STUDIES ON A NEUROTOXIC PROTEIN: THE 5 Å STRUCTURE OF α-BUNGAROTOXIN

Thesis by

Steven A. Spencer

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1978

(Submitted November 14, 1977)

Acknowledgments

I wish to thank my friends and colleagues--Roger Koeppe, Monty Krieger, Mike Ross, Russ Timkovich, John Chambers, Jerry Tobler, Mike Klymkowsky, Dave McKay, Dave Agard and Mel Jones--for their helpful suggestions during the course of this work and for their patience during my frequent periods of ill humor. I especially with to thank Tony Kossiakoff, whose unique grasp of reality helped me to retain a modicum of sanity during my years of graduate study.

I am grateful to Carol Timkovich and Jessie Koeppe for their moral support and for the many excellent (and unrepaid) meals they served me.

My thanks to Dee Barr for her help and patience during my stay at Caltech. Her extraordinary efforts in typing this thesis are gratefully acknowledged.

I wish to thank the staff of the Chemistry Division for their help during my graduate career. Special thanks go to the following people: Instrument Shop--Bill Schuelke, Delmer Dill, Guy Duremberg and Tony Stark; Mechanical Shop--Jim Olson, Lou Borbon, Fern Masse and Tom Menkel. Without their skills, good nature and patience, my task would have been considerably more difficult.

My thanks to the National Institutes of Health for financial support during most of my graduate studies.

Finally, I wish to thank my advisor, Robert Stroud, without whose support this work would not have been possible.

Abstract

The neurotoxic venom from the Formosan Banded Krait, Bungarus multicinctus, was fractionated into nineteen components by ion-exchange and gel-filtration chromatography. The thirteen main protein components were run on SDS-polyacrylamide gels. Six of these components produced banding patterns similar to α-bungarotoxin (α-BuTX), while four produced banding patterns similar to β-bungarotoxin. Three components did not fit either pattern.

Crystallization studies of α-BuTX were conducted, and crystals were obtained throughout the pH range 2.5 to 7.5. However, only near pH 4 were crystals routinely obtained which were large enough to allow the α-BuTX structure to be determined by x-ray diffraction methods.

The pH 4 crystals contained four molecules in the asymmetric unit. The unit cell dimensions were 69.9 x 76.7 x 44.8 Å. The crystals appeared to orthorhombic, but space group identification was difficult because the ℓ = odd reflections were generally weak throughout reciprocal space. These weak reflections resulted from the presence within the unit cell of two subcells, each of which had approximately $P2_12_12_1$ symmetry. This feature of the crystals considerably complicated the structure determination process. Further complications arose from the presence within the asymmetric unit of a noncrystallographic symmetry axis.

The α-BuTX structure was solved at 5 Å resolution by the method of multiple isomorphous replacement. Except for a long tail at the

carboxyl terminus, the toxin molecule fit within an ellipsoid approximately 40 x 28 x 21 Å. The main features of the electron density map were interpreted by comparison to the structure of another neurotoxin, erabutoxin b.

Table of Contents

Table of Tables

CHAPTER I

Introduction

"We, the unwilling,
are led by the unknowing
to do the impossible
for the ungrateful.
We have done so much
with so little for so long,
that we are now qualified
to do anything with nothing."

--Author unknown

Historical Background[1]

Interest in snake venoms undoubtedly dates back to the first encounter between mankind and these venomous reptiles; but until recently, this interest has been confined largely to discovering ways to neutralize the venoms or otherwise minimize their toxic effects. Indeed, some of the earliest medical writings--Egyptian papyri dating back to about 1600 B.C.--described "cures" for various insect and animal bites. However, such remedies were not sufficiently good to save Cleopatra, perhaps the most famous victim of snakebite, who used an asp--believed to be the Egyptian Cobra, Naja haje (Pollard, 1956)--to commit suicide in 30 B.C.

Systematic studies of venoms did not appear until well into the Renaissance period. The first of these, Deux Livres des Venins Anvers, was written by Jaques Grevin in Paris and published in 1568. In these two volumes he described both plant and animal venoms. Not until 1664 did a study devoted entirely to snake venoms appear. This work, entitled Osservazioni Intorno Alle Vipere, was written by Francesco Redi. He proved that, for the snakes he had studied, the venoms were not toxic if taken orally, but had to be injected to be lethal.

The first truly biochemical studies of snake venoms did not occur until the middle of the 19th century. In 1860, the American Silas

---

[1] For a more detailed account of the development of knowledge about plant and animal venoms, see the article by Leake (1956), from which most of this historical introductory material was obtained.

Wier Mitchell published a study of rattlesnake venom called Researches Upon the Venom of the Rattlesnake, and in 1886 he published a more general work along with E. T. Reichert entitled Researches Upon the Venoms of Poisonous Serpents. In these studies, it was first demonstrated that snake venoms contained proteins. The effects of these venoms on blood and nerve tissue were also described.

Immunological studies of snake venoms were first conducted in the 1890's by Leon C. A. Calmette at the Pasteur Institute in Paris. He obtained antivenom sera and studied the possibility of immunizing animals against snakebite. Also, about this same time, Sir Thomas R. Fraser in Edinburgh produced an antivenom directed specifically against cobra venom.

With the advent of modern biochemical techniques for purifying proteins and fractionating protein mixtures, the emphasis has shifted from studying whole venoms to studying the individual venom components. Considerable effort has been expended in attempting to learn the mechanism of these toxins, not only for the purpose of counteracting their effects, but also to allow them to be used as probes in the study of various biological processes. It is one class of these toxins, the neurotoxins, which is the subject of this thesis.

Neurotoxic Snake Venoms

The neurotoxic snake venoms which have been studied the most have been obtained from snakes in the families *Elapidae* (cobras, kraits) and *Hydrophiidae* (sea snakes), although some neurotoxic venoms from *Viperidae* (vipers) and *Crotalidae* (pit vipers, rattlesnakes) have been examined. These venoms have been found to contain one or more of the following neurotoxic proteins: 1) "α-type" toxins, such as α-bungarotoxin (*Bungarus multicinctus*) and cobrotoxin (*Naja naja atra*), which produce a nondepolarizing postsynaptic block at the neuromuscular junction by binding to the acetylcholine receptor (Chang and Lee, 1963); 2) "β-type" toxins, such as β-bungarotoxin, which block the release of acetylcholine at the presynaptic membrane but do not affect postsynaptic transmission (Chang and Lee, 1963); and 3) cardiotoxins, which have been reported to act on sodium-potassium ATPase (Vincent *et al.*, 1976), and are therefore not strictly neurotoxins. In addition, various enzymatic activities have been found in these venoms (see, for example, Tu, 1973), but most of these are not directly related to their toxic action.

α-Toxins: The α-type toxins are small polypeptides (60-74 amino acids) linked together by four or five disulfide bonds. One of the members of this group, α-bungarotoxin (α-BuTX) was the subject for the x-ray structural study described in this thesis.

