly typing this manuscript on such short notice.

Support by the National Institutes of Health is gratefully acknowledged.

TABLE OF CONTENTS

CHAPTER	PAGE
LIST OF TABLES.	vi
LIST OF FIGURES.	viii
INTRODUCTION.	1
1 BLOOD COAGULATION AND FRAGMENT 1.	3
A. The Hemostatic Response.	3
B. The Activation of Prothrombin and the Significance of Studying Fragment 1.	5
2 THE CRYSTALLIZATION OF PROTEINS.	17
3 THE PREPARATION OF FRAGMENT 1 CRYSTALS.	25
4 MOUNTING A FRAGMENT 1 CRYSTAL FOR X-RAY STUDIES..	29
5 THE PREPARATION OF ISOMORPHOUS HEAVY ATOM DERIVATIVE CRYSTALS.	35
A. Screening for Possible Isomorphous Heavy Atom Derivatives of Fragment 1.	42
B. Promising Platinum Derivatives of Fragment 1.	43
6 THE 3.5 Å RESOLUTION INTENSITY DATA COLLECTION OF THE F1PT2 FRAGMENT 1 CRYSTAL.	53
7 PROCESSING THE F1PT2 DATA.	66
A. Converting the Intensity Data to Structure Amplitudes.	66
B. The F1PT2-F1N6 Difference Patterson Vector Map.	76
CONCLUSIONS.	89
LIST OF REFERENCES.	93

LIST OF TABLES

TABLE		PAGE
1	Precipitating Agents Used in Protein Crystallizations.18
2	Common Hard Acids Used in Protein Crystallizations.39
3	Common Soft Acids Used in Protein Crystallizations.40
4	Hard Acid Heavy Atom Derivative Trials.44
5	Soft Acid Heavy Atom Derivative Trials.45
6	Principal Axial Intensity Distribution Differences Between Fragment 1 + 10 mM Pt(NH ₃) ₂ (NO ₂) ₂ Br ₂ and Native Fragment 1.	51
7	Principal Axial Intensity Distribution Differences Between Fragment 1 + Saturated K ₂ Pt(C ₅ H ₅ N) ₂ Cl ₂ and Native Fragment 1.	52
8	The Average Cell Parameters of FlN6 (Native Fragment 1) and of FlPT2.58
9	The Number of Reflections in Each Data Collection 2θ Shell.	58
10	The Course of Events in the FlPT2 Data Collection.65
11	Absorption Table Used in the FlPT2 Data Collection.69
12	Decay Correction "S" Factors for the Three Monitor Reflections.	71
13	Estimate of I°(t)/I(t) for Low Angle Decay Corrections.	71
14	Calculating the FlN6-FlPT2 Scale Factor for the P-DATA Program.72
15	Averaging the Background of the FlPT2 Intensity Data in 2θ Shells.	74
16	Results of Averaging the Background of the FlPT2 Data with Respect to 2θ and φ.	77

TABLE		PAGE
17	The Unique Vectors Relating Equivalent Positions in the Unit Cell of the F1PT2-F1N6 Difference Patterson.	84

LIST OF FIGURES

FIGURE		PAGE
1	The Hemostatic Response.	2
2	The Blood Clot.6
3	"Prothrombinase Complex".7
4	The Products of PT Activation.	8
5	(A) Fragment 1 and (B) Fragment 2.10
6	The Conformational Change due to Calcium Binding and the Calcium Mediated PT-PL Interaction.13
7	The Sugar Chains of PT.	14
8	The Sugar Moieties of PT after Sialidase Treatment. (N-Acetyl Neuraminic Acid Removal)..16
9	A Typical Protein Solubility Curve.	20
10	The Vapor Diffusion Crystallization Technique Using the (A) Hanging Drop and (B) Standing Drop Methods.	23
11	Experimental Vapor Diffusion Set-up for Growing Large Fragment 1 Crystals.26
12	Crystals of Fragment 1.	27
13	Morphology of an Idealized Fragment 1 Crystal.30
14	Stages in the Mounting of a Fragment 1 Crystal.32
15	The Principal Axial Intensity Distributions of Native Fragment 1.	34
16	A Phase Circle Diagram of a Native Protein and an Isomorphous Heavy Atom Derivative.36
17	A Phase Circle Diagram of a Native Protein and Two Isomorphous Heavy Atom Derivatives.38
18	Principal Axial Intensity Distributions of (A) 10 mM Pt(NH ₃) ₂ (NO ₂) ₂ Br ₂ and (B) Native Fragment 1.	49

FIGURE	PAGE
19	Principal Axial Intensity Distributions of (A) Saturated $K_2Pt(C_5H_5N)_2Cl_2$ and (B) Native Fragment 1. 50
20	The Goniostat of a Four-Circle Diffractometer. .54
21	The Intensity of the FlPT2 Monitor Reflections vs. Time. 60
22	ω -Profile of (7,7,13) Taken Before the 3-D FlPT2 Data Collection. 61
23	Intensity Absorption Plots of (0,0,4) and (0,0,20) Before and After the FlPT2 Intensity Data Collection. 63.
24	Absorption Curves of (0,0,4) and (0,0,20) Before and After the FlPT2 Intensity Data Collection. 68
25	The Average Background of the FlPT2 Intensity Data vs. 2θ 75
26	The $ F ^2$ Distribution for FlN6 ($B=0$, $K=1$) and for FlPT2 ($B=4.36$, $K=1.3$). 81
27	$w = \frac{1}{4}$ Harker Section of the FlPT2-FlN6 Difference Patterson Map ('X' indicates Harker vector). 86
28	$w = \frac{1}{2}$ Harker Section of FlPT2-FlN6 Difference Patterson Map ('X' indicates Harker vector). . .87
29	$u = \frac{1}{2}$ or $v = \frac{1}{2}$ Harker Section of FlPT2-FlN6 Difference Patterson Map ('X' indicates Harker vector). 88

INTRODUCTION

The primary concern of an x-ray crystallographic investigation of a protein is to generate an electron density map, $\rho(x,y,z)$, of the molecule. This density, $\rho(x,y,z)$, is calculated via the following Fourier summation¹:

$$\rho(x,y,z) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} |F(hkl)| \exp i\alpha(hkl) \exp \{-2\pi i(hx+ky+lz)\} \quad (1)$$

where $|F(hkl)|$ is the amplitude of a reflection (hkl) , $\alpha(hkl)$ is the phase of the reflection, V is the volume of the unit cell, and the summations extend over all the observable reflections. The major obstacle to calculating this density map is the determination of the phases, $\alpha(hkl)$, which cannot be measured. However, this phase information can be obtained by locating the binding sites of a heavy atom compound within two or more isomorphous derivative crystals and then using the multiple isomorphous replacement method^{2,3}. These crystals have essentially the same unit cell, space group, and molecular structure as the parent crystals but have one or more heavy atoms introduced at specific loci. This thesis documents the results of a search for suitable isomorphous heavy atom

derivatives of bovine prothrombin fragment 1 (hereafter referred to as fragment 1), a cleavage product of prothrombin (hereafter referred to as PT) arising in blood coagulation.

CHAPTER 1

BLOOD COAGULATION AND FRAGMENT 1

A. The Hemostatic Response

The coagulation of blood results from the detailed interaction of the blood vessel, the blood platelet, and the blood coagulation factors found in the plasma. The interaction of these three systems⁴ is represented in Figure 1. The blood coagulation response is initiated when the blood vessel wall is punctured. The puncture exposes subendothelial components such as collagen and the von Willebrand factor to the plasma⁴. When a blood platelet contacts the collagen surface in the presence of the von Willebrand factor, the platelet releases a variety of components⁵—most notably ADP—which stimulates the flattening of the platelet at the vessel opening. The surface puncture also activates the molecular cascade of glycoproteins⁶ by activating factor XII to factor XIIa (Figure 1). The penultimate step of this molecular cascade is the conversion of PT to thrombin; thrombin then converts fibrinogen (factor I) to fibrin (Figure 1).

In the presence of the ADP, collagen, and thrombin, the platelets aggregate to form a plug at the blood vessel

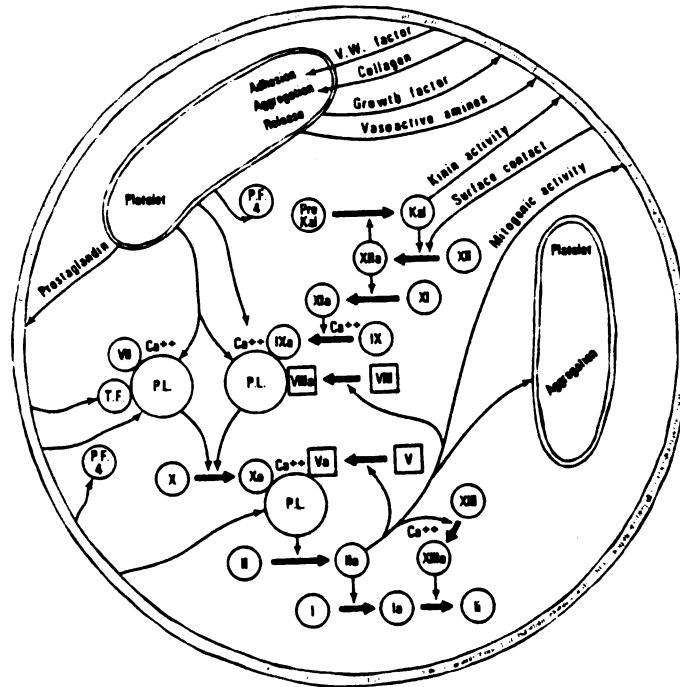


Figure 1. The Hemostatic Response. Phospholipid = P.L., Kallikrein = Kal, platelet factor 4 = P.F.4, tissue factor = T.F., prothrombin (PT) = II, thrombin = IIa, fibrinogen = I, fibrin = Ia, cross-linked fibrin = Ii; (taken from reference 4).

surface with the fibrin strands reinforcing it. The appearance of this plug⁵ is represented in Figure 2. Red blood cells are often trapped in this clot. Within a few hours to a few days factors within the cell will convert the zymogen plasminogen to plasmin, which dissolves the fibrin and the clot⁷.

B. The Activation of Prothrombin and the Significance of Studying Fragment 1

One of the key steps in the blood coagulation cascade is the conversion of PT to thrombin. In this step, PT, a zymogen of molecular weight 74,000, is activated through a "prothrombinase" complex¹⁰ to form thrombin. The complex, as Figure 3 shows, contains the proteolytic enzyme factor Xa (factor X is a zymogen, factor Xa is the active enzyme), the cofactor factor Va, the phospholipid (hereafter referred to as PL) surface of platelets and Ca⁺² ions. A model representation of the activation components^{11,12} is shown in Figure 4. Factor Va, bound to the PL surface of the platelet, contains the binding site for Pt; the domain of PT which ultimately becomes PT fragment 2 (hereafter referred to as fragment 2) binds to factor Va (Figure 3). Factor Xa cleaves PT at sites "b" and "c", which leads to the production of PT fragment 1-2 (fragment 1-2) and the active enzyme thrombin (Figure 4). In the human hemostatic response, an additional cleavage of thrombin occurs at amino acid 323, b' (Figure 4). The

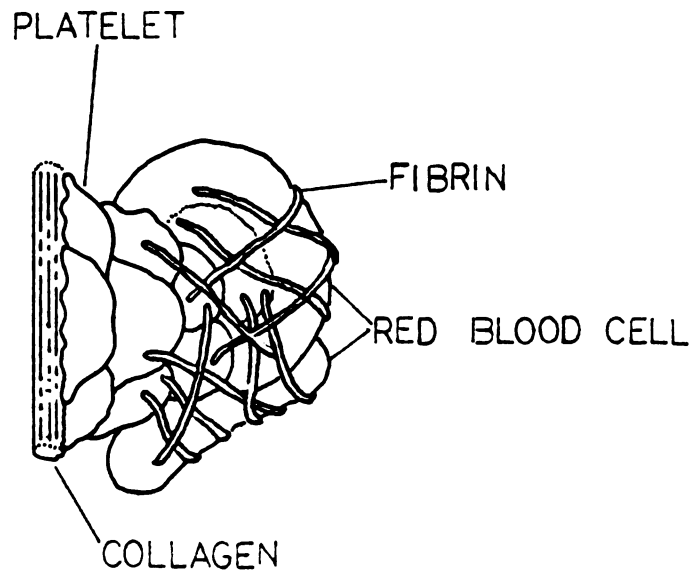


Figure 2. The Blood Clot.

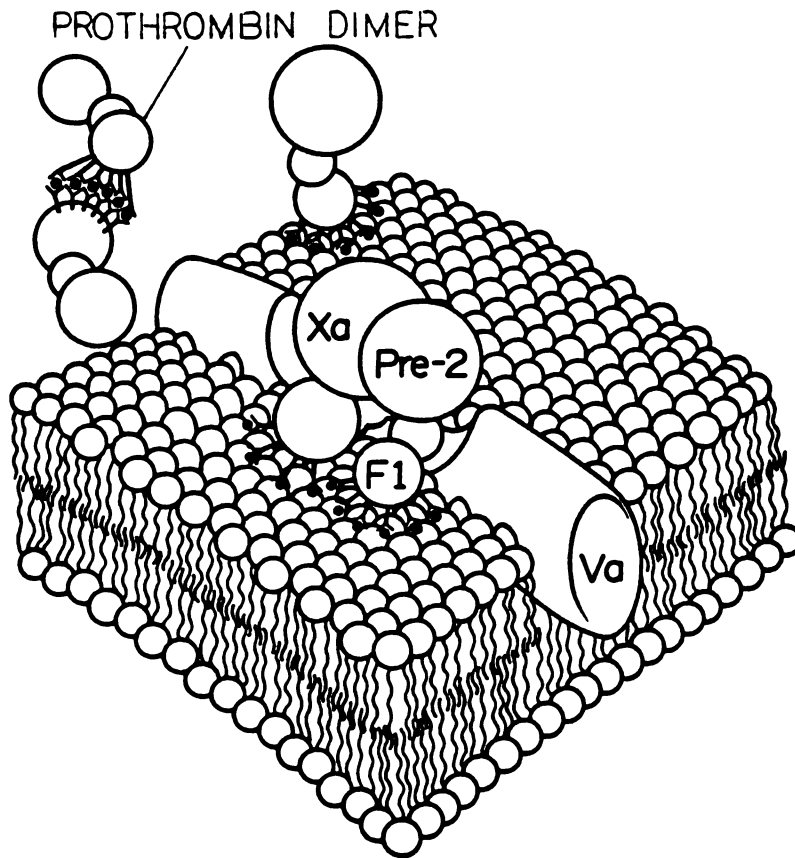


Figure 3. "Prothrombinase Complex".
PT = Pre-2+F2+F1, where Pre-2 = prethrombin 2,
F2 = fragment 2, and F1 = fragment 1; Ca^{+2}
ions are represented by filled circles
binding γ -carboxyglutamic acid residues
(Gla); Va = factor Va; Xa = factor Xa;
see text for further details.

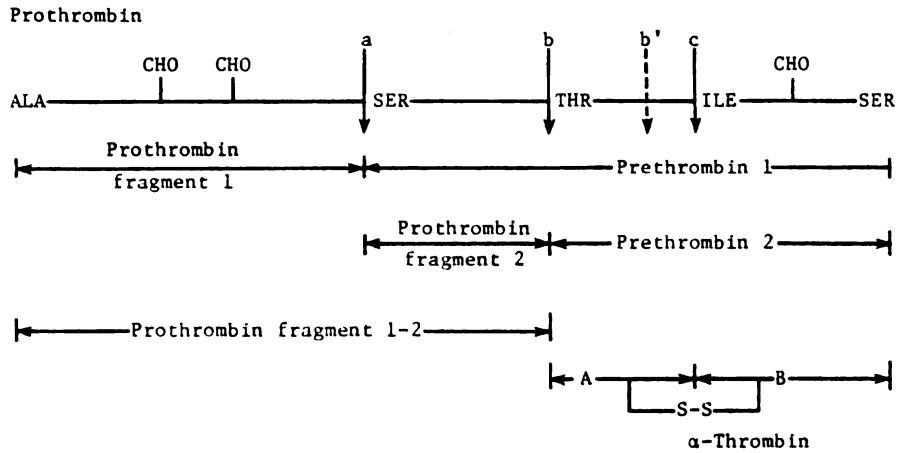


Figure 4. The products of PT Activation. CHO represents a carbohydrate side chain; a, b, and c represent the sites of cleavage of bovine and human prothrombin; an additional cleavage, b' occurs in human prothrombin.

active thrombin can also cleave fragment 1-2 at site "a" to produce fragment 1 (Figure 4), the molecule studied in this investigation (see also Figure 3).

Although fragment 1 (MW 23,000) is simply a by-product of the molecular clotting cascade, understanding the molecular structure of this glycoprotein is of the utmost importance for several reasons. First, there is an extraordinarily high degree of sequence homology between fragment 1 and other glycoproteins involved in the blood coagulation cascade. When the primary sequence of PT was determined completely by Magnusson et al.¹³, it was immediately obvious that there was sequence duplication between fragment 2 and the last 100 residues of fragment 1. This striking sequence homology can be seen in Figure 5. Magnusson named the characteristic sequences "kringle" structures (after the shape of a Danish pastry). The kringle loop also appears five times in plasminogen⁷, the zymogen converted to plasmin (which in turn dissolves fibrin and the clot), and once in urokinase^{13a}, the enzyme which converts plasmin to plasminogen. In plasminogen, all five kringles are found in the "pro" fragments rather than in what ultimately is the active enzyme^{15,16}, as in the case with PT. The sequence homology between the kringle structures of plasminogen are even more strikingly similar than that between fragment 1 and fragment 2. The presence of the kringle in PT, plasminogen, and urokinase clearly suggests that

there must be some overriding structural significance to the kringle fold. Thus, an x-ray crystallographic structure determination of fragment 1 will elucidate i) the unique folding features of kringle structures, and possibly ii) the manner in which these features translate into a biological function.

The N-terminal region of fragment 1 also possesses a great deal of sequence homology to the N-terminal regions⁷ of factors VII, IX, X, and protein C. The amino terminus of all of the aforementioned proteins contains in every case γ -carboxyglutamic acid residues (Gla)¹⁷⁻²⁰, and in all cases Gla residues in these proteins have been found to be responsible for binding Ca^{+2} ions at the PL surface. The Gla residue results from a post-ribosomal modification of the glutamic acid (Glu) residues of these proteins. The modification is mediated by vitamin K^{21,22}, and for this reason PT, factor VII, factor IX, factor X and protein C are all "vitamin K" related proteins. The sequence homology in the Gla-rich N-terminal regions of these proteins suggests that they all probably share a similar Ca^{+2} -PL binding mode. Moreover, an x-ray study of fragment 1 should reveal important structural features of the Gla residues as they might pertain to Ca^{+2} -PL binding.

A three-dimensional structural study of fragment 1 will also reveal the functionality of the two different types of Gla residues present. Studies by Nelsestuen²³, and Prendergast and Mann²⁵ have shown that one type of Gla

residue induces a protein conformation change (half-life ~100 min. at 0°C). These Gla residues bind 3 or 4 Ca⁺² ions per protein molecule and display low cation specificity. In addition to binding Ca⁺², these residues will bind Mg⁺², Mn⁺², Sr⁺², and Ba⁺² and the trivalent cations La⁺³ and Gd⁺³. Another type of Gla residue is involved in the actual protein-PL binding. These residues are highly specific for Ca⁺². For instance, Mg⁺², which is highly effective in causing a PT conformational change, is totally ineffective at inducing protein-PL binding. These studies as well as fluorescence and cyclic dichroism²⁶, immunological studies²⁷, and NMR work^{28,29} are consistent with the schematic model shown in Figure 6.

Finally, fragment 1, like all the other proteins in the blood coagulation cascade, contains carbohydrate moieties, and a study of their intact structure might reveal why these residues are of importance to biological activity. Studies of serum glycoprotein clearance in the liver³⁰, for example, have shown that the carbohydrate moieties of glycoproteins are signals in cellular recognition processes. From Figure 4, it is evident that there are three carbohydrate chains in PT; all of these sugars are linked to asparagine (Asn) residues —Asn-77 (in fragment 1), Asn-101 (in fragment 1), and Asn-376 (in thrombin). The structures of these sugar chains in PT were characterized by Mizuochi, et al.³¹ and they are shown in Figure 7. The sugar chains A and B are

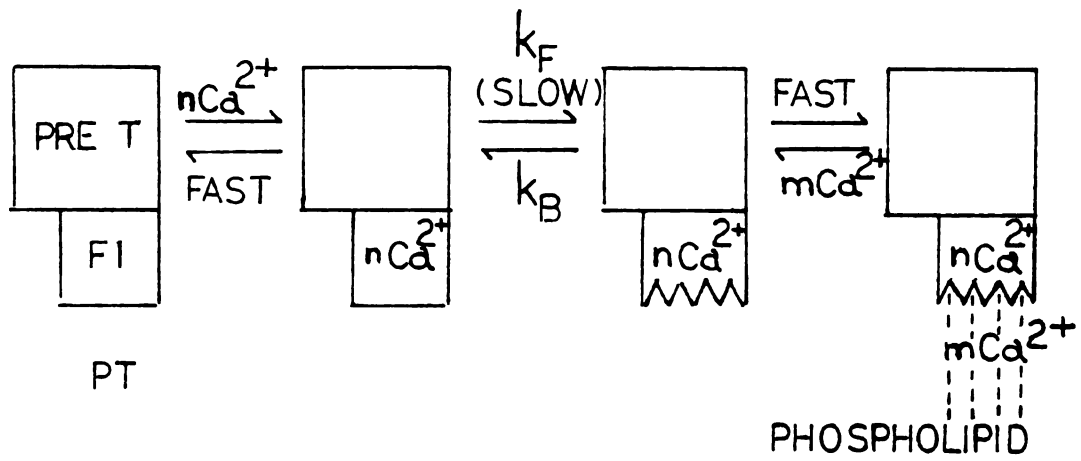


Figure 6. The Conformation Change, due to Calcium Binding and the PT-PL Interaction.