This class of toxins has received considerable attention in recent years because of their tight, specific binding to the nicotinic acetylcholine receptors found in mammalian neuromuscular synapses and in the electric organs of electric fish (Chang and Lee, 1963; Changeux, Kasai and Lee, 1970; Lester, 1970; Miledi, Molinoff and Potter, 1971; Raftery, Schmidt and Clark, 1972). These groups and others have used the α-toxins both as affinity ligands to purify the receptor and as an assay for it. The properties of these toxins are described in greater detail in the next section.

β-Toxins: The β-type toxins are quite different from the α-toxins. The β-toxin which has been studied most extensively, β-bungarotoxin, contains about 191 amino acids arranged in two polypeptide chains which are linked together by disulfide bridges (Kelly and Brown, 1973). It is not known how many of the ten disulfide bonds (Lee *et al.*, 1972) are involved in these interchain bonds. It is interesting that a second β-toxin, crotoxin, which is found in the venom of the Brazilian rattlesnake (*Crotalus durissus terrificus*), also contains two polypeptide chains; but in this case they are not crosslinked (Hendon and Fraenkel-Conrat, 1971).

Both β-bungarotoxin and crotoxin have been shown to possess phospholipase A activity (Strong *et al.*, 1976 and Hendon and Fraenkel-Conrat, 1971, respectively). No sequence for either toxin has yet been published, but it has been reported (B. D. Howard, personal communication) that the larger chain of β-bungarotoxin has a sequence nearly the same as that of phospholipase A. This result agrees with the finding of Horst, Hendon and Fraenkel-Conrat (1972) that the phospholipase activity of crotoxin resides on the larger subunit.

The mechanism by which the β-toxins affect the release of acetylcholine has not yet been discovered, although it has been found that β-bungarotoxin inhibits oxidative phosphorylation in mitochondria (Wernicke, Vanker and Howard, 1975). Exactly how this would prevent the presynaptic release of acetylcholine is not clear, although depletion of the presynaptic energy stores could conceivably interfere with the release mechanism.

The function of the phospholipase activity is also not known. It has recently been shown that the toxic activity of β-bungarotoxin can be destroyed without destroying the phospholipase activity (Howard and Truong, 1977), thus demonstrating that this activity alone is not sufficient to explain its toxicity. The opposite case has not yet been shown, however, thereby suggesting that both the phospholipase activity and the small polypeptide chain are intimately involved in the toxic mechanism. These results are in agreement with the findings of Hendon and Fraenkel-Conrat (1971), who found that the large subunit of crotoxin alone displayed hemolytic activity while the combination of the large and small subunits was neurotoxic. No activity has ever been associated with the small subunit alone.

Cardiotoxins: The cardiotoxins are similar to the α-toxins in certain ways. Cardiotoxin from cobra venom (Naja naja atra) is a short, basic polypeptide (60 amino acids) crosslinked by four disulfide bonds (Lee, 1972). It is less lethal than the α-toxins (Vincent _et_ _al_., 1976), but does show some sequence homology to them (Lee, 1972). It has been suggested that the α-toxins evolved from an ancestor molecule similar to the cardiotoxins (Strydom, 1973).

The α-Toxins

Properties

The α-toxins are small, highly basic, and extremely stable proteins. The basicity of these proteins (Table 1-1) results from the large number of lysine and arginine residues which they contain. The stability is probably a result of the high degree of crosslinking produced by the presence of four or five disulfide bonds

* * * * * * *

TABLE 1-1

Isoelectric points of some α-toxins

Toxins	pI	Source
Erabutoxin _a_	9.15	Tu, Hong and Solie (1971)
Erabutoxin _b_	9.34	Tu, Hong and Solie (1971)
Erabutoxin _c_	9.24	Tamiya and Abe (1972)
^{125}I-α-bungarotoxin	9.19	Clark _et_ _al_. (1972)
Naja _nivea_ toxins α, β, δ	>8.4	Botes _et_ _al_. (1971)
Cobrotoxin	~9.1	Chen, Lo and Yang (1977)

* * * * * * *

in such short polypeptides.

As examples of the stability of these toxins, cobrotoxin retained 25% activity after heating at 100° C for 30 minutes (Yang, 1965). Erabutoxins *a* and *b* (*Laticauda semifasciata*) retained full activity under the same conditions, as well as over the pH range of 1 to 11 (Tu, Hong and Solie, 1971). In addition, cobrotoxin remained fully active in 8 M urea (Yang, 1967), and *Naja haje* toxin I retained full activity when treated in anhydrous formic acid or 1 N HCl for 100 minutes (Chicheportiche *et al.*, 1972).

Sequence comparisons

A large number of α-toxins have been sequenced (see, for example, Tu, 1973, and Yang, 1974), and it has become clear that they can be further subdivided into two classes (Tu, 1973): 1) The Type I or short toxins containing 60-62 amino acids and four disulfide bonds; and 2) the Type II or long toxins comprised of 71-74 amino acids and crosslinked by five disulfide bonds. Sequence comparisons between these two types of toxins (Table 1-2) show that they are homologous, and probably arose from a common ancestor (see Strydom, 1973). The main differences between them are the addition in the long toxins of both an "extra" disulfide bridge linking residues 29 and 33 (α-BuTX numbering) and about seven amino acid residues at the carboxyl terminus.

The many sequences of both long and short toxins from various species have revealed a number of residues which are conserved or

TABLE 1-2

Sequences for several α-type neurotoxins[1]

Sequence number[2]:			4								10										20
1) α-Bungarotoxin[3] (*Bungarus multicinctus*)	I	V	C	H	-	T	T	A	T	I	P	S	S	A	V	T	C	P	P	G	E
2) *Naja nivea* toxin α[5]	I	R	C	F	-	-	-	I	T	P	D	V	T	S	Q	A	C	P	D	G	-
3) *Dendroaspis polylepis* toxin γ[6]	R	T	C	-	N	-	-	K	T	F	S	D	Q	S	K	I	C	P	P	G	E
4) *Naja melanoleuca* toxin *b*[7]	I	R	C	F	-	-	-	I	T	P	D	V	T	S	Q	I	C	A	D	G	-
5) Cobrotoxin[8] (*Naja naja atra*)	L	E	C	H	N	Q	Q	S	S	Q	T	P	T	T	T	G	C	S	G	G	E
6) Erabutoxin *b*[9] (*Laticauda semifasciata*)	R	I	C	F	N	Q	H	S	S	Q	P	Q	T	T	K	T	C	P	S	G	E
7) *Naja haje* toxin α[10]	L	Q	C	H	N	Q	Q	S	S	Q	P	P	T	T	K	T	C	P	-	G	E
8) Cardiotoxin[11] (*Naja naja atra*)	L	K	C	-	N	-	-	K	L	V	P	L	F	Y	K	T	C	P	A	G	K

TABLE 1-2 (continued)

										30										40										50								
1)	N	L	C	Y	R	K	M	W	C	D	A	F	C	S	S	R	G	K	V	V	E	L	G	C	A	A	T	C	P	S	K	K	P	Y	E	E	V	T
2)	H	V	C	Y	T	K	M	W	C	D	N	F	C	G	M	R	G	K	R	V	D	L	G	C	A	A	T	C	P	K	V	K	P	G	V	N	I	K
3)	N	I	C	Y	T	K	T	W	C	D	A	W	C	S	Q	R	G	K	R	V	E	L	G	C	A	A	T	C	P	K	V	K	A	G	V	E	I	K
4)	H	V	C	Y	T	K	T	W	C	D	N	F	C	A	S	R	G	K	R	V	D	L	G	C	A	A	T	C	P	T	V	K	P	G	V	N	I	K
5)	T	N	C	Y	K	K	R	W	R	D	-	H	-	-	-	R	G	Y	R	T	E	R	G	C	G	-	-	C	P	S	V	K	N	G	I	E	I	N
6)	S	S	C	Y	H	K	Q	W	S	D	-	F	-	-	-	R	G	T	I	I	E	R	G	C	G	-	-	C	P	T	V	K	P	G	I	K	L	S
7)	T	N	C	Y	K	K	R	W	R	D	-	H	-	-	-	R	G	S	I	T	E	R	G	C	G	-	-	C	P	S	V	K	K	G	I	E	I	N
8)	N	L	C	Y	K	M	F	M	V	A	T	-	-	-	-	P	K	V	P	V	K	R	G	C	I	D	V	C	P	K	S	S	L	V	L	K	Y	V

TABLE 1-2 (continued)

		60										70				
1)	C	C	S	T	D	K	C	N	H	P	P	K	R	Q	P	G
2)	C	C	S	R	D	N	C	N	P	F	P	T	R	K	R	S
3)	C	C	S	T	D	D	C	D	K	F	Q	F	G	K	P	R
4)	C	C	S	T	D	N	C	N	P	F	P	T	R	N	R	P
5)	C	C	T	T	D	R	C	N	N	-	-	-	-	-	-	-
6)	C	C	E	S	E	V	C	N	N	-	-	-	-	-	-	-
7)	C	C	T	T	D	K	C	N	N	-	-	-	-	-	-	-
8)	C	C	N	T	D	R	C	N	-	-	-	-	-	-	-	-

TABLE 1-2 (continued)

[1] The IUPAC one-letter amino acid code is used.

[2] Residue numbers refer to the α-bungarotoxin sequence.

[3] Mebs *et al.* (1972).

[4] Conserved residues are surrounded by solid boxes, while functionally conserved residues are surrounded by dashed boxes. Residues conserved among either the long or short toxins are also in a box. More sequences were used to make this comparison than are shown in this table. For a more complete set of sequences, see Tu (1973).

[5] Botes (1971).

[6] Strydom (1972).

[7] Botes (1972).

[8] Yang, Yang and Huang (1969).

[9] Sato and Tamiya (1971). The positions of Ser 21 and Glu 20 are reversed from the original sequence results. This change is based on the results of Tsernoglou and Petsko (1977).

[10] Botes and Strydom (1969).

[11] Cardiotoxin is not an α-toxin. Its sequence, from Narita and Lee (1970), is included for reference only.

conservatively substituted (Table 1-2). Several of these (Cys, Pro, Gly) are of obvious structural importance, but the remainder are probably involved with the mechanism of toxicity. As discussed in the section on "Chemical Modifications" (see below), some, but not all, of these "nonstructural" conserved residues have been shown to be necessary for activity.

Binding to the acetylcholine receptor

Both long and short neurotoxins have been reported to have dissociation constants from the acetylcholine receptor of around 10^{-9} to 10^{-10} M (Chicheportiche *et al.*, 1975; Maelicke and Reich, 1975). These values of K_d are considerably smaller than that of *d*-tubocurarine ($K_d = 6.6 \times 10^{-7}$ M as determined by Raftery *et al.*, 1975), a small-molecule neurotoxin with a toxic action similar to that of the α-toxins. Such tight binding suggests that the α-toxins conform closely to the receptor surface around the acetylcholine binding site.

There is some evidence that the long and short neurotoxins react differently with the receptor despite their similar dissociation constants. In the case of the *Naja haje* toxins, the short toxin (Toxin I) binds more rapidly and dissociates more rapidly than the long toxin (Toxin III) (Chicheportiche *et al.*, 1975). This result is consistent with the observation that erabutoxin and cobrotoxin produce neuromuscular blocks which are slowly reversible, whereas α-bungarotoxin appears to bind irreversibly to the receptor (Lee, 1972). It also indicates that the structural differences between the long and short toxins have a significant effect upon their activity.

Chemical modifications[2]

Disulfide bonds: The five disulfide bonds in α-bungarotoxin are shown in Figure 1-1. It was shown in the case of one long toxin (*Naja haje* toxin III) that the "extra" disulfide bond present in long toxins, but not in short ones (Cys 29 to Cys 33 in Figure 1-1), could be reduced without significant loss of activity (Chicheportiche *et al.*, 1975). The other four disulfide bridges were found to be essential for the activity of cobrotoxin (Yang, 1967).

It was also found that totally reduced cobrotoxin could regain full activity after 72 hours of air oxidation (Yang, 1967). In this study, it was observed that only 23% activity was regained when 72% of the disulfide bonds had reformed. This result indicates that all four bonds must be intact to obtain activity[3].

Amino groups: The α-amino group of cobrotoxin was shown to be unimportant for activity by selective modification with trinitrobenzene sulfonate (TNBS) (Chang *et al.*, 1971b). These workers also found that conversion of the lysines to homoarginine by reaction

---

[2] This section contains a short summary of the results of some modifications. For a more complete description, see Tu (1973) and Yang (1974)

[3] With 72% of the disulfide bonds formed, about 27% of the molecules would have all four bonds intact, assuming each bond forms independently of the others. This value compares favorably with the 23% activity observed.

with O-methylisourea did not affect activity. This result was confirmed with erabutoxins a and b, in which all but one of their lysines were modified with the same reagent (Tu, Hong and Solie, 1971).

Modification of the lysine residues of cobrotoxin with a reagent which did not preserve the positive charge (TNBS) completely abolished activity (Chang et al., 1971b). It was found in this study that Lys 26[4] could be modified with TNBS with no loss of activity, but once both Lys 26 and Lys 52 were modified, activity was completely destroyed. This result is surprising, since Lys 26 is conserved in all small and most long toxins. Since no α-toxin has yet been found which lacks both Lys 52 and Lys 26, it is possible that either Lys 26 or Lys 52 is sufficient for activity. It is not clear why both of these residues are conserved in all small toxins, but one or the other can be omitted in the long ones.

Arginine: One of three arginine residues in erabutoxin a was modified with 1,2-cyclohexanedione without reducing its activity (Tu, Hong and Solie, 1971). The modified residue was not identified. In cobrotoxin Arg 27 was modified by reaction with phenylglyoxal without loss of activity, but modification of the conserved residue, Arg 36, destroyed activity (Yang, 1974).

---

[4] α-Bungarotoxin sequence numbering used throughout. See Table 1-2.

Histidine: His 4 of cobrotoxin was photo-oxidized without causing a loss in activity, but similar modification of His 31 greatly reduced activity (Yang, 1974). Since all α-toxins have an aromatic ring in this position, it is probable that the photo-oxidation of His 31 disrupted the tertiary structure in this region and that this, not the loss of the imidazole functionality, produced the loss of activity.