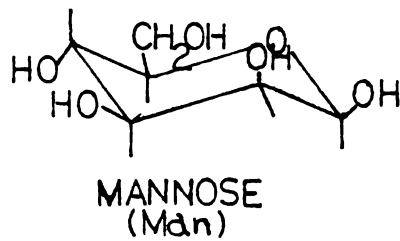
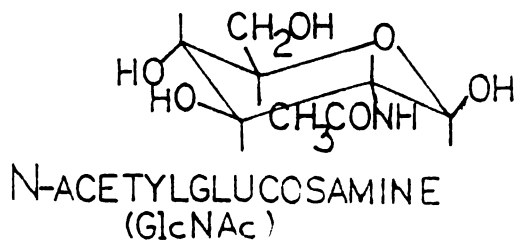
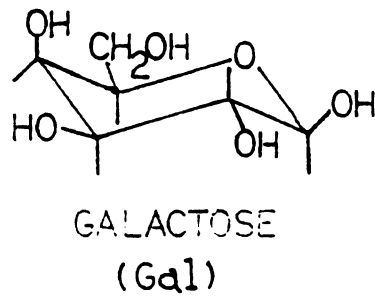
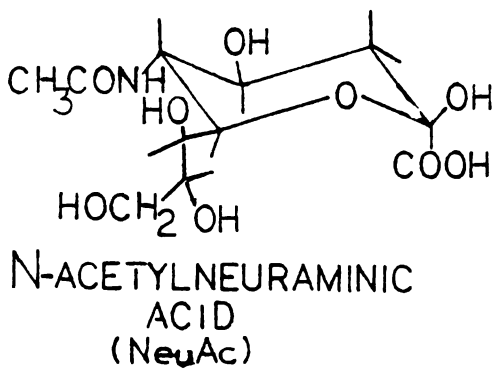
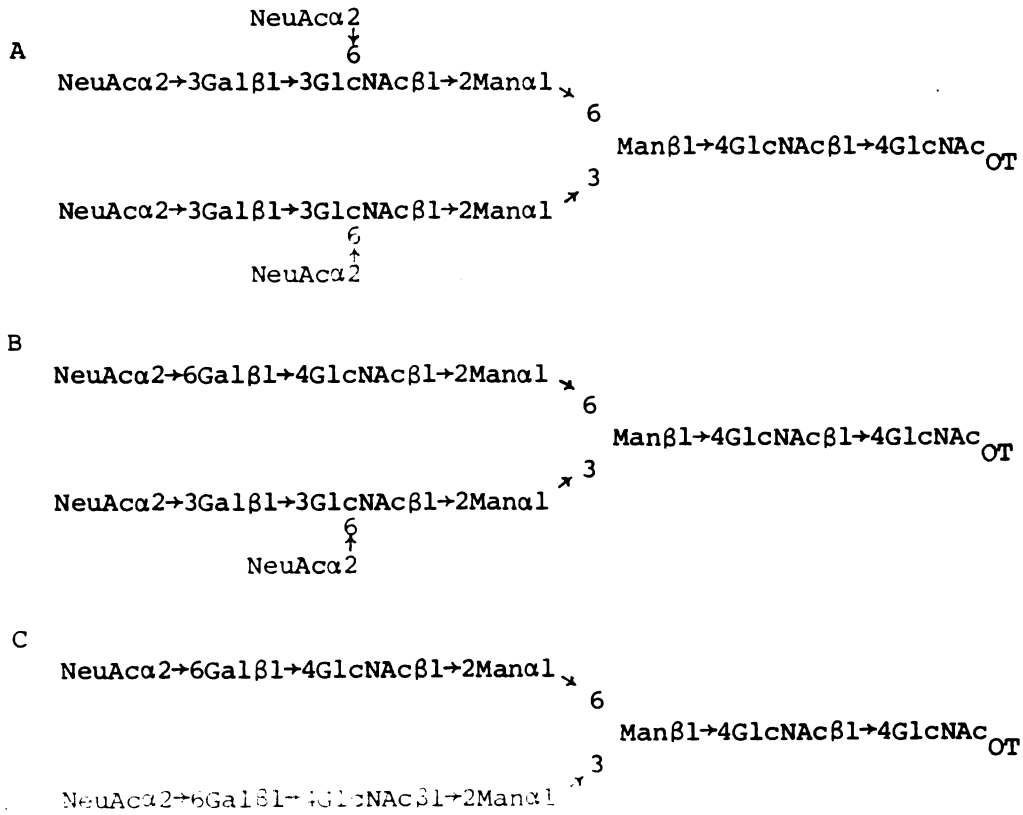
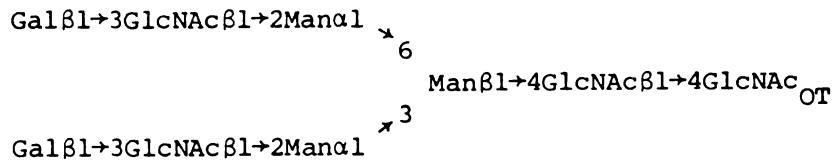


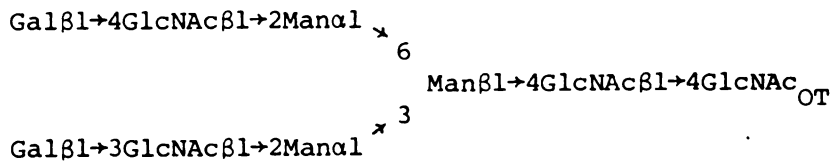
Figure 7. The Sugar Chains of PT.

found in fragment 1 and the sugar chain C is found in thrombin. Sialidase treatment of these sugar chains [removal of the N-Acyl derivatives of neuraminic acid (Figure 7)] yields the oligosacchrides N-1, N-2, and N-3, which are shown in Figure 8. While N-3 is typical of Asn-linked sugars found in other glycoproteins, the Gal β 1 \rightarrow 3GlcNac grouping in the outer portion of N-1 and N-2 is unique to glycoproteins. Due to the unique structures of these sugars, their contribution to the function of PT and fragment 1 must be considered significant.

N-1



N-2



N-3

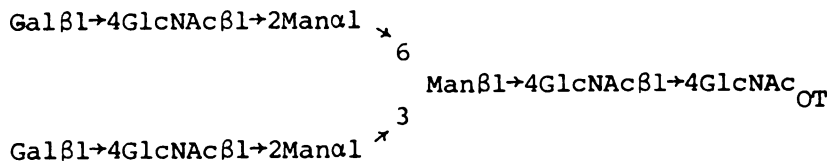


Figure 8. The Sugar Moieties of PT after Sialidase Treatment. (N-Acetyl Neuraminic Acid Removal).

CHAPTER 2

THE CRYSTALLIZATION OF PROTEINS

A protein crystallization is generally achieved by slowly approaching macromolecular insolubility with the aid of a precipitating agent³⁶. The regular packing of molecules in a crystal lattice diminishes the entropy of the system, which is energetically unfavorable, but it also minimizes the free energy which is energetically favorable. It is this free energy minimization which is the driving force behind the formation of crystals or amorphous precipitates. Amorphous precipitates result when the system only attains a local free energy minimum, and thus generally occurs when the molecules are forced out of solution quickly.

The most important parameters in a protein crystallization system are the nature and concentration of the precipitating agent, the pH and buffer system, the protein concentration, and the temperature, usually in descending order.

The three major types of precipitating agents are salts, organic solvents, and polyethyleneglycols (PEG)³². Lists of these precipitating agents³² are shown in Table 1.

TABLE 1. Precipitating Agents Used in Protein Crystallizations

A. Organic Solvents Used in Crystallization

Ethanol	Acetonitrile
Isopropanol	Dimethyl sulfoxide
2-Methyl-2,4-pentanediol (MPD)	2,5-Hexanediol
Dioxane	Methanol
Acetone	1,3-Propanediol
Butanol	1,3-Butyrolactone

B. Salts Used in Crystallization

Ammonium or sodium sulfate
Lithium sulfate
Lithium chloride
Sodium or ammonium citrate
Sodium or potassium phosphate
Sodium or potassium or ammonium chloride
Sodium or ammonium acetate
Magnesium sulfate
Cetyltrimethyl ammonium salts
Calcium chloride
Ammonium nitrate
Sodium formate

C. PEG* Used in Crystallization

PEG 1,000
PEG 4,000
PEG 6,000
PEG 20,000

*The number given is the average molecular weight of the polyethyleneglycol.

As can be seen from Figure 9, the solubility of a protein is logarithmic and decreases exponentially as the ionic strength is increased. The decrease in solubility on either side of the solubility maximum can be used to obtain crystals. Crystallization at the low ionic strength side is known as salting-in and occurs when organic solvents are used³² (Table 1A). The most popular organic solvents of those listed in Table 1A are ethanol and MPD. Salting in solvents decrease the dielectric constant of the medium; this decreases the ionization of charged protein groups and forces molecules to compensate for their local charges with the opposite charges of a neighbor. Salting-out precipitants are salts (Table 1B) of which ammonium sulfate is most popular³². They monopolize water molecules which would otherwise hydrate the protein. When the concentration of these salt ions becomes too high, macromolecules are forced to neutralize their surface charges by binding to one another. PEG shares properties of both salting-in and salting-out precipitants³²—like organic solvents they decrease the dielectric constant, and like salts they monopolize water molecules.

Another important variable in a protein crystallization is pH. In fact, certain pH ranges may destroy the integrity or activity of the macromolecules. Moreover, a seemingly slight change in pH can make a difference between growing or not growing crystals. Fragment 1, for example, does not yield crystals at pH 6.0, but nice hexagonal crystals are

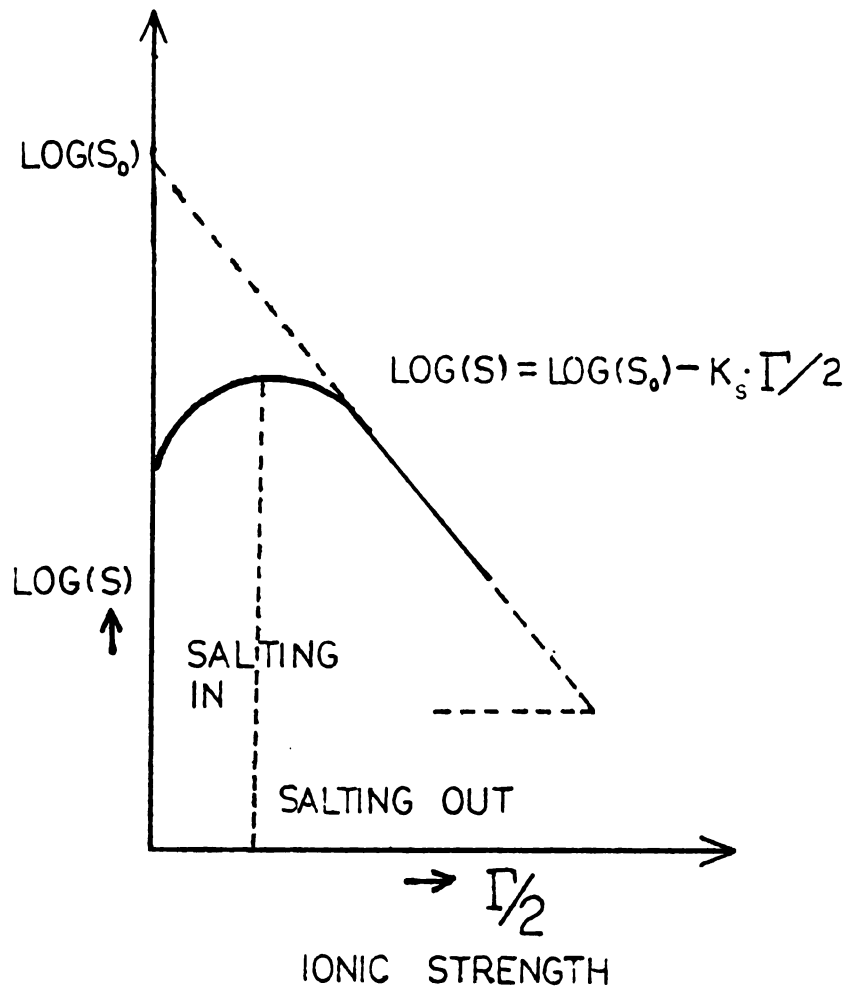


Figure 9. A Typical Protein Solubility Curve.

obtained at pH 6.5. The pH range at which crystals are obtained limits the choice of possible buffer systems since the buffer present should offer the protein maximum protection from pH fluctuation.

Protein concentration is an important variable in a crystallization experiment because it can make the difference between small microcrystals ($0.1 \text{ mm} \times 0.1 \text{ mm} \times 0.1 \text{ mm} \sim 1 \times 10^{-3} \text{ mm}^3$) and large ones ($1.5 \text{ mm} \times 0.6 \text{ mm} \times 0.4 \text{ mm} \sim 0.36 \text{ mm}^3$) under otherwise similar conditions. Protein crystallizations in the literature have reported crystals grown at concentrations from one to several hundred milligrams per milliliter; however the most common range is from 5 to 30 mg/ml³³. Crystallization experiments conducted at less than 1 or 2 mg/ml usually do not work. Large crystals of fragment 1 ($\sim 1.2 \text{ mm} \times 0.6 \text{ mm} \times 0.3 \text{ mm}$) were not obtained until the protein concentration was changed from 10 mg/ml to 25 mg/ml.

Although temperature is certainly important to a protein crystallization it is the least important of the factors mentioned since it is difficult to vary. Although crystallizations have been reported for proteins over the complete range from 0 to 40°C, most work³³ is performed at either 4°C in a cold room or at room temperature, 25°C

There are 3 major protein crystallization set-up types—batch, dialysis, and vapor diffusion. In a batch set-up all of the components of the crystallization are present in the same container. α -Chymotrypsin³⁴ and

ribonuclease T₁³⁵ were grown by the batch method. In a dialysis crystallization the protein solution is placed inside a semi-permeable membrane and the macromolecule is brought slowly to the precipitation point due to the influx of the precipitant which is outside the membrane. As the differential concentration inside and outside the membrane decreases, the approach to equilibrium decreases. Bulk dialysis techniques require milliliter quantities of material which renders them impractical for many proteins. Zeppezauer and his co-workers³⁶ developed cells with which dialysis experiments can be carried out on the 20 to 50 μ l scale; this method proved to be quite useful for KDPG aldolase³⁷. The most common method used to grow protein crystals at the present-time employs the vapor diffusion technique. The vapor diffusion technique employing both hanging drops³⁸ and standing drops³⁹ is shown in Figure 10. In this method, a drop of protein solution containing precipitant (if the precipitant is non-volatile) is mounted away from a reservoir containing a more concentrated precipitant solution. Water slowly distills from the drop into the reservoir and the effect is slow crystallization or precipitation. The fragment 1 crystals employed in this study were grown using this technique⁴⁰.

Finding the proper combination of crystallization parameters and set-up which will yield x-ray data-collection quality crystals is generally a slow, empirical process.

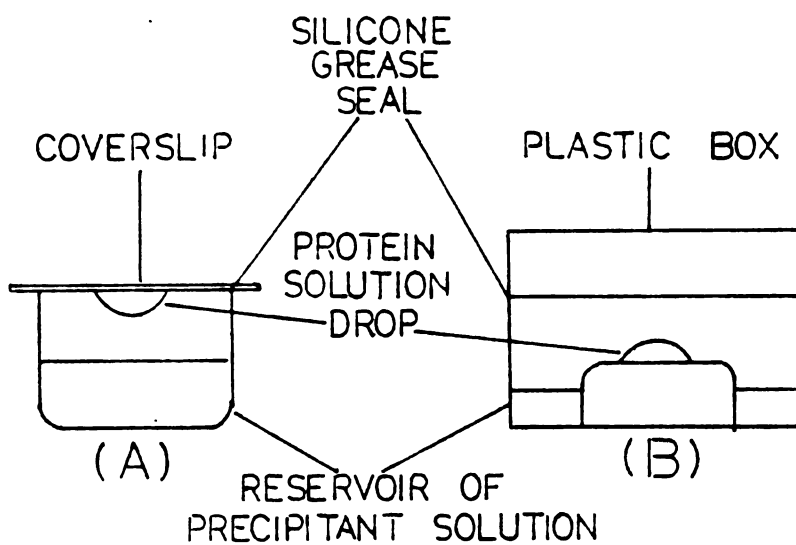


Figure 10. The Vapor Diffusion Crystallization Technique Using the (A) Hanging Drop and (B) Standing Drop Methods.

The amount of protein available, and the stability of the protein under certain pH ranges may help rule out set-up and parameter possibilities, but basically a protein is a puzzle with respect to employable crystallization conditions.

CHAPTER 3

THE PREPARATION OF FRAGMENT 1 CRYSTALS

Crystals of fragment 1 have been obtained by Aschaffenburg et al.,⁴¹ Olsson et al.,⁴² and Kung, Tulinsky and Nelsestuen⁴⁰. The fragment 1 crystals employed in this study were prepared by modifying the third group's procedure. Lyophilized fragment 1 was generously supplied to us by Dr. G.L. Nelsestuen, Professor of Biochemistry at the University of Minnesota-St. Paul. The experimental set-up used to grow the crystals is shown in Figure 11.

Fragment 1 Crystallization Procedure

The seal between the lid and the bottom of a clean 1" x 2" x 0.7" plastic box is coated with some silicone grease (Dow Corning high vacuum grease). A clean siliconized glass vial (0.7" x 0.7" x 0.4") is placed at the bottom of the box with the open end up. Two ml of the well solution is then added by pipette. The well solution is 32 (w/v)% PEG 4000 with 0.1 M sodium phosphate buffer⁴³ at pH 6.5. A protein solution (51.6 mg/ml) is prepared by dissolving fragment 1 in 0.1 M sodium phosphate buffer. An 18 μ l drop of the well solution is placed in the bottom of the vial and 25 μ l of the protein solution is added with some slight mixing. The resulting drop is 30 mg/ml in protein concentration. The lid is then placed on the box and the box is sealed with parafilm. The box is then left undisturbed for two weeks at which time large crystals, as those shown in Figure 12, usually are present. The crystal system of fragment 1 is tetragonal, $\bar{a} = \bar{b} = 77.56 (\pm 0.04) \text{ \AA}$, $\bar{c} = 85.23 (\pm 0.05) \text{ \AA}$, space group $P4_12_12$ or its enantiomorph, with 8 molecules/unit cell (1 molecule/asymmetric unit).

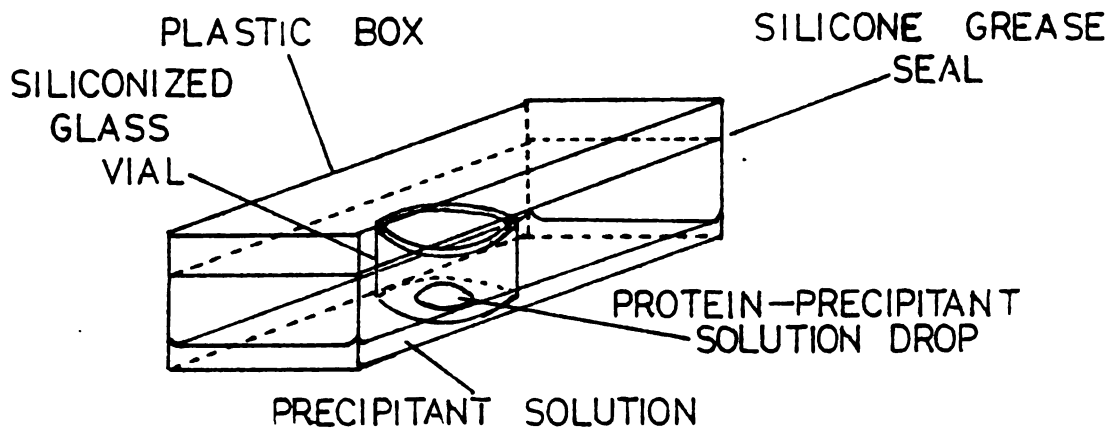


Figure 11. Experimental Vapor Diffusion Set-Up for Growing Large Fragment 1 Crystals.