Carboxyl groups: Six of the seven carboxyl groups in cobrotoxin were modified by carbodiimide coupling with glycine methyl ester without loss of activity (Chang et al., 1971b). The seventh carboxyl group, Glu 20, was unreactive until the toxin was placed in 5 M guanidine-HCl. Modification of this final carboxyl group destroyed activity.

The fact that Glu 20 is unreactive suggests that it is "buried" and is possibly involved in an ion-pair bond with some other residue. Since this residue is semi-conserved (an Asp or Glu is almost always present on one side or the other of Gly 19), it may have some structural function.

It is surprising that Glu 41, which is functionally conserved in both the long and short toxins, could be modified without affecting activity. Since neither asparagine nor glutamine appears at this position, it is unlikely that the mere presence of the amide bond to the glycine ester is sufficient to replace the carboxyl group. It would seem, therefore, that either this residue is not involved in the toxic reaction and is conserved instead for some other purpose (such as the folding of the toxin), or that the glycine ester is somehow capable of substituting for the carboxyl functionality.

Tyrosine: The conserved tyrosine, Tyr 24, was shown to be relatively unreactive in cobrotoxin, and is presumably buried within the molecule (Chang *et al.*, 1971a). These same workers also showed that nitration of this residue with tetranitromethane in guanidine-HCl destroyed the toxin's activity, but that nitration of the second tyrosine in cobrotoxin had no effect. A similar result was found for *Naja haje* toxin I (Chicheportiche *et al.*, 1972).

Tryptophan: All α-toxins contain an invariant Trp 28. A number of workers have tried various means of modifying this residue, most of which destroyed the activity of the toxins (see Tu, 1973). In one case (*Naja haje* toxin I), the formation of N-formyl tryptophan only reduced activity about 50% (Chicheportiche *et al.*, 1972). The formyl group was then removed, but activity was not restored. This result suggests that nonspecific side reactions produced the loss in activity, and that modification of the ring nitrogen had no effect. This means that the nitrogen is probably not necessary for toxicity, and that the bulky ring itself, not its functionality, is what is important.

α-Bungarotoxin

α-Bungarotoxin (α-BuTX) is one of the neurotoxic components in the venom of the Formosan Banded Krait (*Bungarus multicinctus*). It is a long, α-type toxin containing 74 amino acids and five disulfide bridges. It is, to this author's knowledge, the only long α-toxin to be crystallized, and it is only the second neurotoxin--erabutoxin *b* being the first (Low *et al.*, 1976; Tsernoglou and Petsko, 1976)[5]--for which the tertiary structure has been determined by x-ray crystallography.

The structural study of this toxin was undertaken for two reasons. The first was that, at the time, no toxin structure had been solved. It was hoped that a knowledge of the three-dimensional arrangement of the amino acid residues conserved during evolution, as well as those shown by chemical means to be important, would lead to an understanding of the mechanism by which these toxins act.

The second reason for this study was that the tight and apparently irreversible binding of this toxin to the acetylcholine receptor (Chang and Lee, 1963) indicated that it must conform closely to the structure of the receptor around the acetylcholine binding

---

[5] It was originally believed that the toxin from the Japanese sea snake, solved by Low *et al.* (1976), and the toxin from the Philippines sea snake, solved by Tsernoglou and Petsko (1976), were different. This does not now appear to be the case (Tsernoglou and Petsko, 1977), and so the structures determined by these two groups are assumed to be the same.

site. It was therefore hoped that a knowledge of the toxin structure would lead to insights into the structure of the acetylcholine receptor.

At the present time, neither of these goals has been fully realized, and it was perhaps a little naive to believe that they would be. The structure of erabutoxin has contributed significantly to the first objective, but little progress has been made on the second. The structure of α-bungarotoxin will have to be solved at higher resolution before it can make a significant contribution to either of these objectives.

Despite the fact that this work has not achieved either of its major goals, it is hoped that the structure of α-bungarotoxin will eventually provide a framework on which to base further studies of the α-toxins, and that further knowledge of these highly specific molecules will lead to a better understanding of the neural transmission processes.

References

Botes, D. P. (1971), J. Biol. Chem. 246, 7383-7391.

Botes, D. P. (1972), J. Biol. Chem. 247, 2866-2871.

Botes, D. P. and Strydom, D. J. (1969), J. Biol. Chem. 244, 4147-4157.

Botes, D. P., Strydom, D. J., Anderson, C. G., and Christensen, P.A. (1971), J. Biol. Chem. 246, 3132-3139.

Chang, C. C. and Lee, C.-Y. (1963), Arch. Int. Pharmacodyn. 144, 241-257.

Chang, C. C., Yang, C. C. Hamaguchi, K., Nakai, K., and Hayashi, K. (1971a), Biochim. Biophys. Acta 236, 163-173.

Chang, C. C., Yang, C. C., Nakai, K., and Hayashi, K. (1971b), Biochim. Biophys. Acta 251, 334-344.

Changeux, J.-P., Kasai, M., and Lee, C.-Y. (1970), Proc. Nat. Acad. Sci. USA 67, 1241-1247.

Chen, Y.-H., Lo, T.-B., and Yang, J. T. (1977), Biochemistry 16, 1826-1830.

Chicheportiche, R., Rochet, C., Sampieri, F., and Lazdunski, M. (1972), Biochemistry 11, 1681-1691.

Chicheportiche, R., Vincent, J.-P., Kopeyan, C., Schweitz, H., and Lazdunski, M. (1975), Biochemistry 14, 2081-2091.

Clark, D. G., Macmurchie, D. D., Elliott, E., Wolcott, R.G., Landel, A. M., and Raftery, M.A., (1972), Biochemistry 11, 1663-1668.

Hendon, R. A. and Fraenkel-Conrat, H. (1971), Proc. Nat. Acad. Sci. USA 68, 1560-1563.

Horst, J., Hendon, R. A., and Fraenkel-Conrat, H. (1972), Biochem. Biophys. Res. Commun. 46, 1042-1047.

Howard, B. D. and Truong, R. (1977), Biochemistry 16, 122-125.

Kelly, R. B. and Brown, F. R. III (1973), J. Neurobiol. 5, 135-150.

Leake, C. D., "Development of Knowledge About Venoms", in Venoms (E. E. Buckley and N. Porges, Eds.), American Association for the Advancement of Science, Washington, D.C., 1956, pp. 1-4.

Lee, C.-Y. (1972), Ann. Rev. Pharmacol. 12, 265-286.

Lee, C.-Y., Chang, S. L., Kau, S. T., and Luh, S.-H. (1972), J. Chromatogr. 72, 71-82.

Lester, H. A. (1970), Nature 227, 727-728.

Low, B. W., Preston, H. S., Sato, A., Rosen, L. S., Searl, J. E., Rudko, A. D., and Richardson, J. W. (1976), Proc. Nat. Acad. Sci. USA 73, 2991-2994.

Maelicke, A. and Reich, E. (1975), Cold Spring Harbor Symposium on Quantitative Biology 40, 231-235.

Mebs, D., Narita, K., Iwanaga, S., Samejima, Y., and Lee, C.-Y. (1972), Hoppe-Seyler's Z. Physiol. Chem. 353, 243-262.

Miledi, R., Molinoff, P., and Potter, L. T. (1971), Nature 229, 554-557.

Narita, K. and Lee, C.-Y. (1970), Biochem. Biophys. Res. Commun. 41, 339-343.

Pollard, C. B., "Venom Research: A Challenge to the Various Sciences" in Venoms (E. E. Buckley and N. Porges, Eds.), American Association for the Advancement of Science, Washington, D.C., 1956, pp. 5-8.

Raftery, M. A., Schmidt, J., and Clark, D. G. (1972), Arch Biochem. Biophys. 152, 882-886.

Raftery, M. A., Vandlen, R. L., Reed, K. L., and Lee, T. (1975), Cold Spring Harbor Symposia on Quantitative Biology 40, 193-202.

Sato, S. and Tamiya, N. (1971), Biochem. J. 122, 453-461.

Strong, P. N., Goerke, J., Oberg, S. G., and Kelly, R. B. (1976), Proc. Nat. Acad. Sci. USA 73, 178-182.

Strydom, D. J. (1972), J. Biol. Chem. 247, 4029-4042.

Strydom, D. J. (1973), Comp. Biochem. Physiol. 44B, 269-281.

Tamiya, N. and Abe, H. (1972), Biochem. J. 130, 547-555.

Tsernoglou, D. and Petsko, G. A. (1976), FEBS Letters 68, 1-4.

Tsernoglou, D. and Petsko, G. A. (1977), Proc. Nat. Acad. Sci. USA 74, 971-974.

Tu, A. T. (1973), Ann. Rev. Biochem. 42, 235-258.

Tu, A. T., Hong, B., and Solie, T. N. (1971), Biochemistry 10, 1295-1304.

Vincent, J.-P., Schweitz, H., Chicheportiche, R., Fosset, M., Balerna, M., Lenoir, M.-C., and Lazdunski, M. (1976), Biochemistry 15, 3171-3175.

Wernicke, J. F., Vanker, A. D., and Howard, B. D. (1975), J. Neurochem. 25, 483-496.

Yang, C. C. (1965), J. Biol. Chem. 240, 1616-1618.

Yang, C. C. (1967), Biochim. Biophys. Acta 133, 346-355.

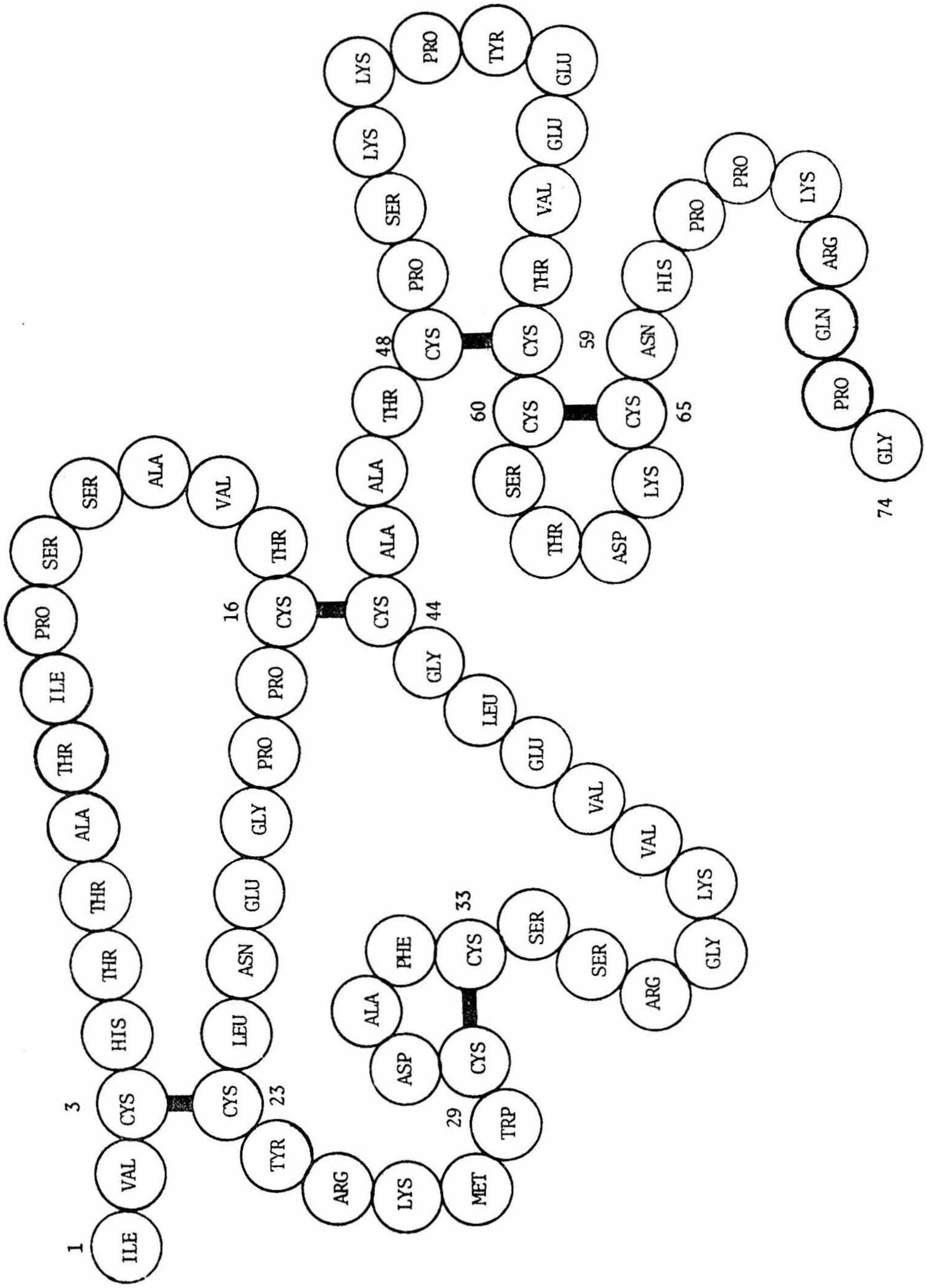
Yang, C. C. (1974), Toxicon 12, 1-43.

Yang, C. C., Yang, H. J., and Huang, J. S. (1969), *Biochim. Biophys. Acta* *188*, 65-77.

Figure Caption

Figure 1-1: The amino-acid sequence and disulfide-bonding pattern in α-bungarotoxin. The sequence is from Mebs *et al*. (1972).

Figure 1-1


1
ILE
VAL
3
CYS
HIS
THR
THR
ALA
THR
ILE
PRO
SER
SER
ALA
VAL
THR
16
CYS
PRO
PRO
GLY
GLU
ASN
LEU
23
CYS
TYR
ARG
LYS
MET
TRP
29
CYS
ASP
ALA
PHE
33
CYS
SER
SER
ARG
GLY
LYS
VAL
VAL
GLU
LEU
GLY
44
CYS
ALA
ALA
THR
48
CYS
PRO
SER
LYS
LYS
PRO
TYR
GLU
GLU
VAL
THR
59
CYS
60
CYS
SER
THR
ASP
LYS
65
CYS
ASN
HIS
PRO
PRO
LYS
ARG
GLN
PRO
GLY
74

CHAPTER II

Fractionation of Bungarovenom and Partial Characterization of Its Components

"Purify, purify, purify!"