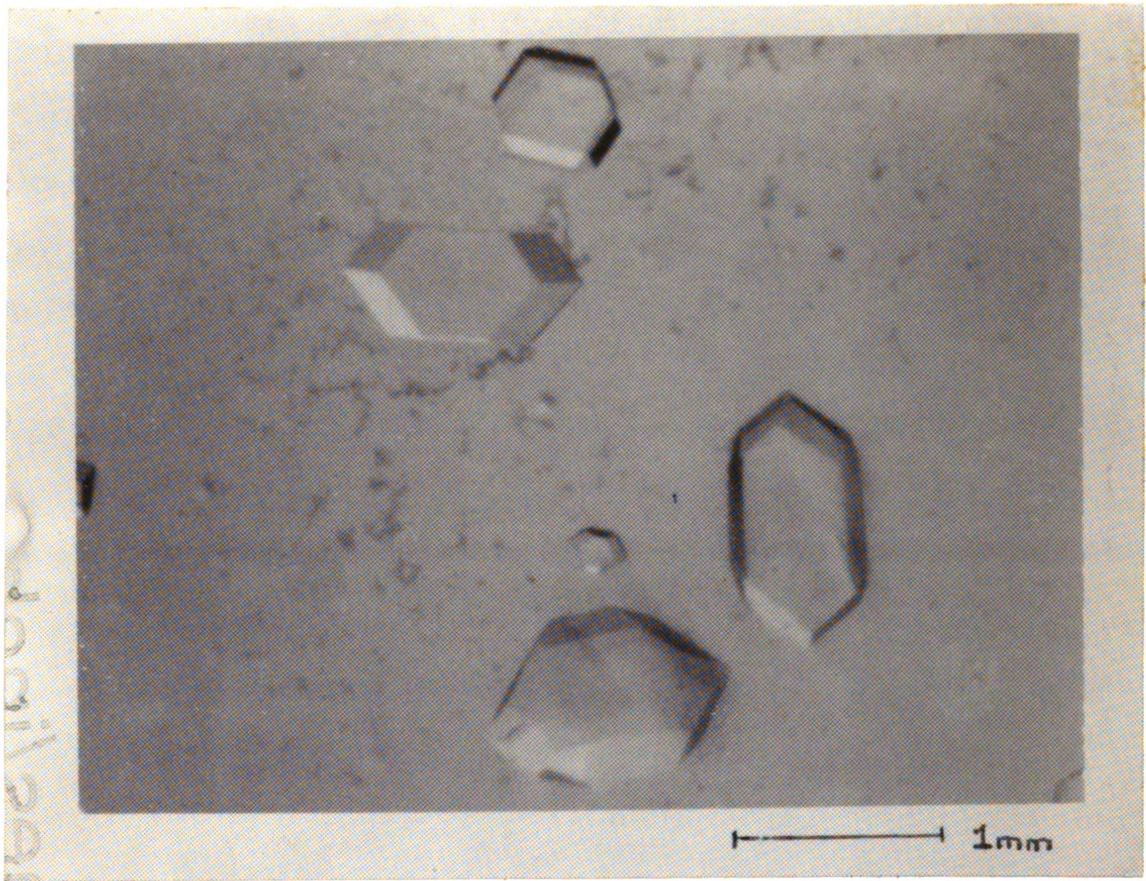


Figure 12. Crystals of Fragment 1.

One of the main objects of the structural investigation of fragment 1 is to conduct Ca^{+2} ion binding studies. Since mono- and dibasic phosphate will form insoluble calcium phosphates in the presence of Ca^{+2} ions, the phosphate-PEG mother liquor (growing solution) had to be exchanged with another solution which would maintain crystal stability and permit diffusion of Ca^{+2} ions into the crystals. It was also important at this point to choose a suitable pH for the native crystal form. Previous flow cell experiments by Kung and Tulinsky⁴⁴ established that the fragment 1 diffraction pattern showed small changes in intensity between pH 6-7 but was fairly constant thereafter. With this in mind and the fact that crystals dissolved when stored in PEG solutions less concentrated than those they were grown in, we chose 39(w/v)% PEG 4000 in 0.1 M tris-maleate buffer at pH 7.5 as our crystal storage solution for the native crystals. The procedure used to bring the fragment 1 crystals to this harvesting solution state is as follows:

Fragment 1 Harvesting Procedure

The drop size of the approximately 40 μl drop containing protein crystals is gradually increased to 640 μl . Initially, 100 μl of the well solution is carefully added to the drop. Six to ten hour time periods separate a 200 μl and later a 300 μl addition of the growing solution. Next, the well solution in the plastic crystallization boxes is replaced with the crystal storage solution, the boxes are resealed and the crystals are left undisturbed for 3-4 days. The growing solution is then exchanged with the crystal storage solution. This is done by removing a small amount of the growing solution and replacing it with the same amount of crystal storage solution. The initial solution exchange is 100 μl , followed by

200 μ l, 400 μ l, and complete volume exchanges, respectively, at 6-10 hour intervals between each. The crystals are now at their native state and available for use.

CHAPTER 4

MOUNTING A FRAGMENT 1 CRYSTAL FOR X-RAY STUDIES

Since protein crystals are a two phase system—protein and mother liquor—one should not mount a protein crystal in a conventional way. Thus, protein crystals are sealed in a capillary tube in the presence of mother liquor. The crystal is usually mounted in a morphologically and crystallographically significant manner so that important families of reflections [i.e. (hh0), (h00), (001)] can be located efficiently.

The idealized morphology of a fragment 1 crystal is shown in Figure 13. The procedure for mounting a fragment 1 crystal for data collection follows. It is a modification of a method proposed by King⁴⁵.

The Mounting of a Fragment 1 Crystal:

- a) Clean and dry a 1.5 mm diameter × 80 mm glass x-ray capillary tube (distributed by the Charles Supper Company). Shorten the tube from the sealed end side by about 30 mm to yield a 50 mm sealed tube. Prepare a clay support for the tube by placing 2 pieces of clay (to support both ends of the tube) on a glass slide. Rest the tube on the clay support (Figure 14(a)).
- b) Start to fill the capillary tube with crystal storing solution approximately 5 mm from the sealed end and add enough solution until the liquid reaches the enlarged open end. This is done by means of a narrow capillary tube connected to rubber tubing that has a mouthpiece

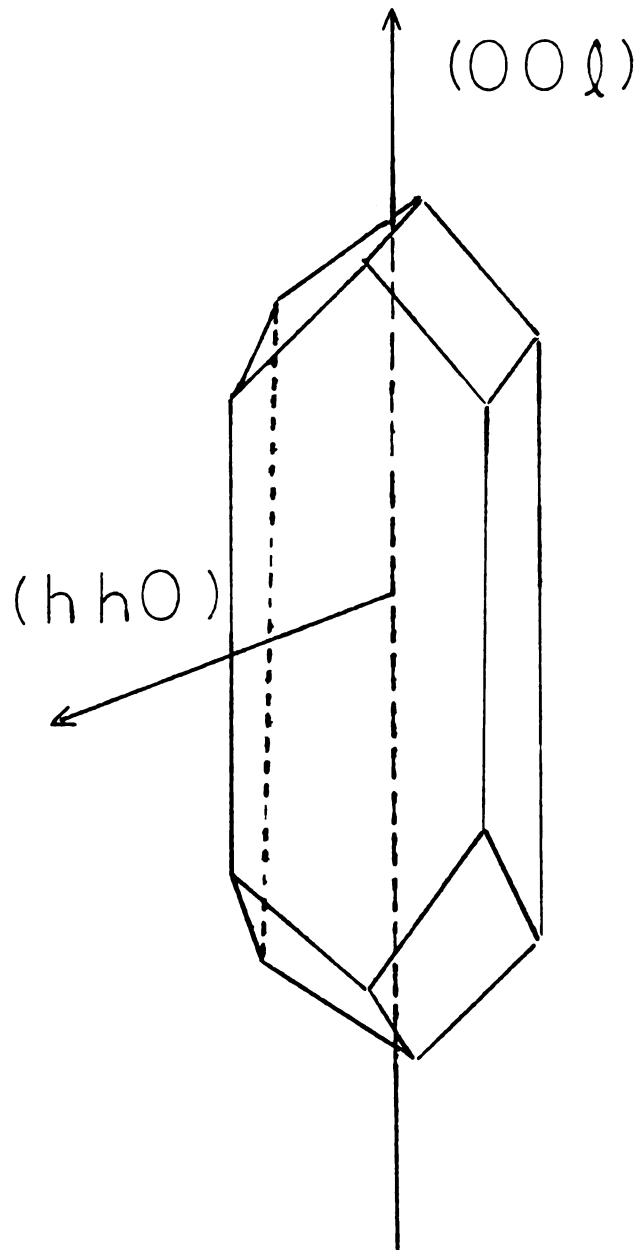


Figure 13. Morphology of an Idealized Fragment 1 Crystal.

at the end (Figure 14(b)). (If a heavy atom derivative is to be mounted, the heavy atom storing solution is used.)

- c) The fragment 1 crystal to be mounted is then transferred with the capillary device shown in Figure 14(b) to the x-ray capillary on the platform just below the air-liquid interface at the open, wide end. The platform is then placed upright in some plasticene. The fragment 1 crystal falls by gravity to the bottom liquid-air interface (Figure 14(c)).
- d) The platform is then placed under a microscope and, while viewing the crystal, the crystal is gently moved beyond the liquid-air interface. Sharp hard objects may chip, crack, or shatter the crystal. A blunt object with a surface of $\geq 0.5 \text{ mm}^2$ such as the rounded tip on a sealed capillary tube of 0.5 or 0.7 mm diameter works well.
- e) Storing solution is removed so that only a small amount remains between the open end and the crystal. The crystal is then oriented so that one broad face (Fig. 13) touches the wall of the capillary tube and the long sides of the crystal (Fig. 13) are parallel with the walls of the tube. Clean, straight broombristles, straight human hair, or a closed capillary tube can be used for this purpose.
- f) The remainder of the mother liquor is then withdrawn and the crystal is "dried" with thin filter paper strips. This is done by simply allowing the paper strips to touch the crystal. Only a small droplet is necessary to hold the crystal to the capillary tube wall. An excessively wet mount can cause crystal slippage and will increase the absorption of x-rays. A small plug of mother liquor is placed about 2-3 mm above the crystal and the capillary is sealed with wax previously softened by a low flame (Figure 14(d)).
- g) The capillary tube is then mounted with the glass sealed tip downward into a small piece of plasticene which sits on the goniometer head. The goniometer head is then optically aligned so the c^* axis of the crystal coincides with the goniometer axis.

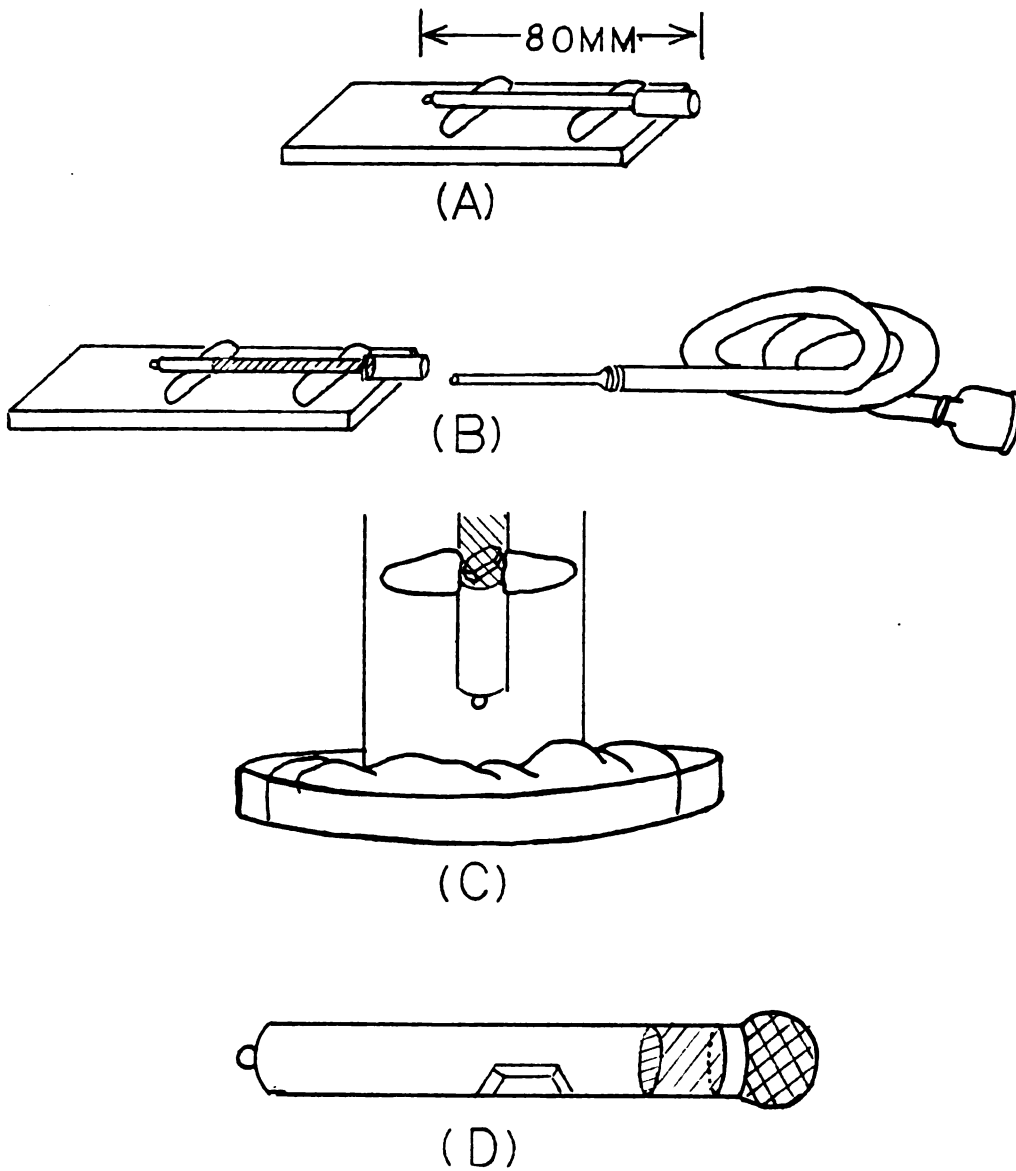


Figure 14. Stages in the Mounting of a Fragment 1 Crystal.

This manner of mounting a fragment 1 crystal makes recording the principal axial intensity distributions of $(hh0)$, \bar{a}^* , and \bar{c}^* reflections convenient. The $(hh0)$ direction is perpendicular to the broad face; the \bar{a}^* direction can be found 45° away and the \bar{c}^* direction is along the long edge at the crystal. The principal axial intensity distributions of native fragment 1 are shown in Figure 15. The axial run-outs show that the crystals scatter x-rays well to 3.5 \AA resolution (25° in 2θ); x-ray photographs of other similar crystals indicate that they scatter x-rays well to 2.5 \AA resolution (36° in 2θ)⁴⁰.

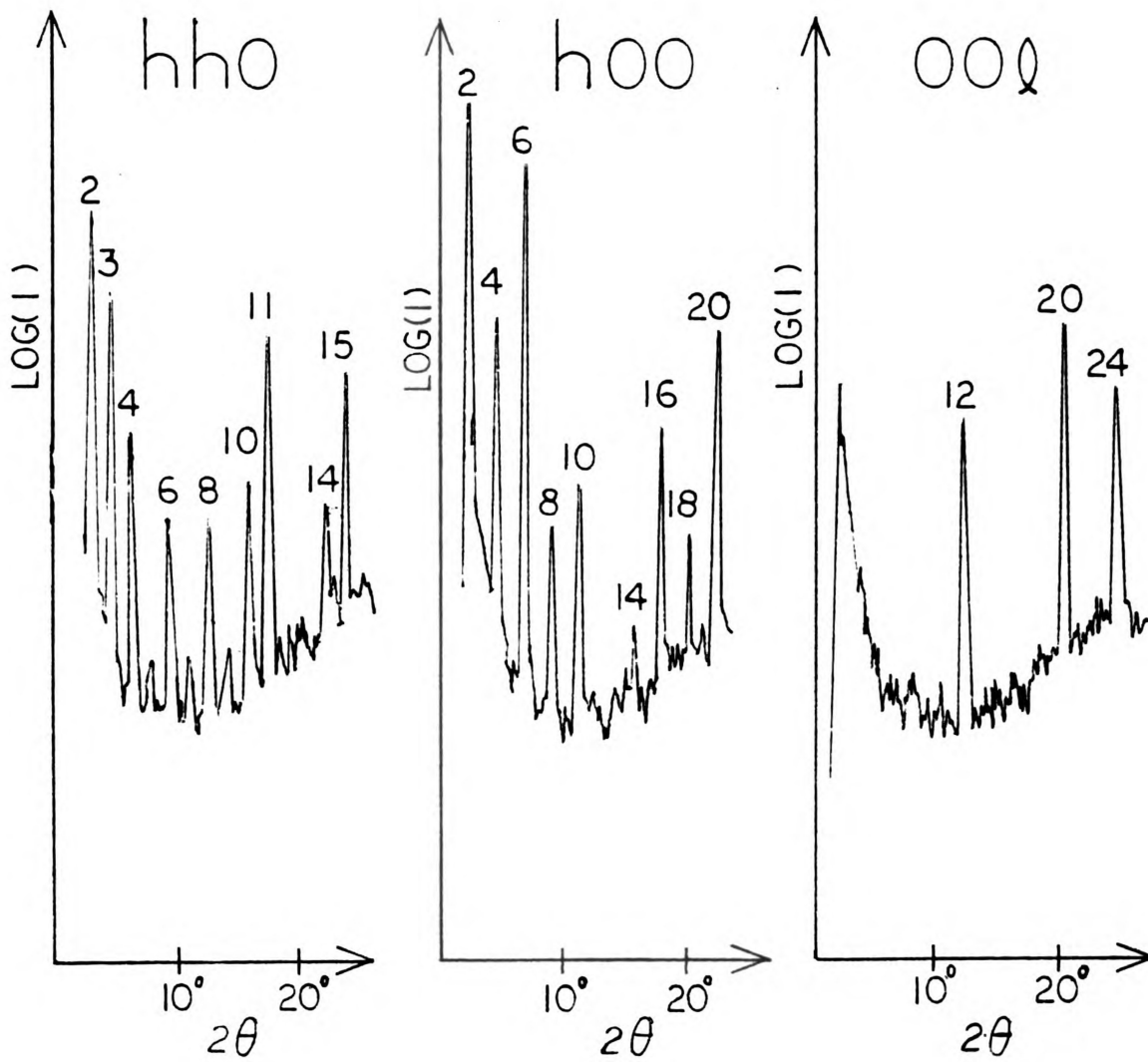


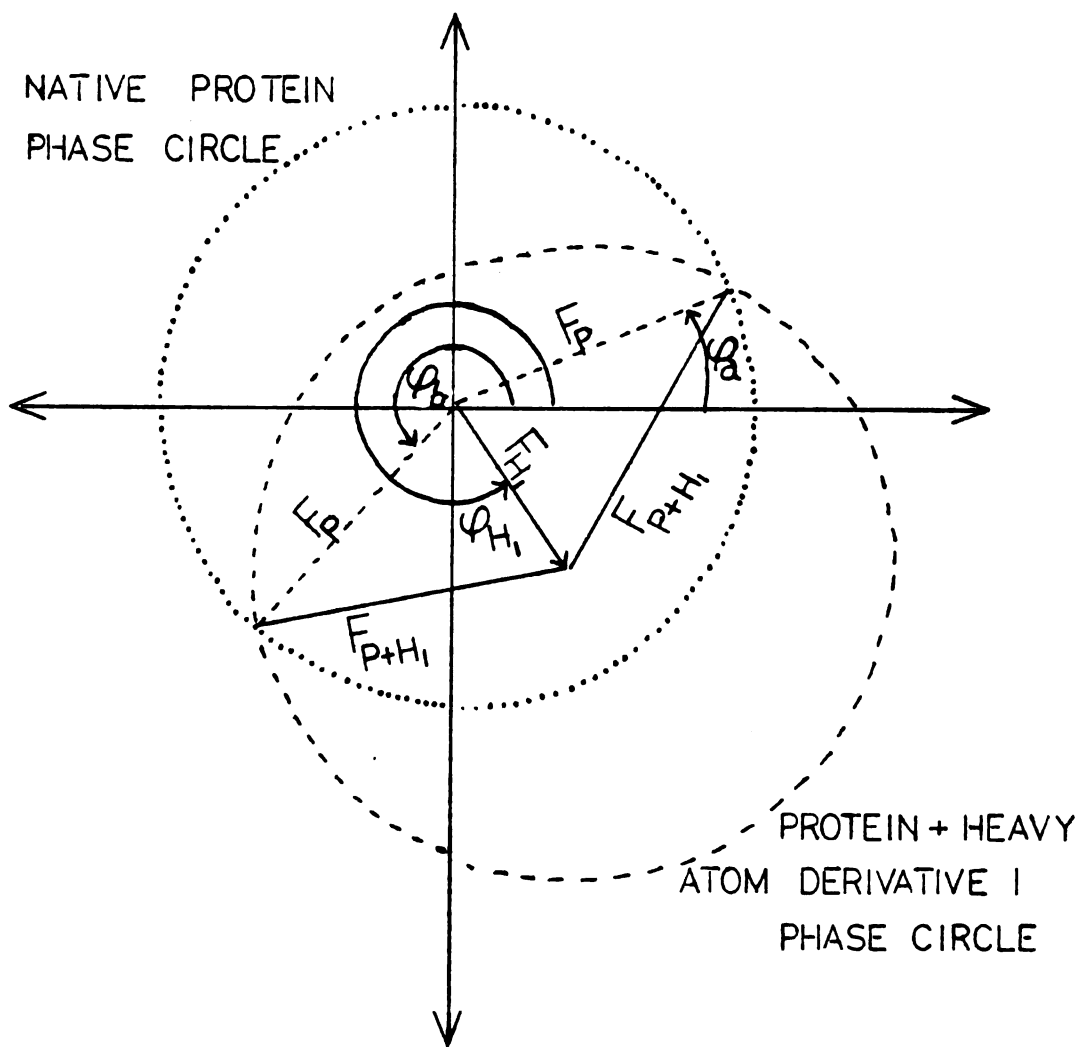
Figure 15. The Principal Axial Intensity Distributions of Native Fragment 1.

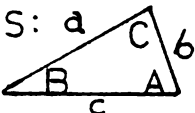
CHAPTER 5

THE PREPARATION OF ISOMORPHOUS HEAVY ATOM DERIVATIVE CRYSTALS

In order to calculate the electron density map for a protein (Equation 1), one needs to know the phase $\alpha(hkl)$ for each structure factor amplitude $|F(hkl)|$. However, there exists no experimental means to measure the phase of electromagnetic radiation. This difficulty is circumvented if two or more unique isomorphous heavy atom derivative crystals of the protein can be prepared. Knowing the positions of the heavy atoms in the unit cell, and having intensity data available from the native and 2 or more derivative crystals, it is possible to phase the protein data by means of the multiple isomorphous replacement method.

Consider Figure 16, a construction devised by Harker², which represents the information available from native protein intensity data, the intensity data from one isomorphous heavy atom derivative and a difference Patterson map between the two. Knowing $|F_P|$, $|F_{P+H}|$, and both $|F_H|$ and ϕ_H , the Law of Cosines (Figure 16) can be employed to obtain two possible phase



LAW OF COSINES:  $a^2 = b^2 + c^2 - 2bc \cos(A)$

TWO SOLUTIONS FOR THE PHASE OF $|F_P|$

1. $|F_{P+H_1}|^2 = |F_P|^2 + |F_{H_1}|^2 - 2|F_P||F_{H_1}|\cos(\varphi_{H_1} - \varphi_b) \therefore \varphi_b$
 A SOLUTION

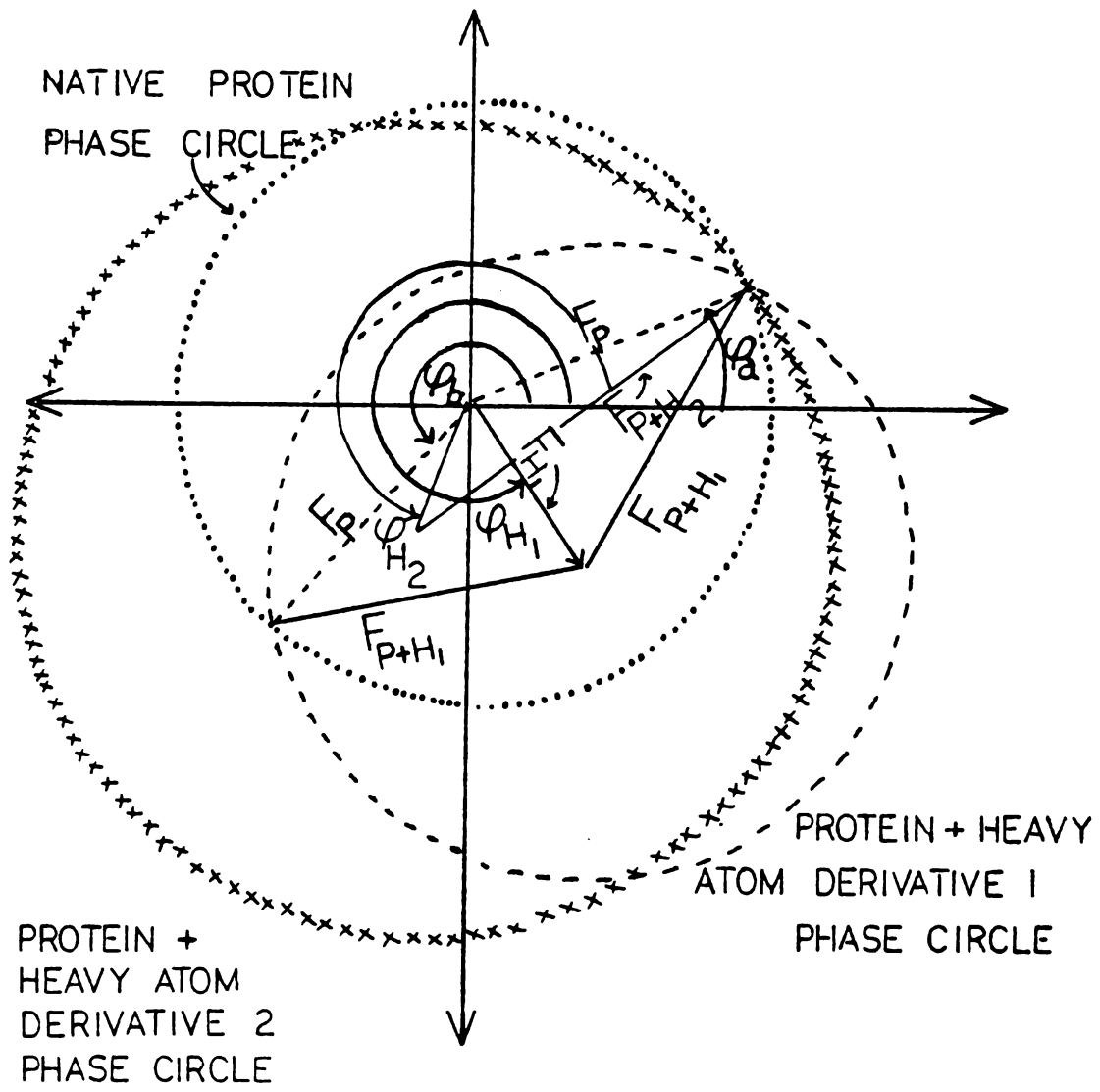
2. $|F_{P+H_1}|^2 = |F_P|^2 + |F_{H_1}|^2 + |F_P||F_{H_1}|\cos(2\pi - \varphi_{H_1} - \varphi_a) \therefore \varphi_a$
 A SOLUTION

Figure 16. A Phase Circle Diagram of a Native Protein and an Isomorphous Heavy Atom Derivative.

solutions for $|F_p|$, ϕ_a and ϕ_b . The processing of the intensity data of a second heavy atom derivative eliminates the phase ambiguity. In Figure 17, for example, the second heavy atom derivative shows that ϕ_a is the correct phase.

Isomorphous heavy atom derivatives of macromolecular crystals are generally prepared by diffusing the reagents into the crystals. This is possible because macromolecular crystals have solvent channels by which ions in solution can enter the crystal. The heavy atom compounds used to make derivatives can be divided into hard acids and soft acids⁴⁶. The common hard acids used in protein crystallizations are uranyl, plumbous, and trivalent lanthanide salts. (A list of hard acid compounds can be found in Table 2). These compounds bind to hard ligands frequently found in protein experiments—buffer anions, amino acid hydroxyl groups, and the terminal carboxyl groups of amino acids⁴⁷. The soft acid heavy atom compounds contain Hg, Pt, Au, or Ag as the major element; Table 3 contains examples of commonly used soft acid compounds in protein crystallography. The soft acid ligands these compounds tend to bind to are sulfur atoms in amino acids, nitrogen-containing aromatics, and terminal amino and imino groups⁴⁷.

The most important parameters in a heavy atom soaking experiment are the concentration of the reagent, the duration of the soak, and the similarity of the heavy atom soaking solution to the native crystal storage



$$3. |\vec{F}_{P+H_2}| = |\vec{F}_P| + |\vec{F}_{H_2}| - 2|\vec{F}_P| |\vec{F}_{H_2}| \cos(2\pi - \phi_{H_2} - \phi_P)$$

$\therefore \phi_A$ IS THE SOLUTION

Figure 17. A Phase Circle Diagram of a Native Protein and Two Isomorphous Heavy Atom Derivatives.

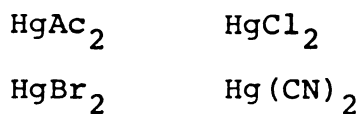
TABLE 2. Common Hard Acids Used in Protein Crystallizations

—Uranyl Salts	$(O=U=O^{2+})$	in:	$UO_2(NO_3)_2$ $(UO_2)_2Cl_4$ $UO_2(Ac)_2$
—Plumbous Salts	Pb^{2+}	in:	$PbAc_2$ $Pb(NO_3)_2$ $PbCl_2$
—Trivalent Lanthanide Salts			La^{+3} } chlorides Sm^{+3} } acetates Eu^{+3} } Tb^{+3} } nitrates

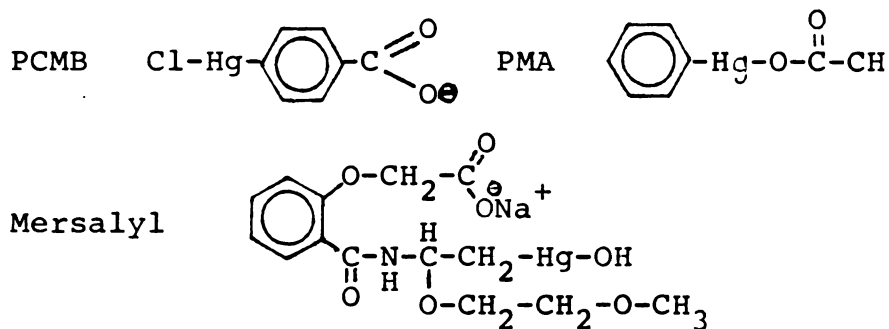
TABLE 3. Common Soft Acids Used in Protein Crystallizations

Hg Compounds

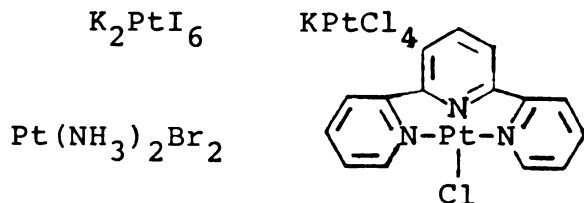
a) Ionic Mercury Compounds:



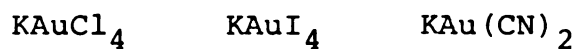
b) Ionic-Covalent Mercury Compounds:



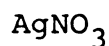
Pt Compounds



Au Compounds



Ag Compounds



solution. Concentrations of heavy atom reagents reported in the literature range from 0.05 mM⁴⁵ to 100 mM⁴⁹ but most are from 1-10 mM⁴⁷. It is desirable to keep the concentration of the reagents as dilute as possible but still affect a change in the diffraction pattern. Too dilute a concentration may not cause changes in the intensities and too strong a concentration can cause excessive binding and destroy the diffracting quality of the crystal. The durations of a crystal soak in heavy atom reagent solutions have been reported as short as 1 hour⁵⁰ and as long as 6-9 months⁵¹, but the most common time is 1-2 weeks⁴⁷. Finding both the effective reagent concentration and soaking duration is an empirical process that requires trial-and-error testing. The similarity of the heavy atom soaking solution to the native crystal storage solution is very important since one wants to keep the crystals as isomorphous as possible with the native crystal. In particular, if the pH of the heavy atom solution is different from that of the native solution observed intensity changes in the diffraction pattern may be due to pH-induced structural changes to amino acid groups rather than due to the binding of the heavy atom. An important study by Mavridis, Tulinsky, and Liebman⁵² showed that axial intensity distributions of pH conformers of α -chymotrypsin even as close in pH as 2.5 and 3.5 had substantial intensity differences. A safe practice is to have the native crystal storage solution and the

heavy atom soaking solution differ only in the presence of the heavy atom compound.

A. Screening for Possible Isomorphous Heavy Atom Derivatives of Fragment 1

Fragment 1 crystals were introduced to heavy atom soaking solutions by exchanging the storage solution for the heavy atom solution in a manner similar to that used in the crystal harvesting procedure. After 1 week to 3 weeks the crystals were removed, mounted, and the diffraction pattern of the $(hh0)$, \bar{a}^* , \bar{c}^* , and $(h0h)$ reflections was recorded on a Picker four-circle diffractometer controlled by Digital Equipment Corporation (DEC) PDP-8 computer (FACS I system). If a heavy atom derivative showed substantial changes in the intensity distribution throughout the scattering range, a data collection to 3.5 Å resolution was carried out.

Many lanthanide ion salts were used in attempts to obtain heavy atom derivatives of fragment 1. These experiments were conducted because the lanthanide ions are similar to Ca^{+2} in ionic radius and have been successfully used by many workers to gain information about the Ca^{+2} ion mediated Gla-PL interaction⁵³. The use of $^{139}\text{La}^{+3}$ enables the metal ion-Gla interaction to be followed by NMR⁵³. Many lanthanides also have favorable luminescence properties which allow the use of changes in intrinsic fluorescence of the blood coagulation factors as indirect measures of metal-ion

binding^{25,54,55}. Two types of Eu^{+3} binding sites have been found in fragment 1 in solution⁵³, and Tb^{+3} binds to only 2 of the 10 Glu residues in the protein⁵⁶.

The I_3^- ion was also investigated at several different concentrations in attempts to obtain suitable heavy atom derivatives. Selective iodination of a single tyrosine residue of α -chymotrypsin was accomplished using this reagent⁵⁷. Since fragment 1 has only 3 tyrosine residues, the procedure was given serious consideration.

A list of heavy atom derivative compounds that were studied along with observations is presented in Tables 4 and 5.

B. Promising Platinum Derivatives of Fragment 1

In the heavy atom derivative survey, two promising platinum derivatives were found—saturated $\text{K}_2\text{Pt}(\text{C}_5\text{H}_5\text{N})_2\text{Cl}_2$ and 10 mM $\text{Pt}(\text{NH}_3)_2(\text{NO}_2)_2\text{Br}_2$. A 4.4 Å resolution data set ($2\theta < 20$) was collected in the 10 mM $\text{Pt}(\text{NH}_3)_2(\text{NO}_2)_2\text{Br}_2$ derivative (hereafter referred to as FlPT1) and a 3.5 Å resolution data set was collected with the saturated $\text{K}_2\text{Pt}(\text{C}_5\text{H}_5\text{N})_2\text{Cl}_2$ derivative (called FlPT2). The data of both heavy atom derivatives were processed to yield difference Patterson maps. The difference Patterson map of the $\text{Pt}(\text{NH}_3)_2(\text{NO}_2)_2\text{Br}_2$ derivative proved to be crystallographically inconsistent and uninterpretable; however the map of the $\text{Pt}(\text{C}_5\text{H}_5\text{N})_2\text{Cl}_2$ derivative corresponded to one heavy atom position per fragment 1 molecule. The

TABLE 4. Hard Acid Heavy Atom Derivative Trials

<u>Lanthanide Compounds</u>		
<u>Compound</u>	<u>Soaking Time</u>	<u>Result</u>
ErCl ₃ -sat'd.	17 days	low angle changes: pattern absent after 8° in 2θ.
Er(NO ₃) ₃ -sat'd.	1 week	big changes in (2,2,0) and (2,0,0); diffraction pattern absent there- after.
Er(NO ₃) ₃ -1 mM	11 days	diffraction pattern completely destroyed.
LaCl ₃ -sat'd.	16 days	big changes in (2,2,0) and (2,0,0).
LaCl ₃ -1 mM	12 days	native run-outs obtained.
LaCl ₃ -½sat'd.	4 weeks	native run-outs obtained.
PrCl ₃ -sat'd.	18 days	big low angle changes but diffraction pattern absent after 5° in 2θ.
SmCl ₃ -8 mM	1 week	big changes in (2,2,0) and (2,0,0) but diffraction pattern absent thereafter.
TbCl ₃ -1 mM	2 weeks	native run-outs obtained.
<u>Miscellaneous Compounds</u>		
<u>Compound</u>	<u>Soaking Time</u>	<u>Result</u>
PbCl ₃ -sat'd.	1 week	native run-outs obtained.
BaCl ₃ -sat'd.	5½ weeks	mostly low angle changes in run-outs.

TABLE 5. Soft Acid Heavy Atom Derivative Trials.

<u>Platinum Compounds</u>		
<u>Compound</u>	<u>Soaking Time</u>	<u>Result</u>
$K_2Pt(CN)_4$ -20 mM	3½ weeks	native run-outs obtained.
K_2PtI_6 -sat'd.	6 weeks	diffraction pattern destroyed.
$K_2Pt(C_5H_5N)_2Cl_2^a$ -sat'd.	8 days	<u>promising-good</u> changes throughout run-outs.
$K_2Pt(C_5H_5N)_2Cl_2$ -sat'd.	3 weeks	diffraction pattern destroyed.
$K_2Pt(C_5H_5N)_2Cl_2$ -sat'd.	2 weeks	run-outs the same as those of 8 day soak.
K_2PtCl_4 -20 mM	1 week	diffraction pattern destroyed.
K_2PtCl_4 -5 mM	20 days	diffraction pattern of higher order reflections destroyed; remainder of pattern looks like native.
$Pt(NH_3)_2(NO_2)_2Br_2$ -10 mM	6 days	low angle changes in run-outs; try a longer soak.
$Pt(NH_3)_2(NO_2)_2Br_2$ -10 mM	2 weeks	<u>promising-nice</u> changes on c^* axis reflections to 20° in 2θ ; also changes in (hh0), a^* , and (h0h) run-outs.
$Pt(NH_3)_2(NO_2)_2Br_2$ -10 mM	3 weeks	essentially the same pattern as that of the 2 week soak.

Table 5 Continues.

Table 5 Continued.

<u>Mercury Compounds</u>		
<u>Compound</u>	<u>Soaking Time</u>	<u>Result</u>
PMA ^b -20 mM	2 weeks	diffraction pattern destroyed.
PMA-5 mM	11 days	crystal too small; diffraction pattern weak.
PMA-5 mM	1 week	native run-outs obtained
PCMBS ^c -20 mM	4½ weeks	native run-outs obtained.
PCMBS-40 mM	11 days	native run-outs obtained.

<u>Gold Compounds</u>		
<u>Compound</u>	<u>Soaking Time</u>	<u>Results</u>
KAuBr ₄ -sat'd.	3 weeks	diffraction pattern destroyed.
KAuBr ₄ (<<1 mM)	13 days	run-outs very similar to those of native.
KAuI ₄ -sat'd.	3 weeks	diffraction pattern destroyed.
KAuCl ₄ -sat'd.	22 days	native run-outs obtained.

<u> KI₃ </u>		
<u>Compound</u>	<u>Soaking Time</u>	<u>Result</u>
KI ₃ -0.4 mM	12 days	changes on \bar{a}^* only.
KI ₃ -4 mM	2 weeks	native run-outs obtained.
KI ₃ -10 mM	2 weeks	slight changes in (hh0), \bar{a}^* , \bar{c}^* , and (h0h) run-outs.
KI ₃ -20 mM	18 days	diffraction pattern destroyed.

Notes for Table 5 on Following Page.

Notes for Table 5:

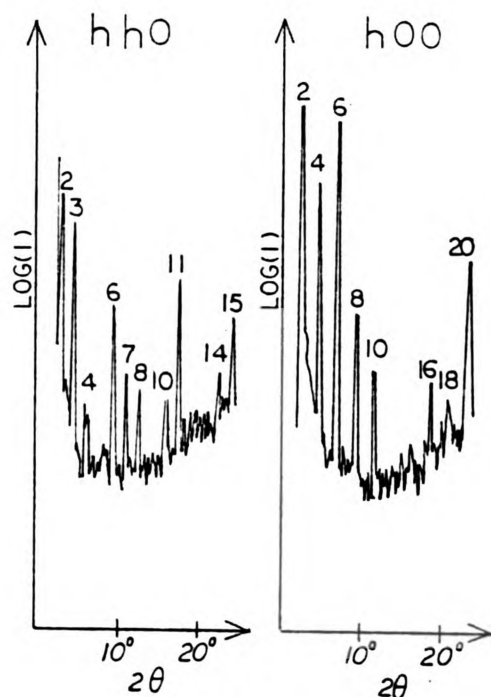
^aThis platinum derivative was originally prepared by Dr. Chang H. Park of this laboratory. However, the 3.5Å data collection and data processing are part of this work.

^bPMA = phenyl mercury acetate.

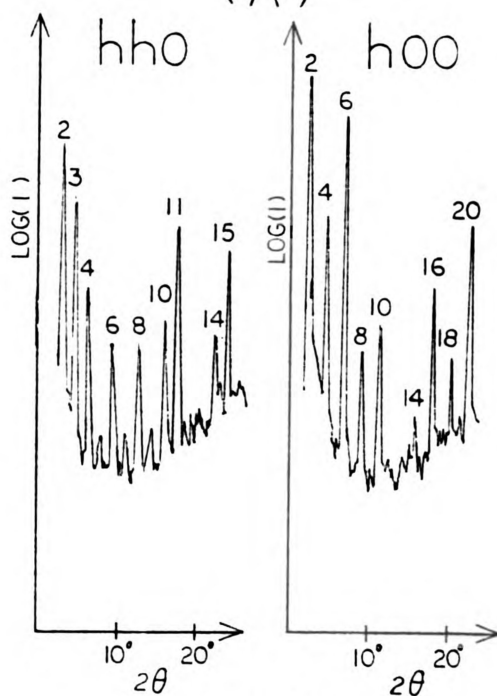
^cPCMBS = parachloromercury benzene sulphonate.

remainder of this thesis deals with the processing of the FlPT2 data to the level of the difference Patterson map.