--Attributed to Dr. Arthur Kornberg

Materials and Methods

Materials

Lyophilized venom from *Bungarus multicinctus* (BgV) was obtained from the Miami Serpentarium Laboratories, Lots BM1E and BM2ETL. The carboxymethyl cellulose was Whatman CM-52. Sephadex G-50 was obtained from Pharmacia.

The sources for electrophoresis reagents were: acrylamide, Eastman 5521; BIS[1], Eastman 8383; TEMED, Eastman 8178; SDS, Pierce Chemical Company "Sequanal Grade"; Tris, Sigma "Trizma Base"; EDTA, Baker 8993; DTT, Calbiochem 233155; pyronin Y, Allied (C.I. 45005); Aniline Blue Black (Acid Black 1), Matheson Coleman and Bell (C.I. 20470).

Molecular weight standards for SDS-polyacrylamide gel calibration were obtained from the following sources: bovine serum albumin, Sigma A4378; bovine trypsin, Worthington TRL, 2x crystallized; hen egg white lysozyme was provided by Miss Lois Kay; horse heart cytochrome *c* was donated by Dr. Richard E. Dickerson.

Initial fractionation of BgV on carboxylmethyl cellulose

Crude BgV was initially fractionated using a procedure similar to that of Clark *et al.* (1972). A typical run was as follows: 477 mg of crude, lyophilized BgV were dissolved in 20 ml 0.1 M ammonium acetate

---

[1] See Appendix A for an explanation of abbreviations.

solution (NH_4OAc), pH 5.8. This solution was introduced onto a 2.5 x 50 cm CM-52 column equilibrated with the same buffer. The column was then connected to a 2 ℓ, linear, 0.1 M to 0.4 M, pH 5.8, NH_4OAc gradient, and 7.5 ml fractions collected. The column was run at 4°C, and had an initial flow rate of 45 ml/hr.

After the first gradient had finished, the column was connected to a 2 ℓ, linear, 0.4 M to 0.8 M, pH 5.8, NH_4OAc gradient. Fractions were collected as before. The absorbance at 280 nm was then measured for every other fraction. The fractions for each peak were pooled.

Further chromatography on carboxymethyl cellulose

Each component which was to be rerun on the CM-52 column was first diluted with distilled water to reduce its salt concentration below that of the starting gradient buffer and then adjusted to the proper pH. This diluted solution was introduced onto a 2.5 x 50 cm CM-52 column which had been equilibrated at 4°C with the starting buffer. The effluent was monitored at 280 nm to insure that the protein did not wash off the column during sample application.

When all the sample was on the CM-52 column, the column was rinsed with starting buffer to re-equilibrate it. The A_{280nm} of the effluent was again monitored to guard against loss of sample. The column was then connected to the appropriate gradient and fractions collected. (Refer to the Results section for the gradients used.) The A_{280nm} of the fractions was then measured, and the fractions corresponding to the major peaks were pooled. Minor peaks were discarded.

Gel-filtration chromatography

Each component which was to be run on Sephadex G-50 was first partially lyophilized to reduce its volume to around 10 ml. This material was then layered at room temperature onto a 2.5 x 55 cm Sephadex G-50 (fine) column equilibrated with 0.1 M acetic acid. The sample was then eluted with the same solution. Unless otherwise indicated, 3 ml fractions were collected.

The A_{280nm} of the fractions was measured and the fractions corresponding to the major peaks were pooled. Minor peaks and those peaks appearing in the salt fraction were discarded.

Acetylcholinesterase assay

The method of Ellman *et al*. (1961) was used to test for acetylcholinesterase activity.

SDS-polyacrylamide gel electrophoresis

Preparation of the SDS-polyacrylamide gels and electrophoresis samples largely followed the method of Fairbanks, Steck and Wallach (1971). Table 2-1 contains a summary of the solutions used.

The main differences between the procedure used in this work and that described by Fairbanks and coworkers were: 1) the ratio of Con AcBis to water was increased to make gels containing 10% acrylamide; 2) the Con AcBis was not deionized, although the acrylamide and BIS were both recrystallized before use; and 3) the gels were overlaid with *sec*-butanol instead of an acrylamide solution. The latter

TABLE 2-1

SDS-polyacrylamide gel solutions[1]

Stock Solutions:

Con AcBis

Acrylamide (40 g)
BIS (1.5 g)
H_2O to 100 ml

10x Buffer (pH 7.4)

1.0 M Tris (40 ml)
2.0 M sodium acetate (10 ml)
0.2 M EDTA (10 ml)
Acetic acid to pH 7.4
H_2O to 100 ml

Other Solutions

20% (w/v) SDS
1.5% (w/v) ammonium persulfate
0.5% (v/v) TEMED

Electrophoresis Buffer:

10x Buffer (100 ml)
20% SDS (50 ml)
H_2O (850 ml)

Gels (10% in acrylamide)

Con AcBis (5 ml)
10x Buffer (2 ml)
20% SDS (1 ml)
H_2O (9 ml)
1.5% (w/v) ammonium persulfate (2 ml)
0.5% TEMED (1 ml)

Sample Solution:

20% (w/v) Sucrose (5 ml)
100 mM Tris-HCl + 10 mM EDTA, pH 8 (1 ml)
0.01% pyronin Y (1 ml)
20% SDS (0.5 ml)
H_2O (0.5 ml)

---

[1] The solutions and terminology are those used by Fairbanks, Steck, and Wallach (1971).

change made it much easier to produce a flat gel surface.

The gels were poured into 120 mm x 7 mm O.D. Pyrex tubes to a height of about 85 mm. After the surfaces were overlaid with sec-butanol, the gels were allowed to polymerize undisturbed for one to two hours. The gel tops were then rinsed several times with electrophoresis buffer. The gels were covered with electrophoresis buffer and stored overnight before use.

The samples were prepared by placing 0.3 - 1 mg of the sample protein in 200 µl of the sample solution. To this solution were added 1.5 mg DTT (for unreduced samples, this was omitted) and sufficient water to bring the volume to 250 µl, thereby diluting the mixture to the concentrations suggested by Fairbanks. The samples were then incubated at 37°C for 30 minutes.

The gel tubes were placed in the electrophoresis apparatus, covered with electrophoresis buffer, and 5 to 100 µl of each sample were layered on the gel surfaces. The gels were run at room temperature with a current of 4 ma/tube.

When the tracking dye was approximately 5 mm from the ends of the gels, the current was switched off. The gels were removed from the tubes by first running a stream of water around them with a hypodermic needle, and then pushing with a plunger which fit tightly inside the tubes.

The tracking dye (pyronin Y) position was marked by puncturing the gel with a scalpel, and then the gels were stained overnight in an aqueous solution containing 1% (w/v) Aniline Blue Black, 20% (v/v) isopropanol, and 20% (v/v) acetic acid. The gels were destained by transverse electrophoresis in a solution containing 20% (v/v) ethanol and 20% (v/v) acetic acid in water. Band positions were measured visually with a millimeter scale or by scanning the gels at 580 nm using a Gilford 240 spectrophotometer equipped with a linear gel transport.