Principal axial intensity distributions of the saturated $K_2Pt(C_5H_5N)_2Cl_2$ derivative and those of the 10 mM $Pt(NH_3)_2(NO_2)_2Br_2$ derivative—with corresponding run-outs of native fragment 1—are shown in Figures 18 and 19 respectively. Substantial differences in intensity exist in these derivatives. Some of the relative differences are summarized in Tables 6 and 7.

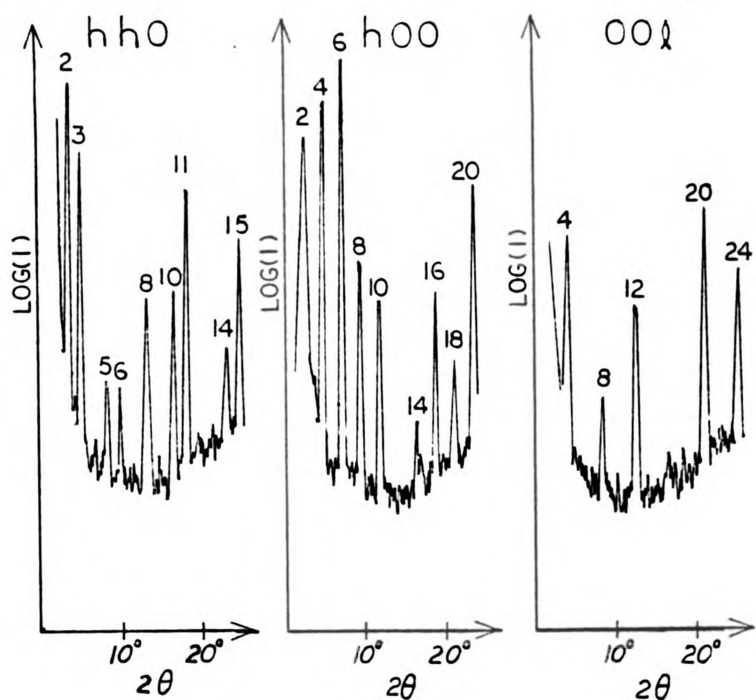


(A)

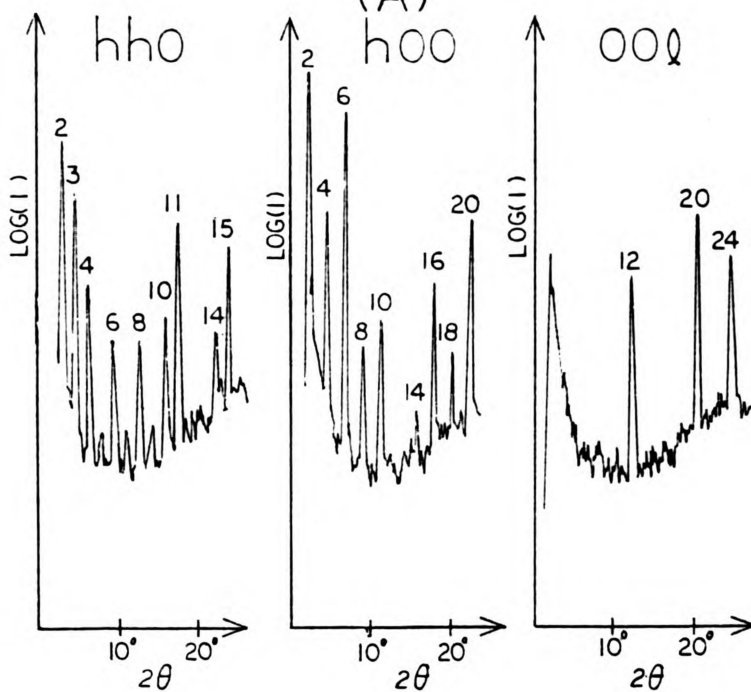


(B)

Figure 18. Principal Axial Intensity Distributions of (A) 10 mM $\text{Pt}(\text{NH}_3)_2(\text{NO}_2)_2\text{Br}_2$ and (B) Native Fragment 1.



(A)



(B)

Figure 19. Principal Axial Intensity Distributions of (A) Saturated $K_2Pt(C_5H_5N)_2Cl_2$ and (B) Native Fragment 1.

TABLE 6. Principal Axial Intensity Distribution Differences Between Fragment 1 +10 mM $\text{Pt}(\text{NH}_3)_2(\text{NO}_2)_2\text{Br}_2$ and Native Fragment 1

<u>Axis</u>	<u>Reflection</u>	<u>Relative Difference from Native Protein</u>
hh0	(3,3,0)	up in derivative
hh0	(4,4,0)	down in derivative
hh0	(6,6,0)	up in derivative
hh0	(7,7,0)	up in derivative
hh0	(10,10,0)	down in derivative
h00	(4,0,0)	up in derivative
h00	(8,0,0)	up in derivative
h00	(14,0,0)	down in derivative
h00	(16,0,0)	down in derivative

TABLE 7. Principal Axial Intensity Distribution Differences Between Fragment 1 + Saturated $K_2Pt(C_5H_5N)_2Cl_2$ and Native Fragment 1

<u>Axis</u>	<u>Reflection</u>	<u>Relative Difference from Native Protein</u>
hh0	(4,4,0)	almost absent from derivative.
hh0	(5,5,0)	substantial in derivative, almost absent in native.
hh0	(6,6,0)	down in derivative.
hh0	(8,8,0)	up in derivative.
h00	(4,0,0)	up in derivative.
h00	(8,0,0)	up in derivative.
h00	(16,0,0)	down in derivative.
001	(0,0,4)	large peak in derivative; absent in native.
001	(0,0,8)	large peak in derivative; absent in native.
001	(0,0,20)	up in derivative.

CHAPTER 6

THE 3.5 Å RESOLUTION INTENSITY DATA COLLECTION OF THE FlPT2 FRAGMENT 1 CRYSTAL

The fragment 1 crystal used in the FlPT2 data collection had dimensions of 1.56 mm × 0.68 mm × 0.40 mm, and it was soaked in a saturated solution of $\text{Pt}(\text{C}_5\text{H}_5\text{N})_2\text{Cl}_2$ for 13 days. The soaking solution was prepared by adding the platinum salt to the native crystal storing solution and adjusting the pH to 7.5. The crystal was mounted in a capillary tube by the method outlined in Chapter 4. These crystals belong to the space group $P4_12_12$ or $P4_32_12$, and thus there are eight equivalent positions in the unit cell⁵⁸. Since anomalous scattering was not considered, only 1/16th of the reciprocal sphere of reflection was surveyed to obtain a unique data set.

The intensity data were measured using monochromatic Cu K_α radiation (1.5418 Å) with a Nicolet P3/F four-circle diffractometer. A schematic drawing of the goniostat of this diffractometer is shown in Figure 20. The goniostat orients the crystal and detector with respect to the stationary x-ray beam, and brings Bragg planes sequentially into reflecting position. W.L. Bragg noticed the similarity of the diffraction of

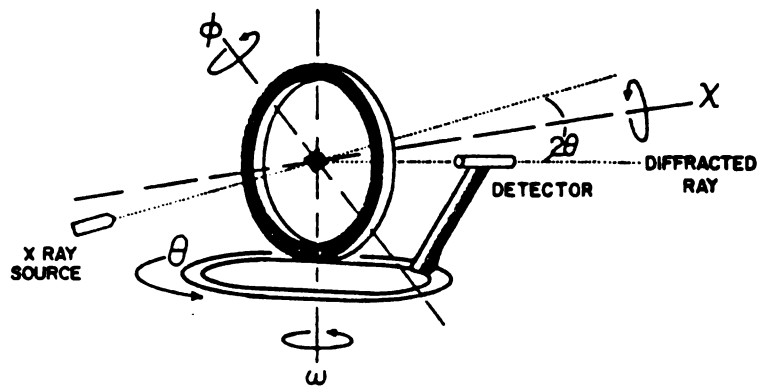


Figure 20. The Goniostat of a Four-Circle Diffractometer (taken from reference 33).

x-rays by crystals to ordinary reflection, and he deduced a simple equation — $n\lambda = 2d \sin \theta$ — by treating diffraction as "reflection" from planes of atoms in a lattice⁵⁹. Since every diffracted x-ray beam from a crystal can be associated with an (hkl) indexed Bragg plane, it is common to refer to a diffracted x-ray from a crystal as a reflection.) The four circles of the goniostat refer to the 3 circles swept out by rotation of the crystal about the ω , ϕ , and χ axes (Figure 20), and the circle swept out by the arm of the detector through rotation about the θ axis (which is coincident with the ω axis). Each diffracted x-ray beam is therefore defined by an (hkl), as well as by 2θ , ω , ϕ and χ angles. The purpose behind a data collection of a crystal is to collect an intensity (and associated data such as background intensity measurements) for each reflection in the data set.

The diffractometer is controlled by a Data General Nova/4 computer. The (hkl) indices of the unique reflections observed with native fragment 1 were stored in a file on disk. The file was sorted by (hkl)-h varying the slowest and l varying the fastest. However, the intensity data was not collected in this order. The 3-D data to 3.5 Å resolution was collected in 5 shells of 2θ in the following order: $(22.5-25.1)^\circ$, $(20.0-22.5)^\circ$, $(18.0-20.02)^\circ$, $(15.0-18.0)^\circ$, and $(2.5-15.0)^\circ$. Once the unit cell dimensions of the crystal ($a, b, c, \cos \alpha, \cos \beta, \cos \gamma$) and the orientation of the crystal on the diffractometer are known, 2θ , ω , ϕ , and

χ values can be calculated. Thus, when the $(22.5-25.1)^\circ$ shell was being collected, the computer simply searched the file until it found the first reflection with a 2θ value in that range, drove the goniostat to those angles, and obtained the intensity information for that reflection. Once the intensity measurement had been made, the next reflection in the file with a 2θ value in that range was found, and the process continued until all reflections in the $(22.5-25.1)^\circ$ range had been measured.

The method used to measure the reflection intensities in the FlPT2 data collection was first proposed by Wyckoff et al.⁶⁰ and is now known as the Wyckoff ω -step procedure. Using this method, the peak was scanned 0.12° on either side of the calculated ω value for the reflection, in 7 steps performed in 0.03° increments. The four largest intensities measured are then summed to contribute to the recorded intensity of that reflection. The procedure also allows the diffractometer to take additional steps (2) on either side of the calculated peak if an intensity exceeds a minimum limit. (I_{\min} , the minimum intensity must be decided upon by the experimenter before the 3-D data are collected.)

The 2θ , ω , ϕ and χ angles which the goniostat drove to in the course of making an intensity measurement resulted from the unit cell parameters and the orientation of the crystal on the diffractometer. The former were obtained by performing a least squares calculation between the

calculated and observed angles in a reflection array. The criteria used to select these reflections were: the intensities are large enough so that the centering routine will center them after typical (5-10%) decay has occurred; the reflections have a relatively high 2θ value ($\sim 17^\circ - 25^\circ$); and, several reflections of low, intermediate, and high χ values with respect to the χ range are chosen.

In this data collection, only reflections with $17.5 < 2\theta < 25^\circ$ were used for the array, and since the χ range of the data collection was from $0^\circ - 90^\circ$, 4 reflections had $0^\circ < \chi < 30^\circ$, 4 reflections had $30^\circ < \chi < 60^\circ$, and 4 reflections had $60^\circ < \chi < 90^\circ$. Each time the least squares calculation was performed, the cell dimensions of the crystal were calculated. A listing of the average cell parameters obtained by least squares for the FlPT2 data collection and for the FlN6 data collection is shown in Table 8, from which it can be seen the two are dimensionally quite isomorphous.

Approximately 2,700 unique reflections were measured during the data collection and it took approximately 36 hours exposure to measure all of the intensities. The number of reflections in each of the five 2θ shells used is shown in Table 9. During the course of the experiment, the alignment of the crystal was monitored by measuring the intensity of 3 reflections periodically (every 100 reflections which amounted to every 75 minutes). If the

TABLE 8. The Average Cell Parameters of FlN6 (Native Fragment 1) and of FlPT2

<u>Crystal</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>α</u>	<u>β</u>	<u>γ</u>
FlN6	77.56	77.54	85.23	90.01	90.09	90.01
σ	± 0.03	± 0.04	± 0.05	± 0.04	± 0.03	± 0.01
FlPT2	77.42	77.46	85.02	90.07	90.08	90.00
σ	± 0.10	± 0.13	± 0.03	± 0.02	± 0.03	± 0.01

TABLE 9. The Number of Reflections in Each Data Collection 2 θ Shell

<u>The FlPT2 Data Collection</u>	
<u>2θ Shell</u>	<u>Number of Reflections</u>
(22.5-25.1) $^\circ$	683
(20.0-22.5) $^\circ$	617
(18.0-20.02) $^\circ$	367
(15.0-18.0) $^\circ$	421
(2.5-15) $^\circ$	645
Total Number of Reflections	2733

intensity of these monitor reflections decreased by greater than 15%, the crystal was considered misaligned. This triggered the re-measurement of the 12 reflections in the reflection array, followed by the calculation of a new orientation matrix and cell parameters from the least squares method. The 3 reflections used for this purpose and their corresponding 2θ values were: (14,0,7) ($2\theta = 17.62$); (7,7,17) ($2\theta = 21.09$); and (2,0,23) ($2\theta = 24.18$). The intensities of these reflections were plotted versus exposure time to the crystal and the resulting slopes were used to obtain decay corrections for the data between $15^\circ < 2\theta < 25^\circ$. A plot of these intensities vs. time is shown in Figure 21. The decay correction is discussed in Chapter 7, Section A.

Several measurements preceded and followed the actual intensity data collection; these measurements include the ω profile of a reflection (7,7,13), the intensity absorption versus ϕ angle [(0,0,4) and (0,0,20)], and the (hk0) zone reflections to 5.9 Å resolution. The ω profile of a reflection assesses the quality of a crystal for 3-D data collection and contains information which is used to initiate the Wyckoff ω -step scanning procedure. The ω -profile of (7,7,13) (Figure 22) was made by measuring the intensity of the reflection at ω values $\pm 0.65^\circ$ from its calculated value in 0.05° intervals of ω . A split profile, for example, may indicate that the crystal is cracked and that 2 closely related but distinct crystal

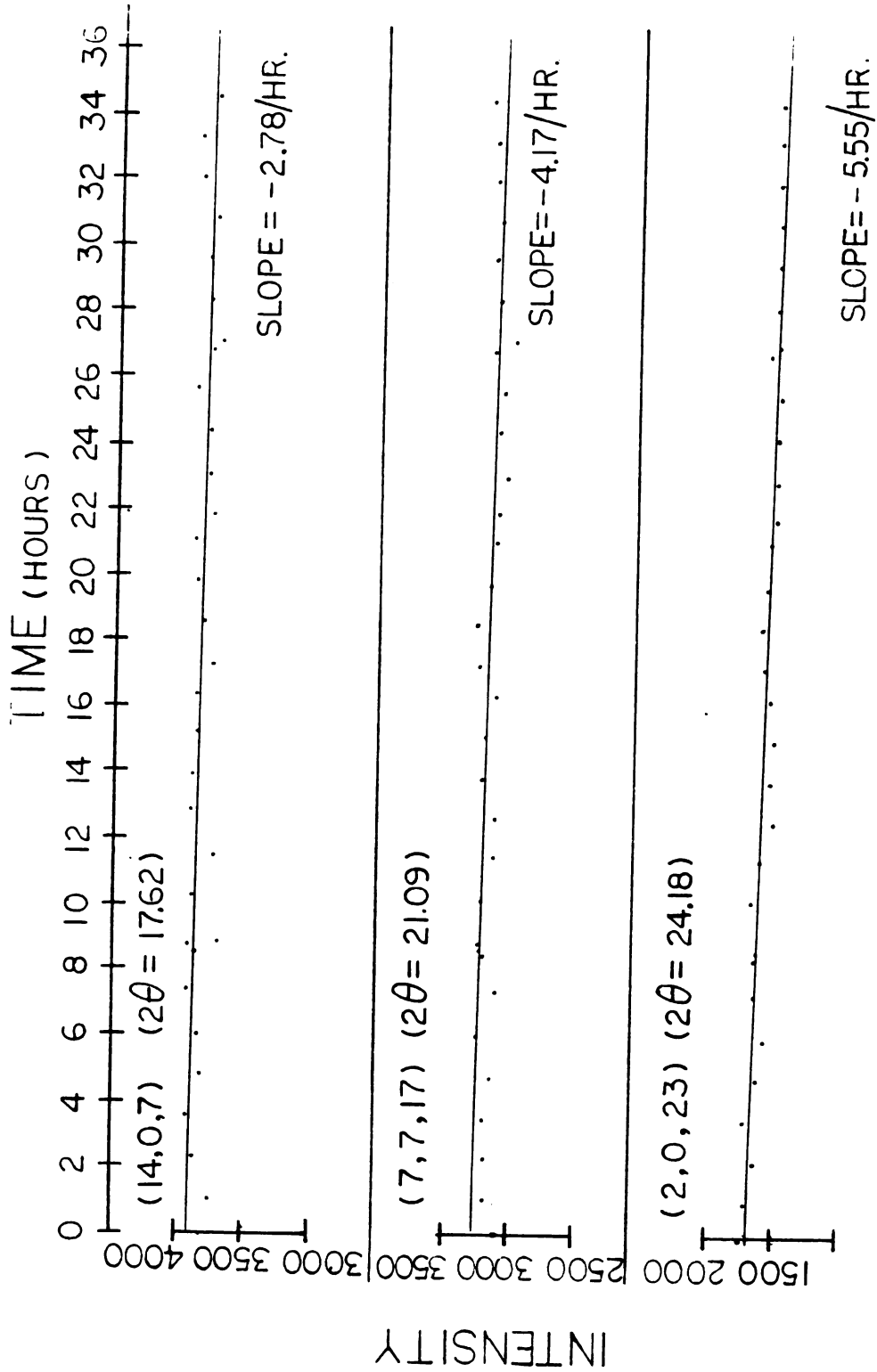


Figure 21. The Intensity of the FlPT2 Monitor Reflections vs. Time.

INTENSITY
(COUNTS/5 SEC.)

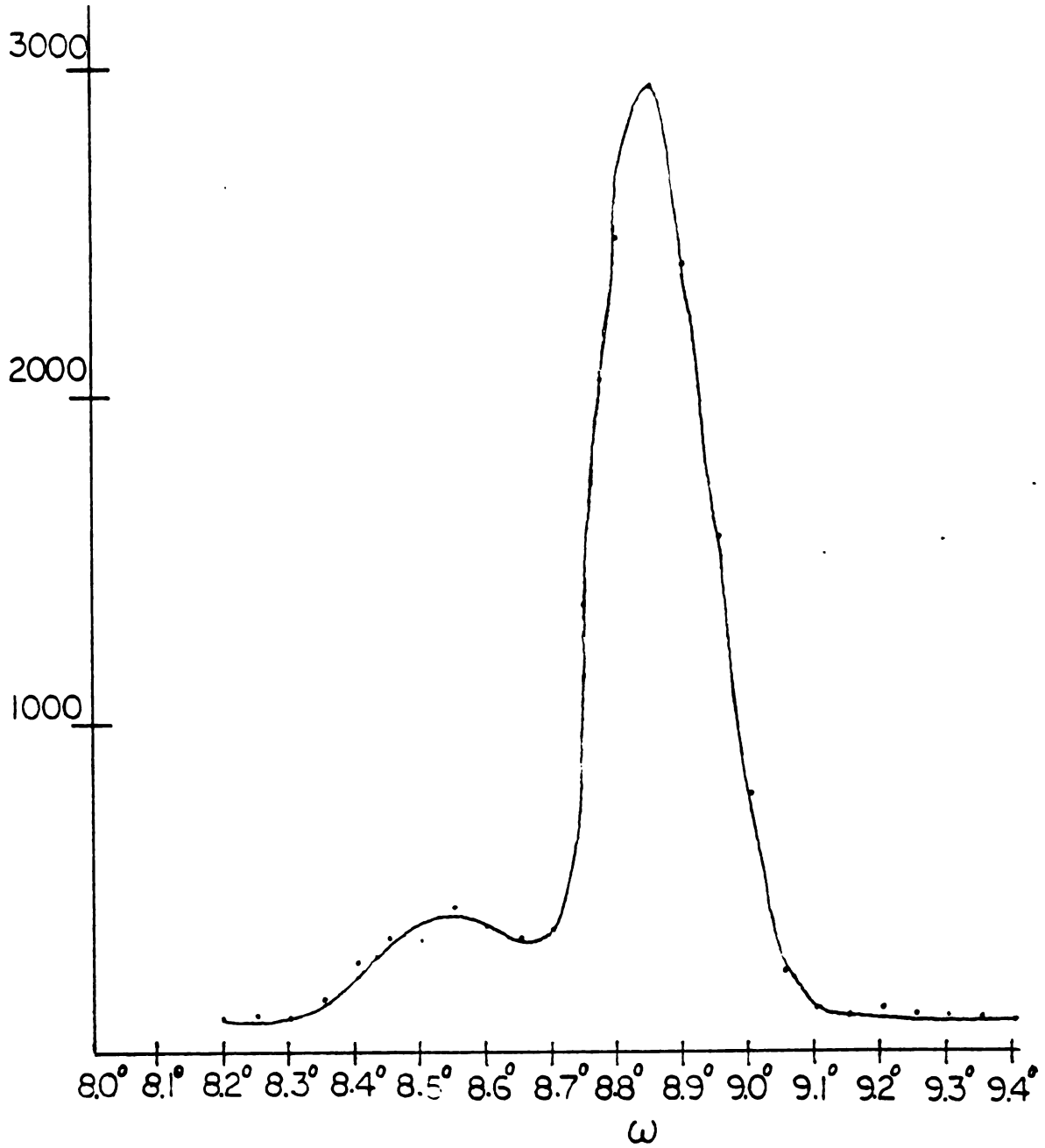


Figure 22. ω -Profile of (7,7,13) Taken Before the 3-D FlPT2 Data Collection.

orientations are contributing to the profile. The breadth of a single peak at half the maximal height (Figure 22) defines the region of the peak which should be scanned. In the FlPT2 data collection, 0.24° was used. From this plot, an offset value can be obtained to measure the background. Due to the shoulder on the left side of the profile for (7,7,13) (Figure 22), 0.45° was used for this experiment.

The intensity absorption versus ϕ angle was measured for reflections (0,0,4) and (0,0,20) to find the region in ϕ of low x-ray absorption so the data collection could be carried out over this ϕ range, and so that an absorption correction could ultimately be applied to the measured intensities. The \bar{c}^* axis reflections (0,0,4) and (0,0,20) were used because they were of fairly high intensity in FlPT2 (Figure 19), and because in fragment 1 crystals (Chapter 4) the \bar{c}^* direction is coincident with the ϕ axis so that the (001) Bragg planes are always in the reflecting position for all values of ϕ at $\chi = 90^\circ$. Intensity measurements were made every 10° in ϕ over an 100° range for both reflections. Plots of $I(\phi)$ versus ϕ taken before and after the data collection on reflections (0,0,4) and (0,0,20) are shown in Figure 23; the 45° range blocked off in the graphs was the ϕ range employed in the data collection. This range was chosen because the intensities were high and fairly constant with respect to ϕ (implying low absorption) as opposed to the regions on the right and

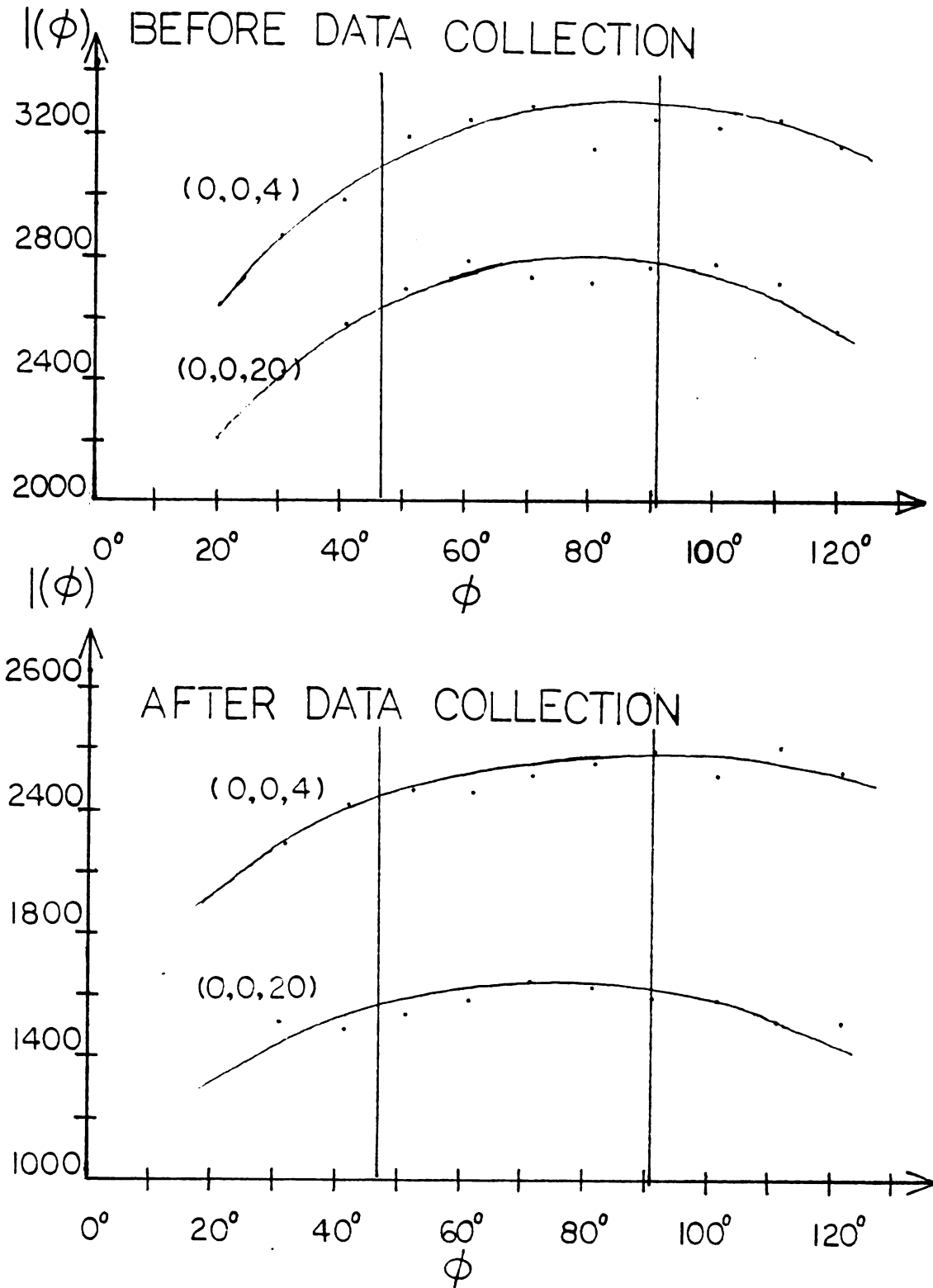


Figure 23. Intensity Absorption Plots of (0,0,4) and (0,0,20) Before and After the FlPT2 Intensity Data Collection.

left of the selected range. The absorption correction applied to the intensity data is discussed in Chapter 7, Section A.

The intensities of the (hk0) reflections at 5.9 Å resolution (15° in 2θ) were measured before and after the data collection to assess decay due to x-ray exposure to reflections in the $2.5^\circ < 2\theta < 15.0^\circ$ range. The use of the (hk0) intensity data for this purpose will be discussed in Chapter 7, Section A.

The flow which transpired during the FlPT2 work is summarized in Table 10.

TABLE 10. The Course of Events in the FlPT2 Data Collection

1. Optically align crystal.
 2. Run-out (hh0) axis.
 3. Run-out (h00) axis.
 4. Make 3-reflection matrix.
 5. Make first 12 reflection matrix.
 6. Plot ω -profile of (7,7,13).
 7. Make intensity absorption measurements of (0,0,4) and (0,0,20).
 8. Make second 12 reflection matrix.
 9. Measure (hk0) reflections at 5.9 Å resolution ($2\theta < 15^\circ$)
 10. Collect 3D data: $22.5^\circ < 2\theta < 25.1^\circ$.
 11. Collect 3D data: $20.0^\circ < 2\theta < 22.5^\circ$.
 12. Collect 3D data: $18.0^\circ < 2\theta < 20.0^\circ$.
 13. Collect 3D data: $15.0^\circ < 2\theta < 18.0^\circ$.
 14. Collect 3D data: $2.5^\circ < 2\theta < 15.0^\circ$.
 15. Make third 12 reflection matrix.
 16. Measure (hk0) reflections at 5.9 Å resolution.
 17. Run-out (hh0) axis.
 18. Run-out (h00) axis.
 19. Run-out (001) axis.
 20. Make intensity absorption measurements of (0,0,4) and (0,0,20).
 21. Plot ω -profile of (7,7,13).
-

CHAPTER 7

PROCESSING THE F1PT2 DATA

A. Converting the Intensity Data to Structure Amplitudes

The first step in processing protein intensity data is to reduce the measurement for each reflection (hkl) to its corresponding structure factor modulus, $|F(hkl)|$. Structure factor amplitudes are often represented as $|F_o|$ or as $|F_{obs}|$ because they are determined from "observed" intensity data. The F1PT2 intensity data was reduced using a program called P-DATA, written by C.D. Buck, of this laboratory. The program converts diffraction intensities from either the Nicolet P3/F or the Picker FACS-I diffractometers by means of the equation

$$|F(hkl)|^2 = \text{CONST} \times \text{ABS} \times \text{DEC} \times \text{LORPOL} \times I(hkl) , \quad (2)$$

where CONST is a constant used to scale the data, ABS is the absorption correction factor, DEC is the decay correction factor which increases intensities which have decreased due to deterioration of the crystal from x-ray exposure, LORPOL is the Lorentz-polarization factor which is a geometrical factor, and $I(hkl)$ is the background-corrected intensity of a reflection (hkl).

The absorption correction applied in the P-DATA program was based on the method proposed by North, Phillips, and Mathews⁶¹. The program uses absorption tables of $I_{\max}/I(\phi)$ versus ϕ to apply the absorption correction. The absorption correction may vary with 2θ , and up to 6 tables are allowed to be used. From the curves drawn to fit the intensity absorption versus ϕ plots of Figure 23, $I_{\max}/I(\phi)$ versus ϕ plots were made for (0,0,4) and (0,0,20) before and after the data collection (Figure 24). All four absorption curves showed the same basic trend (Figure 24) so only one table was used, and it was obtained by averaging the $I_{\max}/I(\phi)$ values of (0,0,4) and (0,0,20) taken before the data collection; this table is presented in Table 11.

The decay correction factor, DEC, corrects the intensity deterioration of a reflection as a function of exposure time and it is often 2θ dependent. The decay curves of the 3 check reflections in Figure 20 are typical of protein reflections—the intensities decrease as a function of exposure time with negative slopes (s). For both the high angle ($15^\circ < 2\theta < 25^\circ$) and the low angle data ($2.5^\circ < 2\theta < 15^\circ$) the DEC factor has the form

$$I^\circ(t) = I(t) \left[\frac{1}{1-St} \right] \quad (3)$$

where $S = -s/I_{CR}^\circ$ and I_{CR}° is the intensity of a check reflection at zero time. The DEC Factors for high angle data are determined by the following means. Values of S for the 3 reflections monitored in Figure 20 —(14,0,7), (7,7,17),

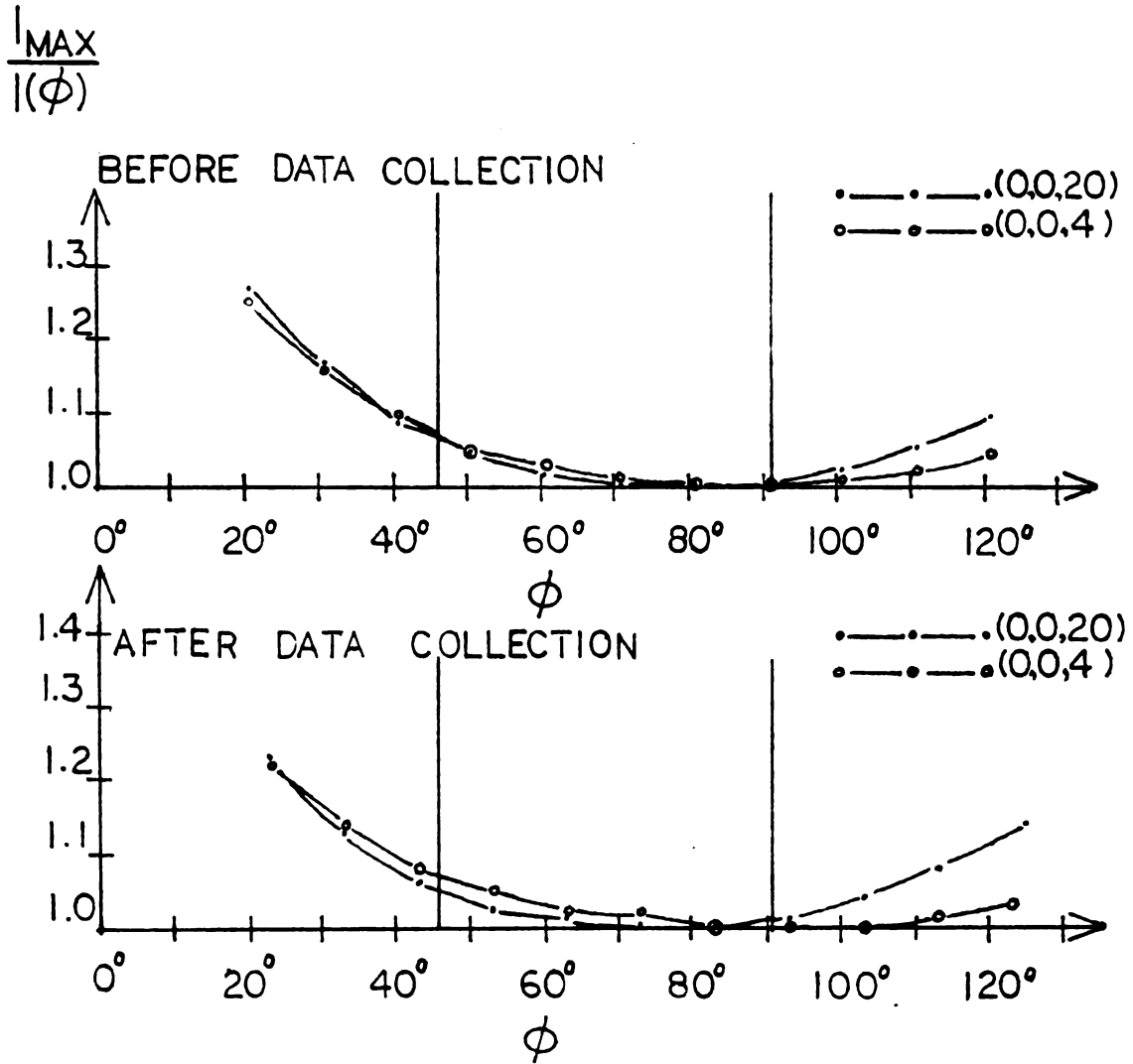


Figure 24. Absorption Curves of (0,0,4) and (0,0,20) Before and After the FlPT2 Intensity Data Collection.

TABLE 11. Absorption Table Used in the FlPT2 Data
Reduction

ϕ	$\frac{I_{\max}}{I(\phi)}$
20.40	1.26
30.40	1.16
40.40	1.10
50.40	1.05
60.40	1.02
70.40	1.00
80.40	1.00
90.40	1.00
100.40	1.02
110.40	1.04
120.40	1.06

and (2,0,23) are readily evaluated, since s and I_{CR} can be obtained from Figure 20; these values are shown in Table 12. The S values for the non-monitor reflections are simply obtained by interpolation using the monitor S values. Since the intensity data for every reflection (hkl) also contains the exposure hours on the crystal at its time of measurement, the DEC factor for a high angle reflection is easily determined by these means. For low angle data decay correction, the S value is calculated from an estimate of $I^\circ(t)/I(t)$ and a knowledge of time. The $I^\circ(t)/I(t)$ is estimated by finding the mean ratio of corresponding Patterson peak heights using the 5.9Å (hk0) intensity data before and after the data collection. The estimation of $I^\circ(t)/I(t)$ is shown in Table 13, and it was found to be $1.4 \pm 0.03(\sigma)$. Knowing that the time between the intensity measurement of the two (hk0) reflection files was 35.8 hours, an S value of $1.07 \times 10^{-3} \text{ hr}^{-1}$ was determined. However, this S value is greater than the S value determined for the $2\theta = 17.62$ monitor reflection (14,0,7), which had an S value of $7.13 \times 10^{-4} \text{ hr}^{-1}$. Since the decay rate for reflections in the 2θ range $(2.5-15)^\circ$ is never larger than that of higher 2θ reflections the S value used in the $(2.5-15.0)^\circ$ data reduction was $7.13 \times 10^{-4} \text{ hr}^{-1}$.

The factor CONST was used to scale the FlPT2 derivative intensity data to the FlN6 native intensity data. A value for CONST was obtained by finding the mean ratio of corresponding Patterson peak heights for the FlN6 and

TABLE 12. Decay Correction "S" Factors for the Three Monitor Reflection

<u>Reflection</u>	<u>2θ Value</u>	<u>S</u>
(14,0,7)	17.62	$7.13 \times 10^{-4} \text{ hr}^{-1}$
(7,7,17)	21.09	$1.28 \times 10^{-3} \text{ hr}^{-1}$
(2,0,23)	24.18	$3.26 \times 10^{-3} \text{ hr}^{-1}$

TABLE 13. Estimate of $\frac{I^{\circ}(t)}{I(t)}$ for Low Angle Decay Corrections

<u>Patterson Peak</u>	<u>Peak Height Before Data Collection</u>	<u>Peak Height After Data Collection</u>	<u>Height Before Height After</u>
1	190	189	1.00
2	131	123	1.06
3	181	178	1.02
4	156	146	1.07
5	161	155	1.04
6	162	161	1.01
7	192	176	1.09
8	212	209	1.01
9	142	139	1.02
10	252	255	0.99
11	174	160	1.09
12	269	255	1.05

Mean Ratio: $1.04 \pm 0.03(\sigma)$

FlPT2 using the (hk0) intensity data at 5.9Å resolution taken before each data collection. The table containing the resulting ratios is shown in Table 14, and the resulting scale factor was $1.01 \pm 0.04(\sigma)$.

Before the P-DATA program was run, the background of the intensity data was averaged. Background measurements made at the left and right side of a reflection peak and the total intensity scan are used to calculate the reported intensity from the equation:

$$I(\text{intensity}) = \left[\text{total scan count} - \frac{\text{sum of background counts}}{\text{background to scan ratio}} \right] \times \text{scan rate} . \quad (4)$$

Since the intensity of a given reflection is calculated by a formula which is dependent on background measurements, it is definitely useful to average the left and right background data in blocks of 2θ , and then 2θ and ϕ in an attempt to improve the intensity data quality.

The background readings were first averaged in shells of 2θ . The size of the shells used in this averaging was in the approximate range of 200 reflections—varying from 166 to 240. The 2θ shells used in the background averaging, the size of these shells, and the resulting average background values are shown in Table 15, and the average background versus 2θ curve is shown in Figure 25. As Figure 25 indicates, there was a definite 2θ dependence to the background measurements.

TABLE 14. Calculating the F1N6-F1PT2 Scale Factor for the P-DATA Program

<u>Peak</u>	<u>Peak Height for F1N6</u>	<u>Peak Height for F1PT2</u>	<u>F1N6 F1PT2</u>
1	190	190	1.00
2	124	131	0.95
3	184	181	1.02
4	201	192	1.05
5	143	142	1.01
6	185	174	1.06
7	265	269	0.98

Mean Ratio: $1.01 \pm 0.04(\sigma)$

TABLE 15. Averaging the Background of the FlPT2 Intensity Data in 2 θ Shells

<u>Shell 2θ Range</u>	<u>Average 2θ in Range</u>	<u>Number of Reflections</u>	<u>BCKG</u>
7.0-10.5	8.75	166	37.0
10.5-13.0	11.75	193	29.9
13.0-15.0	14.0	206	27.9
15.0-16.5	15.75	189	32.0
16.5-18.0	17.25	232	34.6
18.0-19.0	18.5	171	37.6
19.0-20.0	19.5	196	42.0
20.0-20.8	20.4	192	43.9
20.8-21.7	21.25	223	46.2
21.7-22.5	22.1	202	47.4
22.5-23.4	22.95	240	49.2
23.4-24.2	23.8	207	50.1
24.2-25.1	24.65	236	51.6

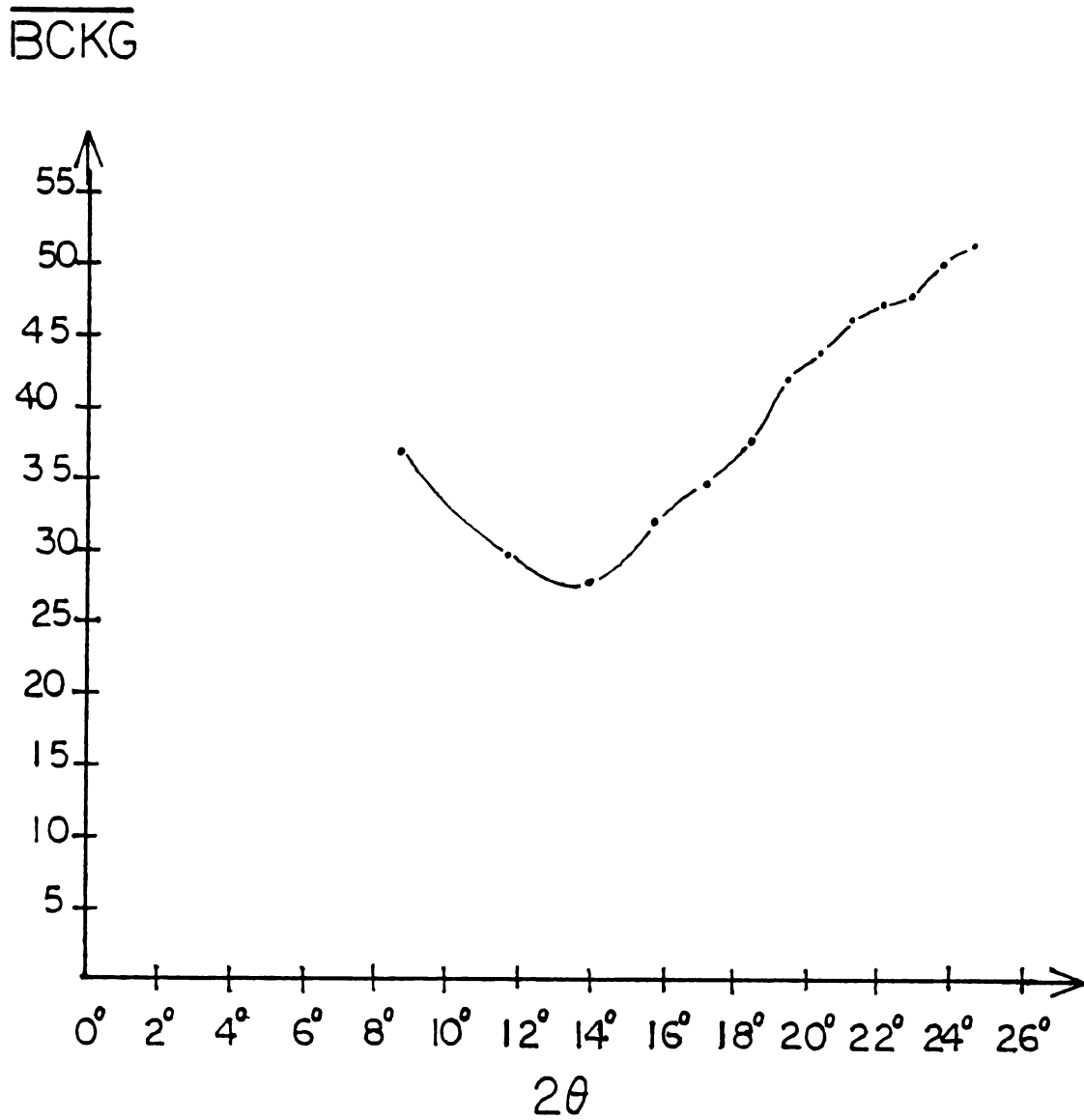


Figure 25. The Average Background of the FlPT2 Intensity Data vs. 2θ .