Results and Discussion

Fractionation plan

The overall scheme used to fractionate BgV and to purify the components is shown in Figure 2-1. All the components appearing at the bottom of the flow chart and not labeled as discarded were lyophilized to dryness.

Initial BgV fractionation

The result of the initial fractionation of BgV on carboxymethyl cellulose is shown in Figure 2-2. The fractionation pattern was qualitatively similar to the patterns reported by other workers (Bosmann, 1972; Clark *et al.*, 1972; Eterovic *et al.*, 1975; Lee *et al.*, 1972) in that α-BuTX eluted first as a single sharp peak followed by a rounded β-BuTX peak which was incompletely separated from several other components. However, the shallower gradients used in this work resulted in a somewhat better resolution of the venom components than reported by the other workers, especially around the β-BuTX peak.

Seventeen components were separated in this initial fractionation. Several of these (I, VI, IX, XVII) contained such a small amount of material that they were discarded without further study. Component II was discarded when it was determined that it had an A_{260nm}/A_{280nm} ratio greater than one, a property which identified this fraction as guanosine. This nucleoside was found in BgV by previous workers (Wei and Lee, 1965; Lee et al., 1972).

Further purification of BgV components

Component III: The result of running Component III on Sephadex G-50 is shown in Figure 2-3. Component IIIa appeared in the void volume, indicating a molecular weight greater than 20,000 daltons[2]. Component IIIb emerged well ahead of the column volume, indicating a molecular weight in the range of 15,000-20,000 daltons.

The material in Component IIIc appeared to bind to the Sephadex, as this component emerged somewhat later than the normal salt fractions. Upon lyophilization, this component produced a minute amount of brown, powdery material, with an A_{260nm}/A_{280nm} ratio of one. It gave no reaction with ninhydrin and was discarded.

Component IV (α-BuTX): The α-BuTX eluted from the G-50 just ahead of the salt fractions (Figure 2-4). The skewing of the peak towards lower fraction numbers (higher molecular weight) was

---

[2] The column was not calibrated, so only relative molecular weights could be determined.

reproducible, and seemed to be characteristic of this toxin. This could have been due to dimerization of the toxin in the concentrated solution applied to the column, followed by gradual dissociation of the dimers as the toxin was diluted by the eluant.

This material was lyophilized and used without further purification in the crystallization experiments. The toxin was not assayed directly for toxicity, but it has been shown to bind specifically to the acetylcholine receptor (M. Klymkowsky, personal communication).

Component V: On G-50, this component produced a small peak in the void volume, which was discarded; and a single, lower molecular weight peak which contained the bulk of the material.

Components VII and VIII: These two components were only partially resolved on the CM-52 column and were therefore run together on the G-50 column (Figure 2-5). The partially resolved peaks from the G-50 run were divided at the minimum point and each rerun on G-50. Component VII produced a peak which eluted just behind the void volume and skewed towards lower molecular weight, while Component VIII eluted primarily as a single peak with a lower apparent molecular weight than Component VII. In addition, a small amount of material corresponding in size to Component VII eluted just ahead of Component VIII. This material was discarded.

Components X to XIII: These components were combined and run on the G-50 column. They produced two large, poorly-resolved peaks, one of which emerged somewhat behind the void volume and the other which emerged just ahead of the salt fractions. Because of the poor resolution obtained, the fractions were pooled and run on the CM-52 column using a linear, 1.8 ℓ, 0.6 M, pH 4.5 to pH 5.8, NH_4OAc gradient (Figure 2-6). 6.9 ml fractions were collected.

Components A and B were individually run on the G-50 column. Component A produced a single peak, while Component B contained both a high and a low molecular weight material. The low molecular weight peak from B contained the bulk of the protein and was the only one which was saved.

Components C and D were combined and rerun on the CM-52 column using a 2 ℓ, linear, pH 7.0, 0.2 M to 0.5 M NH_4OAc gradient (Figure 2-7). 7.9 ml fractions were collected. The first component which eluted from this column (Component 1) bled off the column to some extent, and so it was rerun on the CM-52 using a linear, 1 liter, pH 7.0, 0.15 M to 0.3 M NH_4OAc gradient. This procedure resulted in a single, symmetric peak for this material.

The two components from the CM-52 column (Component 1 and Component 2) were individually run on the G-50 column. Component 1 emerged just ahead of the salt fractions, whereas Component 2 emerged just behind the void volume. Cross contamination between these two components was minimal.

Component XIV (β-BuTX): This material was rerun on the CM-52 column using a linear, 2 ℓ, pH 7.0, 0.2 M to 0.6 M NH_4OAc gradient (Figure 2-8). The major component was dialyzed against distilled water and lyophilized.

Components XV and XVI: These components each eluted from the G-50 column just behind the void volume, but had a shoulder of lower molecular weight material. They were not purified further.

Acetylcholinesterase assay

Lee et al (1972) reported that the acetylcholinesterase activity in BgV eluted from their ion exchange column slightly ahead of the α-BuTX peak. Based on this, the component most likely to contain the esterase activity--Component IIIa--was assayed. Unfortunately, most of the material from Component IIIa failed to redissolve and was undoubtedly denatured. No esterase activity was found.

The purified α-BuTX also showed no esterase activity. The crude venom had an esterase activity of about 18 μmoles hydrolyzed/min/mg of dry venom (Ellman et al., 1961). No other components were assayed for activity.

SDS-polyacrylamide gel electrophoresis

The major BgV components were run on SDS gels, and gave the results shown in Figure 2-9. Figure 2-10 contains a tabulation of the gel bands, along with measured R_F values relative to the pyronin Y band and approximate molecular weights. The latter are only rough estimates, since in most cases internal standards were not run with the samples. This is especially true below 10,000 daltons, where the plot of log (molecular weight) versus R_F deviated significantly from a linear relationship. Differences in R_F value of 0.03 or less between different gels are not significant.

β-BuTX: Reduced and unreduced β-BuTX were run along with several internal molecular weight standards to determine the size of the parent protein and its subunits. The parent (unreduced) molecule ran somewhat ahead of trypsin, and gave an apparent molecular weight of 20,000 daltons. The large subunit ran slightly faster than lysozyme and about the same as horse heart cytochrome c, thus indicating a molecular weight of around 12,000 to 13,000 daltons. The small subunit ran slightly behind α-BuTX, and had an apparent molecular weight of around 8,000 to 9,000 daltons. The sums of the molecular weight limits for the subunits indicate that the parent molecule should be between 20,000 and 22,000 daltons, in good agreement with the measured value.

Staining: The Aniline Blue Black stain produced several different colors which were characteristic of various protein species. The unreduced proteins usually stained blue-black, whereas the reduced proteins stained black. Reduced β-BuTX and related toxins were unusual, in that the low molecular weight subunit stained blue, while the high molecular weight subunit stained black. In general, the reduced proteins stained more intensely than the unreduced ones.

Further Discussion

Venom composition

Based on the gel patterns, the BgV components were divided into three categories: 1) "α-type" toxins--those components giving a single band around 8,000 daltons, whether reduced or not; 2) "β-type" toxins--those components producing one band around 20,000 daltons when unreduced and two bands around 12,000 and 8,000 daltons when reduced; and 3) those fractions not displaying either pattern. The gel patterns in Figure 2-10 are labeled with these three categories.