To see if there was both a ϕ as well as a 2θ dependence to the average background, a second background averaging was performed using the 2θ shells chosen previously and three ϕ intervals of equal size. Although data were collected from 46° - 91° in ϕ , the ϕ ranges were extended 6° on either side; the three ϕ ranges chosen were: $(40-59)^\circ$, $(59-78)^\circ$, and $(78-97)^\circ$. As can be seen from Table 16, there was no definite trend in $\overline{\text{BCKG}}$ with respect to 2θ and ϕ , so the intensity data resulting from the 2θ shell background averaging were used for the P-DATA data reduction.

Another step which must be taken before the intensity data are reduced by P-DATA is to decide on the minimum intensity which is considered "observable". In the P-DATA program, negative intensities and intensities less than the I_{\min} are set to $I_{\min}/2$ for structure factor modulus calculations. From previous work performed in this laboratory, 2-3 times the average corrected negative intensity reported in the 2θ background averaging is usually a good estimate. The averaged corrected negative intensity in the FlPT2 data set was -11.7. Considering this value and examining the chart recorder response to intensity measurements, I_{\min} was chosen to be 25.

B. The FlPT2-FlN6 Difference Patterson Vector Map

In 1935, Patterson published an important paper in which he showed that while the Fourier synthesis in Equation (1) gives peaks corresponding to the distribution

TABLE 16. Results of Averaging the Background of the FLPT2 Data with Respect to 2θ and ϕ

2θ Range	ϕ Range	$\overline{\text{BCKG}}_{2\theta}$	$\overline{\text{BCKG}}_{2\theta, \phi}$	$\frac{ \overline{\text{BCKG}}_{2\theta, \phi} - \overline{\text{BCKG}}_{2\theta} }{ \overline{\text{BCKG}}_{2\theta} } \times 100$
7.0-10.5	40-59		31.7	14.3%
	59-78	37.0	39.8	7.6%
	78-97		39.5	6.8%
10.5-13.0	40-59		26.4	11.7%
	59-78	29.9	31.2	4.3%
	78-97		32.2	7.7%
13.0-15.0	40-59		25.4	9.0%
	59-78	27.9	28.4	1.8%
	78-97		29.9	7.2%
15.0-16.5	40-59		30.0	6.3%
	59-78	32.0	33.2	3.8%
	78-97		32.9	2.8%
16.5-18.0	40-59		33.5	3.2%
	59-78	34.6	34.4	0.6%
	78-97		36.4	5.2%
18.0-19.0	40-59		34.0	9.6%
	59-78	37.6	37.0	1.6%
	78-97		35.2	6.4%
19.0-20.0	40-59		40.0	4.8%
	59-78	42.0	42.2	0.5%
	78-97		43.7	4.0%
20.0-20.8	40-59		42.5	3.2%
	59-78	43.9	44.8	2.1%
	78-97		44.2	0.7%

Table 16 Continues.

Table 16 Continued.

2θ Range	ϕ Range	$\overline{\text{BCKG}}_{2\theta}$	$\overline{\text{BCKG}}_{2\theta, \phi}$	$\frac{ \overline{\text{BCKG}}_{2\theta, \phi} - \overline{\text{BCKG}}_{2\theta} }{ \overline{\text{BCKG}}_{2\theta} } \times 100$
20.8-21.7	40-59		43.8	5.2%
	59-78	46.2	46.9	1.5%
	78-97		48.0	3.9%
21.7-22.5	40-59		44.5	6.1%
	59-78	47.4	48.9	3.2%
	78-97		48.7	2.5%
22.5-23.4	40-59		46.8	4.9%
	59-78	49.2	48.6	1.2%
	78-97		52.6	6.9%
23.4-24.2	40-59		46.8	6.6%
	59-78	50.1	50.9	1.6%
	78-97		52.2	2.2%
24.2-25.1	40-59		47.2	8.5%
	59-78	51.6	51.6	0.0%
	78-97		55.9	1.4%

of atoms in a crystal, a Fourier synthesis using the square of the structure factor amplitudes of the form

$$P(uvw) = \frac{2}{V} \sum_{h=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} |F|^2 \cos 2\pi(hu+kv+lw) \quad (5)$$

produces peaks corresponding to all of the interatomic vectors⁶². This function is called the Patterson function. Thus a peak at the point (u,v,w,) implies that there exists atoms at x_1, y_1, z_1 and x_2, y_2, z_2 such that $u = x_1 - x_2$, $v = y_1 - y_2$, $w = z_1 - z_2$. If there are N atoms in the unit cell of a molecule then the Patterson map will possess N^2 peaks, of which N vectors can be drawn to each of the N atoms, including one vector of zero length from each atom to itself. The N self-vectors are located at the origin so that there are $N^2 - N$ non-origin peaks in the Patterson map.

If the positions of heavy atom binding sites in two derivative crystals and the 3-D intensity data of the native protein and derivative crystals are available, the phases necessary to produce an electron density map could be calculated via multiple isomorphous replacement. If a heavy atom derivative contains the compound binding at specific loci in the protein throughout the crystal, it is possible to determine these positions by means of a Patterson map of the heavy atom derivative. However, such a determination is complicated and very impractical because of the overwhelming number of peaks contributed by the atoms of the protein. A much more practical way

to seek out the heavy atom binding sites is to examine a difference Patterson vector map. This map uses $\left| |F_{P+H}| - |F_P| \right|^2$ as the coefficients for the Patterson synthesis, where $|F_{P+H}|$ is the structure factor modulus of the derivative crystal and $|F_P|$ is the structure factor modulus of the native protein. The difference Patterson map essentially removes protein-protein vector contributions so that basically only heavy atom-protein and heavy atom-heavy atom vectors remain. Moreover, because the heavy atom structure factor amplitudes are so much larger than those of the individual atoms of the protein, the heavy atom-heavy atom peaks dominate the map.

In order to subtract protein-protein contributions in the difference Patterson as thoroughly as possible, the structure factor amplitudes of the native and the derivative crystal should be on the same scale. This is done by fitting the $|F|^2$ distribution of the derivative data to the $|F|^2$ distribution of the native protein. An $|F|^2$ distribution is basically a plot of the $\langle |F|^2 \rangle$ over an interval versus the average 2θ value of that interval. The $|F|^2$ distribution of FlN6 is shown in Figure 26. The data were divided into 14 2θ ranges in the distribution. To scale the distribution of FlPT2 to that of FlN6, two parameters were fitted according to the equation

$$|F_P|^2 = K |F_{P+H}|^2 \exp \left[2B \left(\frac{\sin \theta}{\lambda} \right)^2 \right] \quad (6)$$

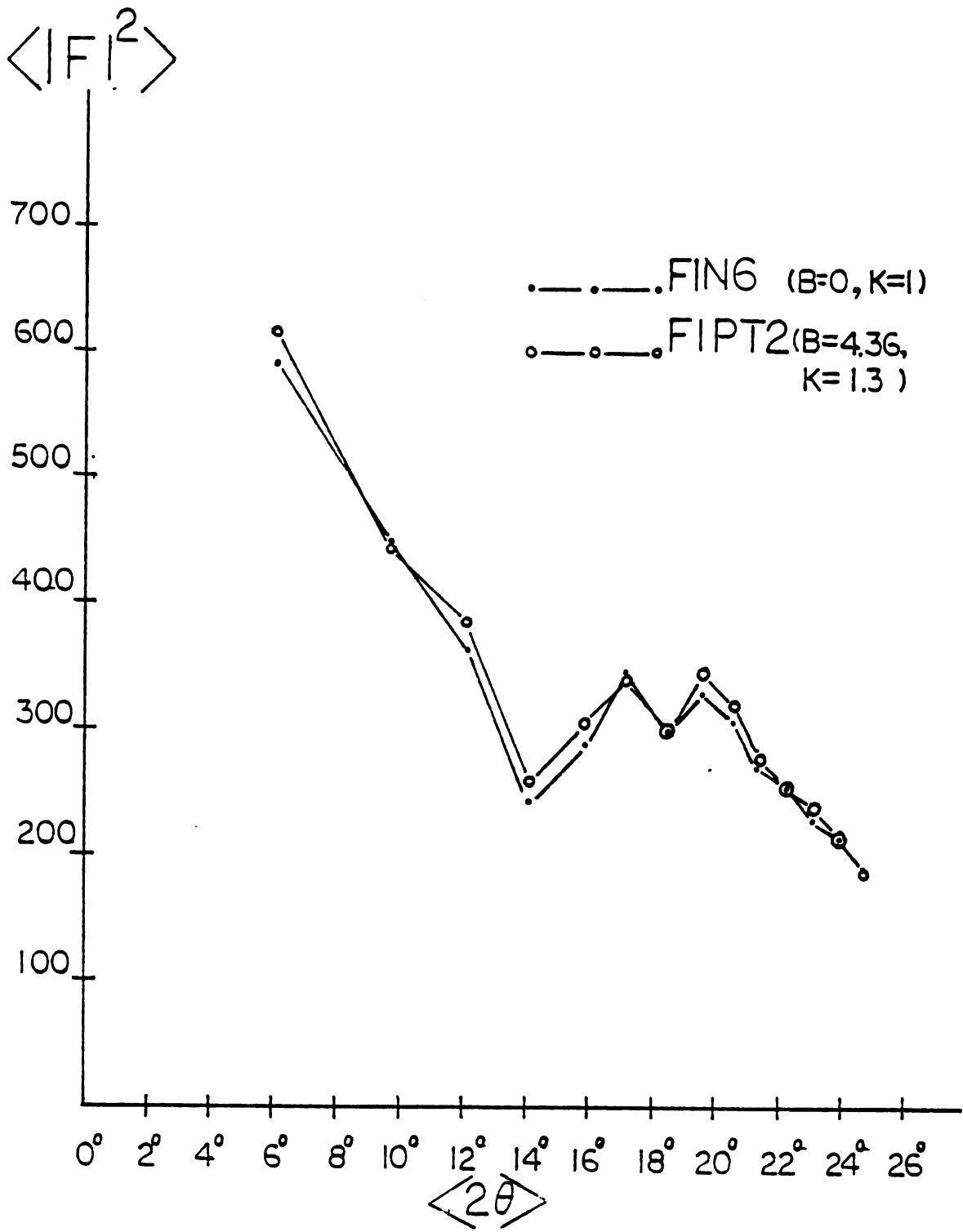


Figure 26. The $|F|^2$ Distribution for F1N6 (B = 0, K = 1) and for F1PT2 (B = 4.36, K = 1.3).

The constant K is a scale factor and is originally estimated from $\langle |F|^2 \rangle_{F1N6} / \langle |F|^2 \rangle_{F1PT2}$ for $2\theta < 14^\circ$. The variable B is called an average isotropic temperature factor and it corrects for differences in the thermal vibrations or disorder between molecules in the different crystals. After a series of manipulations, the F1PT2 $|F|^2$ distribution was brought in close agreement with the F1N6 distribution using a scale factor of 1.3 and a B value of 4.36 (Figure 26).

To calculate a difference Patterson map using the F1PT2 and F1N6 intensity data at 3.5 Å resolution, a file containing the structure factor amplitude differences between the scaled derivative and the native protein was created, $(\sqrt{K}|F_{P+H}| \exp\{B(\sin \theta/\lambda)^2\} - |F_P|)$. The R factor between the two components on the file was 12.5%.

The symmetry of the difference Patterson map is P4/mmm where the four-fold axis is along \bar{c} , with 3 mutually perpendicular mirror planes and a diagonal mirror between \bar{a} and \bar{b} . Due to these symmetry elements, it is only necessary to examine a map which extends from 0 to $\bar{a}/2$, 0 to $\bar{b}/2$, and 0 to $\bar{c}/2$ which is cut along the \bar{a}/\bar{b} diagonal. The axes in Patterson space which correspond to \bar{a} , \bar{b} and \bar{c} are \bar{u} , \bar{v} , and \bar{w} . The full axial lengths of the Patterson map were chosen to be 80, 80, 80 so that the Patterson map which was studied could be produced as 40 planar \bar{u} , \bar{v} sections perpendicular to \bar{w} , and so the distance between the \bar{u} , \bar{v} sections corresponded to approximately

1 Å (0.97 Å/grid unit along \bar{u} and \bar{v} , 1.06 Å/grid unit along \bar{w}).

If a heavy atom compound has truly bound to a unique site (or sites) in the protein with minimal distortion to the protein, then the coordinates of the site (or sites) should be obtainable from an examination of the difference Patterson. Basically, for a site to be accountable, it is necessary to find a set of (x,y,z) coordinates which satisfy the unique set of vectors relating equivalent positions in the unit cell. In the space group of the fragment 1 crystal— $P4_12_12$ or its enantiomer—there are 8 equivalent positions in the unit cell given by:

$$\begin{array}{llll} x, y, z & -x, -y, 1/2+z & 1/2-y, 1/2+x, 1/4+z & 1/2+y, 1/2-x, 3/4+z \\ y, x, -z & -y, -x, 1/2-z & 1/2-x, 1/2+y, 1/4-z & 1/2+x, 1/2-y, 3/4-z \end{array}$$

By taking all possible combinations of these coordinates 64 vectors were obtained. The symmetry elements of the $P4/mmm$ system imply that only 6 of these vectors are unique. The six independent vectors of the FlPT2-FlN6 difference Patterson map are listed in Table 17.

Certain sections or planes of a Patterson map contain vectors due to the symmetry elements relating atoms of different molecules in the unit cell. In an isomorphous heavy atom derivative of a protein, such vectors in the difference Patterson map are due to heavy atom substituents. These sections are called Harker sections⁶³. Moreover,

TABLE 17. The Unique Vectors Relating Equivalent Positions in the Unit Cell of the F1PT2-F1N6 Difference Patterson

<u>Vector</u>	<u>u</u>	<u>v</u>	<u>w</u>
1	-2x	-2y	$\frac{1}{2}$
2	$\frac{1}{2} - (x+y)$	$\frac{1}{2} + (x-y)$	$\frac{1}{4}$
3	(y-x)	(x-y)	-2z
4	-(x+y)	-(x+y)	$\frac{1}{2} - 2z$
5	$(\frac{1}{2} - 2x)$	$\frac{1}{2}$	$(\frac{1}{4} - 2z)$
6	$\frac{1}{2}$	$(\frac{1}{2} - 2y)$	$(\frac{3}{4} - 2z)$

one unique peak can be expected in a Harker section for each unique heavy atom binding site. In the fragment 1 crystal system there is a four-fold screw axis along \bar{c} , and two-fold screw axes along \bar{a} and \bar{b} . Thus, one sizeable and unique vector peak should be found at $w = 1/4$, $w = 1/2$, $u = 1/2$, and $v = 1/2$ for each heavy atom bound to a protein molecule in the lattice. An examination of the $w = 1/4$ Harker section (Figure 27) and the $w = 1/2$ Harker section (Figure 28) reveals a single unique peak above the u/v diagonal in each plane. Since \bar{a} and \bar{b} are identical in the unit cell of fragment 1, \bar{u} and \bar{v} are identical in the difference Patterson map and either the $u = 1/2$ or the $v = 1/2$ Harker section contains both the $u = 1/2$ and $v = 1/2$ Harker peaks. An examination of the $u = 1/2$ or $v = 1/2$ Harker section (Figure 29) reveals two prominent and unique peaks. All of these factors suggest that a single heavy atom substitution has occurred in the protein.

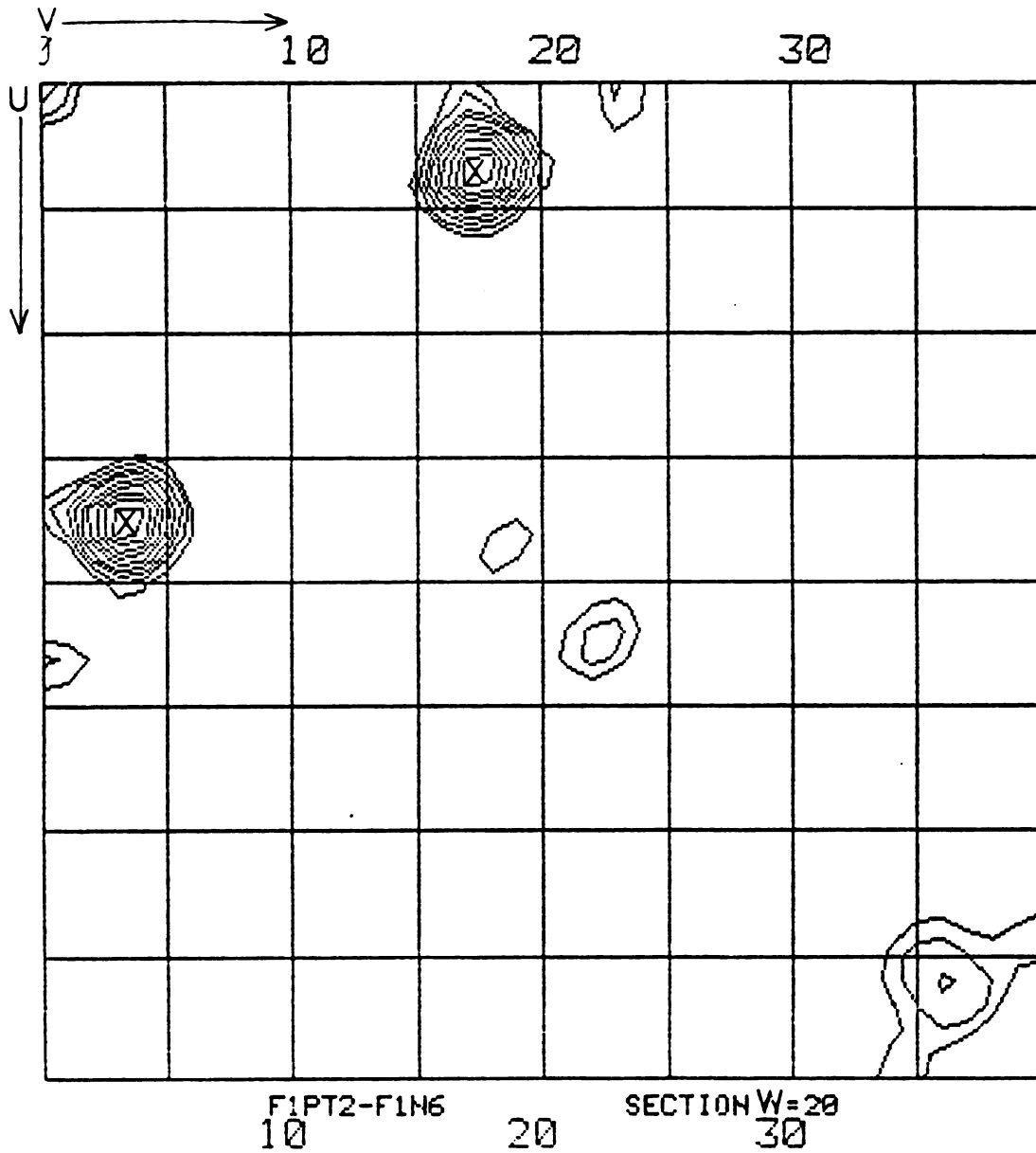


Figure 27. $w = \frac{1}{4}$ Harker Section of F1PT2-F1N6 Difference Patterson Map ('X' indicates Harker vector).

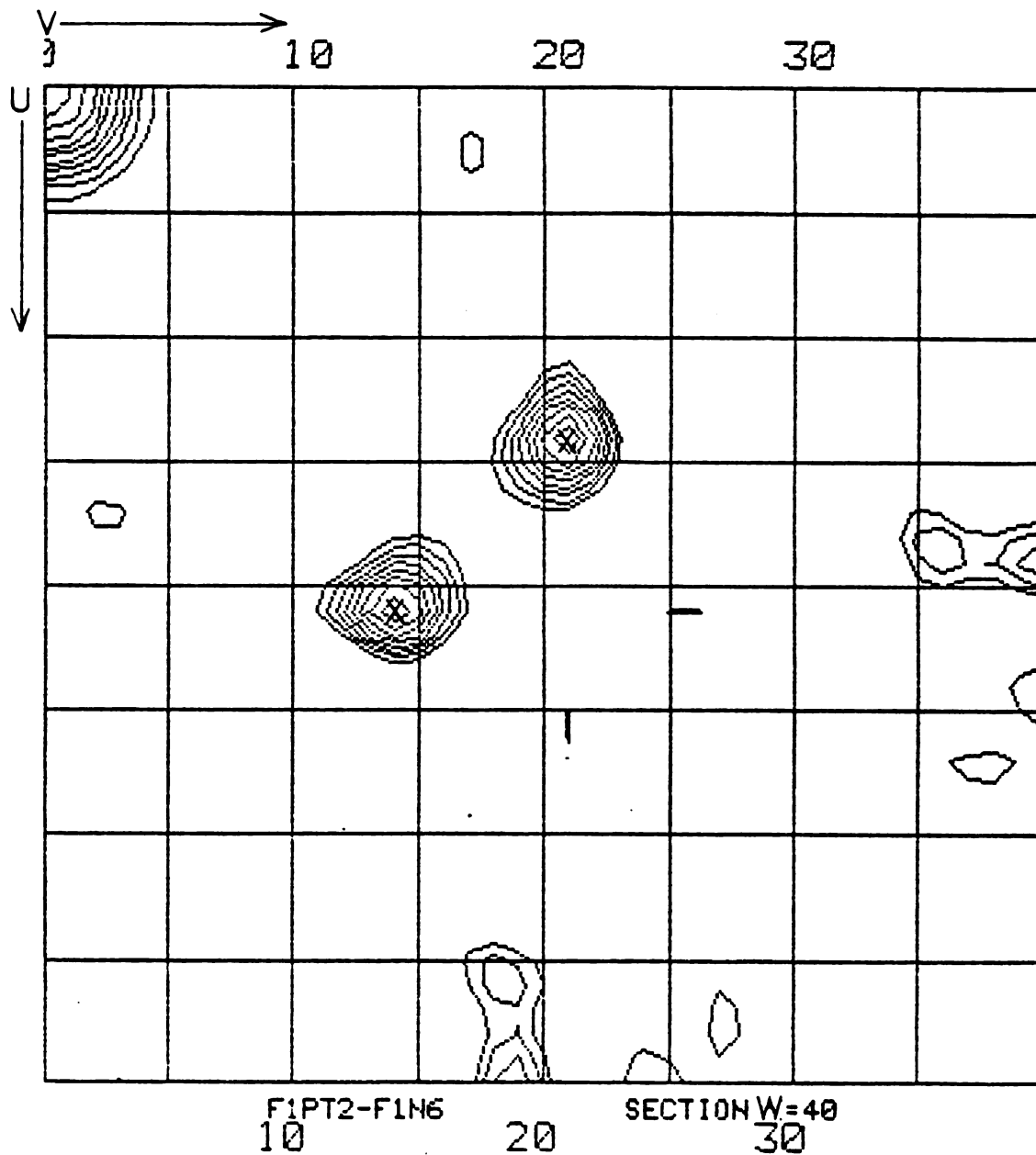


Figure 28. $w = \frac{1}{2}$ Harker Section of F1PT2-F1N6 Difference Patterson Map ('x' indicates Harker vector).

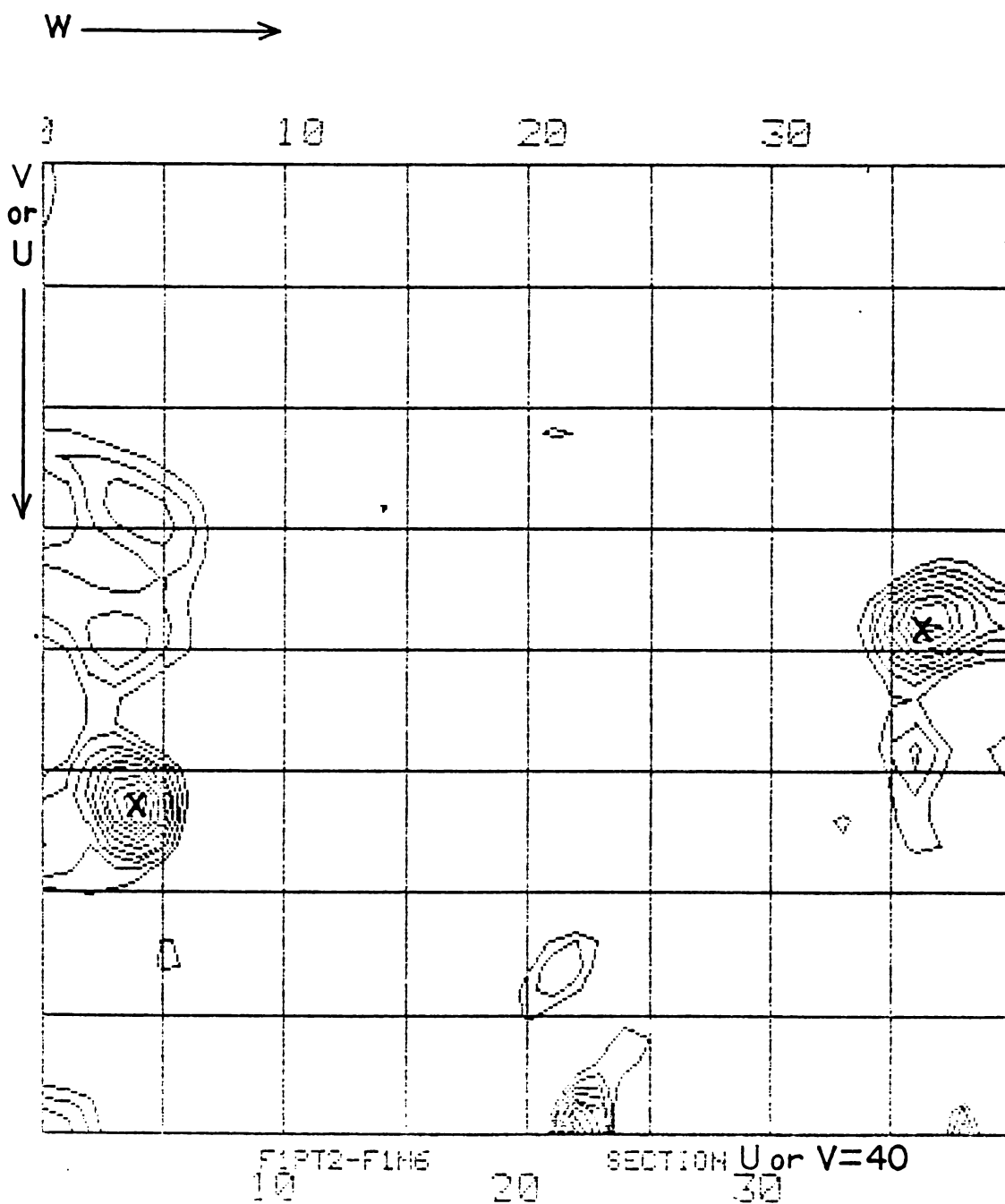


Figure 29. $u = \frac{1}{2}$ or $v = \frac{1}{2}$ Harker Section of F1PT2-F1N6 Difference Patterson Map ('X' indicates Harker vector).

CONCLUSIONS

A careful examination of the F1PT2-F1N6 difference Patterson revealed that only one platinum binding site in the protein is consistent with the map. The fractional coordinates of the platinum position are $\bar{x} = 0.089 \pm 0.001$, $\bar{y} = 0.369 \pm 0.001$, and $\bar{z} = 0.101 \pm 0.001$. The veritable absence of unaccountable major peaks in the Harker sections (Figures 27, 28 and 29), and in the remainder of the difference map suggests that very few protein structural changes occurred due to the binding of the platinum compound. This prominent indication of isomorphism between the native and the derivative protein data suggests that the platinum derivative should provide good phase information.

At the present time there are two other heavy atom derivative crystals of fragment 1 whose difference Patterson maps have been interpreted in this laboratory. A 10 mM HgAc₂ derivative of fragment 1 was consistent with one heavy atom substitution, and a mixed derivative of fragment 1 (saturated Pt(C₅H₅N)₂Cl₂ + 10 mM HgAc₂) contained a platinum binding site identical to the one located in this study, and a mercury binding site identical to the one located in the 10 mM Hg⁺² derivative. Because the Patterson vector peaks

in the FlPT2 derivative (contoured from 75 to 300 in increments of 25) are about half the height of the Patterson peaks in the 10 mM Hg⁺² derivative (contoured from 150 to 500 in increments of 50), if the occupancy of mercury in the mercury derivative is considered to be 100%, the occupancy in the platinum derivative is ~70%. Thus, soaking a crystal in saturated K₂Pt(C₅H₅N)₂Cl₂ for 3 weeks to a month may affect greater changes, and result in a higher platinum occupancy at the substitution site.

Two other research groups are presently working on the structure of fragment 1. Aschaffenburg et al.,⁴¹ at Oxford have grown crystals of fragment 1 which are of the same crystal system (tetragonal) and space group (P4₁2₁2 or its enantiomer P4₃2₁2) as the crystals of this laboratory, with $\bar{a} = \bar{b} = 78.87 \text{ \AA}$ and $\bar{c} = 84.99 \text{ \AA}$, but which are obtained from 2 M phosphate or saturated ammonium sulfate, pH 6.0-7.5. The crystals are reported to be small and poor scatterers of x-rays. The group has reported 6 Å resolution data from native and derivative crystals, but no heavy atom positions have been obtained from these data collections⁶⁵. The fact that these crystals contain phosphate prevents Ca⁺² binding studies to be carried out since calcium phosphates would precipitate.

Olsson et al.,⁴² report that large orthorhombic crystals of space group P2₁2₁2₁ with $\bar{a} = 39.5 \text{ \AA}$, $\bar{b} = 54.0 \text{ \AA}$, and $\bar{c} = 129.0 \text{ \AA}$ were grown out of 8% (w/v) PEG 6,000 at pH 6.0-7.0 in the presence of 100 mM Ca⁺² using 100 mM

sodium cacodylate as a buffer. The fact that they can obtain crystals at such a high Ca^{+2} ion concentration is puzzling because Lundblad and his co-workers⁶⁴ have reported that fragment 1 should be extensively dimerized at such a concentration, and the group's crystal data are not consistent with the formation of dimers. Using 2.8 Å resolution data from the native protein and two platinum derivatives, a 4 Å electron density map was calculated. However, the map is of dubious quality since four of the five heavy atom binding sites are in special positions (a heavy atom bound at a special position makes a contribution to the structure factors of only certain classes of reflections), and since the figure of merit⁴⁷ of the map is only 0.54. (The figure of merit of a protein structure is a measure of how certain the phase angles are known.) Moreover, they were only able to build a balsa wood model of the protein and thus they can make no statements about the folding of the main chain and, more specifically, the kringle loop.

The derivative data at 3.5 Å resolution and the heavy atom positions of FlPT2 and the Hg and Hg-Pt derivatives, along with the native protein data at 3.5 Å resolution, should allow a good quality electron density map to be calculated from which the kringle folding and main chain folding can be followed. In addition, we have also grown crystals of deglycosylated fragment 1, which are isomorphous with fragment 1. Using phases calculated for native

fragment 1, the deglycosolated variant structure factor amplitudes can be assigned phases and a difference electron density map showing the location of the sugar residues can be calculated.

LIST OF REFERENCES

LIST OF REFERENCES

1. A.C.T. North and D.C. Phillips, "Progress in Biophysics and Molecular Biology", Vol. 19, Pergamon Press (Oxford), 1969.
2. D. Harker, *Acta Cryst.* 9, 1 (1956).
3. M.G. Rossman, *Acta Cryst.* 14, 383 (1961).
4. K.G. Mann and M.R. Downing, "The Chemistry and Biology of Thrombin", (R.L. Lundblad, J.W. Fenton, II, K.G. Mann, eds.), Ann Arbor Science (Ann Arbor), 1977.
5. M.B. Zuker, *Sci. Am.* 242, 86 (1980).
6. E.W. Davie and K. Fujikawa, *Ann. Rev. Biochem.* 44, 799 (1975).
7. C.M. Jackson and Y. Nemerson, *Ann. Rev. Biochem.* 49, 765 (1980).
8. K.G. Mann, C.M. Hildelbrandt, and D.N. Fass, *J. Biol. Chem.* 246, 6106 (1971).
9. A. Cox and D.J. Hannah, *Biochim. Biophys. Acta* 207, 49 (1971).
10. M.E. Nesheim, L.S. Hibbard, P.B. Tracy, J.W. Bloom, K.H. Myrmel, and K.G. Mann, "The Regulation of Coagulation", (K.G. Mann, F.B. Taylor, Jr., eds.), Elsevier-North Holland, Inc. (New York), 1980.
11. K.G. Mann, *Methods in Enzymol.* 45, 123 (1976).
12. W.H. Seegers, "The Chemistry and Biology of Thrombin", (R.L. Lundblad, J.W. Fenton, II, K.G. Mann, eds.), Ann Arbor Science (Ann Arbor), 1977.
13. S. Magnusson, T.E. Petersen, L. Sottrup-Jensen, and H. Claeys, "Proteases and Biological Control", (J.F. Davidson, R.M. Rowan, M.M. Samama, and P.C. Desnoyers, eds.), Vol. 3, Raven Press (New York), 1978.

- 13a. M. Trexler and L. Patthy, Proc. Nat'l. Acad. Sci. USA 80, 2457 (1983).
14. J. Stenflo and J.W. Suttie, Ann. Rev. Biochem. 46, 157 (1977).
15. H. Claeys, L. Sottrup-Jensen, M. Zajdel, T.E. Petersen, and S. Magnusson, FEBS Lett. 61, 20 (1976).
16. L. Sottrup-Jenson, H. Claeys, M. Zajdel, T.E. Petersen, and S. Magnusson "Progress in Chemical Fibrinolysis and Thrombolysis", (J.F. Davidson, R.M. Rowan, M.M. Samama, and P.C. Desnoyers, eds.), Vol. 3, Raven Press (New York), 1978.
17. G.L. Nelsestuen, T.H. Zytkovitz, and J.B. Howard, J. Biol. Chem. 249, 6347 (1974).
18. J. Stenflo, P. Fernland, W. Egan, and P. Roepstoriff, Proc. Nat'l. Acad. Sci. 71, 2730 (1974).
19. D. Bucher, E. Negelin, J. Thomsen, and J. Stenflo, FEBS Lett. 68, 293 (1976).
20. J. Stenflo and J.W. Suttie, Ann. Rev. Biochem. 46, 157 (1977).
21. D.V. Shah and J.W. Suttie, Proc. Nat'l. Acad. Sci. 68, 1653 (1971).
22. J.W. Suttie, Science 179, 192 (1973).
23. G.L. Nelsestuen, J. Biol. Chem. 251, 5648 (1976).
24. G.L. Nelsestuen, M. Broderius, and G. Martin, J. Biol. Chem. 251, 6886 (1976).
25. F.G. Prendergast and K.G. Mann, J. Biol. Chem. 252, 840 (1977).
26. J.W. Bloom and K.G. Mann, Biochemistry 17, 4430 (1978).
27. B. Furie, K.L. Provost, R.A. Blanchard, and B.C. Furie, J. Biol. Chem. 253, 8980 (1978).
28. R. Sperling, B.C. Furie, M. Blumenstein, B. Keyt, and B. Furie, J. Biol. Chem. 253, 3898 (1978).
29. P. Esnouf, International Workshop on Regulation of Coagulation (Abst.) Norman, Oklahoma (1979).

30. G. Ashwell and A.G. Morell, "Glycoproteins of Blood Cells and Plasma", (G.A. Jamieson and T.T. Greenwalt, eds.), Lippincott (Philadelphia), 1971.
31. T. Mizuochi, K. Yamashita, K. Fujikawa, W. Kisiel, and A. Kobata, J. Biol. Chem. 254, 6419 (1979).
32. J.D.G. Smit, Journal de Chimie Physique 76, 9 (1979).
33. A. McPherson, Jr. "Preparation and Analysis of Protein Crystals", John Wiley and Sons (New York), 1982.
34. M.N. Liebman, Ph.D. Thesis, Michigan State University, (1977).
35. P.D. Martin, A. Tulinsky, and F.G. Walz, Jr., J. Mol. Biol. 136, 95 (1980).
36. M. Zeppezauer, H. Eklund, and E.S. Zeppezauer, Arch. Biochem. Biophys. 126, 564 (1968).
37. R.L. Vandlen, D.L. Ersfeld, A. Tulinsky, and W.A. Wood, J. Biol. Chem. 248, 2251 (1973).
38. B.R. Reid, G. Koch, Y. Boulanger, B. Hartley, and D. Blow, J. Mol. Biol. 80, 199 (1973).
39. S.H. Kim, G. Quigley, F.L. Suddath, and A. Rich, Proc. Nat'l. Acad. Sci. USA 68, 841 (1971).
40. W.H. Kung, A. Tulinsky, and G.L. Nelsestuen, J. Biol. Chem. 255, 10523 (1980).
41. R. Aschaffenburg, C.C.F. Blake, J.M. Burridge, and M.P. Esnouf, J. Mol. Biol. 114, 575 (1977).
42. G. Olsson, L. Andersen, O. Lindqvist, L. Sjölin, S. Magnusson, T. Petersen, L. Sottrup-Jensen, FEBS Lett. 45, 2 (1982).
43. G. Gomri, "Methods in Enzymology", Vol. 1, Academic Press (New York), 1955.
44. H.J. Kung and A. Tulinsky, unpublished observation of this laboratory.
45. M.V. King, Acta Cryst. 7, 601 (1954).
46. R.G. Pearson, J. Am. Chem. Soc. 85, 3533 (1963).
47. T.L. Blundell and L.N. Johnson, "Protein Crystallography", Academic Press (New York), 1976.

48. G.E. Schulz, M. Elzinga, F. Marx, and R.H. Schirmer, *Nature (London)* 250, 120 (1974).
49. E.T. Adman, L.C. Sieker, and L.H. Jensen, *J. Biol. Chem.* 248, 3987 (1973).
50. P.M. Colman, L.H. Weaver, and B.W. Matthews, *Biochem. Biophys. Res. Comm.* 46, 1999 (1972).
51. M.M. Bloom, G. Bodo, H.M. Dintzis, and J.C. Kendrew, *Proc. Roy. Soc. Lond.* A246, 369 (1958).
52. A. Mavridis, A. Tulinsky, and M.N. Liebman, *Biochem.* 13, 3661 (1974).
53. M.E. Scott, M.M. Sarasuo, H.C. Marsh, D.L. Harris, R.G. Hiskey, and K.A. Koehler, *J. Am. Chem. Soc.* 102, 3413 (1980).
54. G.L. Nelsestuen, *J. Biol. Chem.* 251, 3235 (1976).
55. B.A. Lewis, J. Freyssinet, and J.J. Holbrook, *Biochem. J.* 169, 397 (1978).
56. G.L. Nelsestuen, private communication.
57. P.B. Sigler, *Biochem.* 9, 3609 (1970).
58. N.M. Henry, K. Lonsdale, "International Tables for X-Ray Crystallography", Vol. 1, The Kynoch Press (Birmingham), 1952.
59. G.H. Stout and L.H. Jensen, "X-Ray Structure Determination—A Practical Guide", MacMillan Publishing Co., Inc. (New York), 1968.
60. H.W. Wyckoff, M. Doscher, D. Tsernoglou, T. Inagami, L.N. Johnson, K.D. Hardman, N.M. Allewell, D.M. Kelly, and F.M. Richards, *J. Mol. Biol.* 27, 563 (1967).
61. A.C.T. North, D.C. Phillips, and F.S. Mathews, *Acta Cryst.* A24, 351 (1968).
62. A.L. Patterson, *Z. Krist.* A90, 517 (1935).
63. C.R. Cantor and Paul R. Schimmel, "Biophysical Chemistry", Part II, W.H. Freeman and Co. (San Francisco), 1980.
64. R.C. Tarvers, C.M. Noyes, H.R. Roberts, and R.L. Lundblad, *J. Biol. Chem.* 257, 18 (1982).
65. J.M. Burridge, C.C.F. Blake, R. Aschaffenburg, and M.P. Esnouf, "The Regulation of Coagulation", Elsevier-North Holland, Inc. (New York), 1980.

MICHIGAN STATE UNIV. LIBRARIES



31293107057006