The result of this analysis was that the BgV contained six α-type toxins and four β-type toxins. Lee *et al.* (1972), using an assay for the type of neuromuscular block produced, found six β-type toxins (one of which--Component XVII--was discarded in this work) and four α-type toxins. The shallower gradient used to fractionate the venom in the current work may have resolved some α-type components not separated by Lee and coworkers, hence increasing the number of α-type toxins observed.

The absence of one β-type toxin in the present work, compared to the results of Lee and coworkers, was somewhat disturbing. The "missing" toxin may have been overlooked due to the overlap of two β-type components, or the presence of a β-type toxin in one of the unidentified components. However, the low toxicity of one β-type toxin (Fraction IV_1) produced by Lee _et al_. suggested that this "toxin" may have been an artifact.

β-BuTX molecular weight

The measured molecular weight of 20,000 to 22,000 daltons for β-BuTX is lower than the value of 28,500 daltons reported previously by Lee _et al_. (1972). It is, however, consistent with a more recent determination by Kelly and Brown (1973), who, on the basis of amino acid composition, gel filtration, and sedimentation results, reported a molecular weight of 21,800 daltons. Their values of 8,800 and 12,400 daltons for the subunits are also consistent with the values of 8,000 to 9,000 daltons and 12,000 to 13,000 daltons reported here.

Unidentified Components

Component IIIa contained two proteins of around 100,000 daltons, and obviously did not fit the pattern set by either class of toxin. Most likely, this fraction contained one or several of the enzymes whose presence in BgV was described by Lee and coworkers (1972).

Component IIIb was unusual in that it gave one band around 14,000 daltons when unreduced and two bands around 14,000 and 10,000 daltons when reduced. One possible explanation for this observation is that Component IIIb contained two 14,000 dalton proteins, one consisting of a single chain and the other containing two equal-sized subunits. The molecular weight discrepancy could result from either the reduced or unreduced protein running anomalously on the gels, a situation with some precedent (Fairbanks et al., 1971). It is also possible that one protein actually consisted of 10,000 and 4,000 dalton subunits and the latter diffused out of the gel during staining. This material was not studied further.

Component VII contained at least three proteins--one corresponding in molecular weight to the α-type toxins and whose apparent size was unaffected by disulfide bond reduction, and the other two corresponding to the β-type toxins in molecular weight. The latter were unusual in that only one of the β-size proteins produced two low molecular weight bands upon reduction, while the other retained its original apparent size. No attempt was made to further separate these components or to determine their nature.

Summary

BgV was fractionated into a total of nineteen components, two of which (II, IIIc) were small molecules. Of the thirteen major components examined by SDS-polyacrylamide gel electrophoresis, four gave a pattern similar to the one produced by β-BuTX, six gave a pattern similar to α-BuTX, and three gave patterns not fitting either category.

The α-BuTX fraction gave a clean, single band on gels, and appeared to be free of contamination by other proteins. This toxin fraction was used for the crystallization studies described in the next chapter.

References

Bosmann, H. B. (1972), J. Biol. Chem. 247, 130-145.

Clark, D. G., Macmurchie, D. D., Elliott, E., Wolcott, R. G., Landel, A. M., and Raftery, M. A. (1972), Biochemistry 11, 1663-1668.

Ellman, G. L., Courtney, K. D., Andres, V., Jr., and Featherstone, R. M. (1961), Biochem. Pharmacol. 7, 88-95.

Eterovic, V. A., Hebert, M. S., Hanley, M. R., and Bennett, E. L. (1975), Toxicon 13, 37-48.

Fairbanks, G., Steck, T. L., and Wallach, D. F. H. (1971), Biochemistry 10, 2606-2617.

Kelly, R. B. and Brown, F. R. III (1973), J. Neurobiol. 5, 135-150.

Lee, C.-Y., Chang, S. L., Kau, S. T., and Luh, S.-H. (1972), J. Chromatogr. 72, 71-82.

Wei, A.-L. and Lee, C.-Y. (1965), Toxicon 3, 1-4.

Figure Captions

Figure 2-1: The overall scheme used to fractionate BgV and to purify its components. The NH_4OAc gradients which were used are indicated in the figure. CM-52 is Whatman carboxymethyl cellulose ion-exchange resin. G-50 is Sephadex G-50 (fine). "Dis" means that the component was discarded.

Figure 2-2: Initial fractionation by BgV on a 2.5 x 50 cm CM-52 column. Chromatography was performed at 4°C using a linear, 2 ℓ, 0.1 M to 0.4 M, pH 5.8, NH_4OAc gradient followed by a linear, 2 ℓ, 0.4 M to 0.8 M, pH 5.8, NH_4OAc gradient. 7.5 ml fractions were collected. The first gradient ended at fraction 260.

Figure 2-3: Gel-filtration chromatography of venom Component III on a 2.5 x 56 cm G-50 column. 0.1 M acetic acid was used as the eluant. 3 ml fractions were collected. Chromatography was performed at room temperature.

Figure 2-4: Gel-filtration chromatography of α-BuTX (Component IV) on a 2.5 x 57 cm G-50 column. Conditions were the same as for Figure 2-3.

Figure 2-5: Gel-filtration chromatography of venom Components VII and VIII on a 2.5 x 55 cm G-50 column. Conditions were the same as for Figure 2-3.

Figure 2-6: Ion-exchange chromatography of venom Components X to XIII on a 2.5 x 50 cm CM-52 column. The column was run at 4°C. A linear, 0.6 M, pH 4.5 to pH 5.8 NH_4OAc gradient was used. 7.2 ml fractions were collected.

Figure 2-7: Ion-exchange chromatography of venom components C and D on a 2.5 x 50 cm CM-52 column. The column was run at 4°C. A linear, 2 ℓ, pH 7.0, 0.2 M to 0.5 M NH_4OAc gradient was used. 7.9 ml fractions were collected.

Figure 2-8: Ion-exchange chromatography of β-BuTX (Component XIV) on a 2.5 x 48 cm CM-52 column. The column was run at 4°C. A linear, 2 ℓ, pH 7.0, 0.2 M to 0.6 M NH_4OAc gradient was used. 7.8 ml fractions were collected.

Figure 2-9: SDS-polyacrylamide gels of the thirteen major BgV components. The reduced samples (R), were incubated for 30 minutes at 37°C with DTT before being applied to the gels, whereas the unreduced samples (U) were incubated without DTT. The bar at the bottom of each gel indicates the position of the pyronin Y band. The gels were aligned so the major bands were in approximately the correct relative positions. The very sharp bands and dark dots were defects in the background used for the photograph. The gels were overloaded to emphasize the minor bands.

Figure 2-10: Summary of the bands visible on the SDS-polyacrylamide gels. R_F values are relative to pyronin Y. Differences of 0.03 in R_F between different gels are not significant. The molecular weight scale is only approximate. Dashes indicate minor components. In some cases, the R_F values from two or more gels were averaged. Some of these bands are not visible or not resolved in Figure 2-9.

Figure 2-1

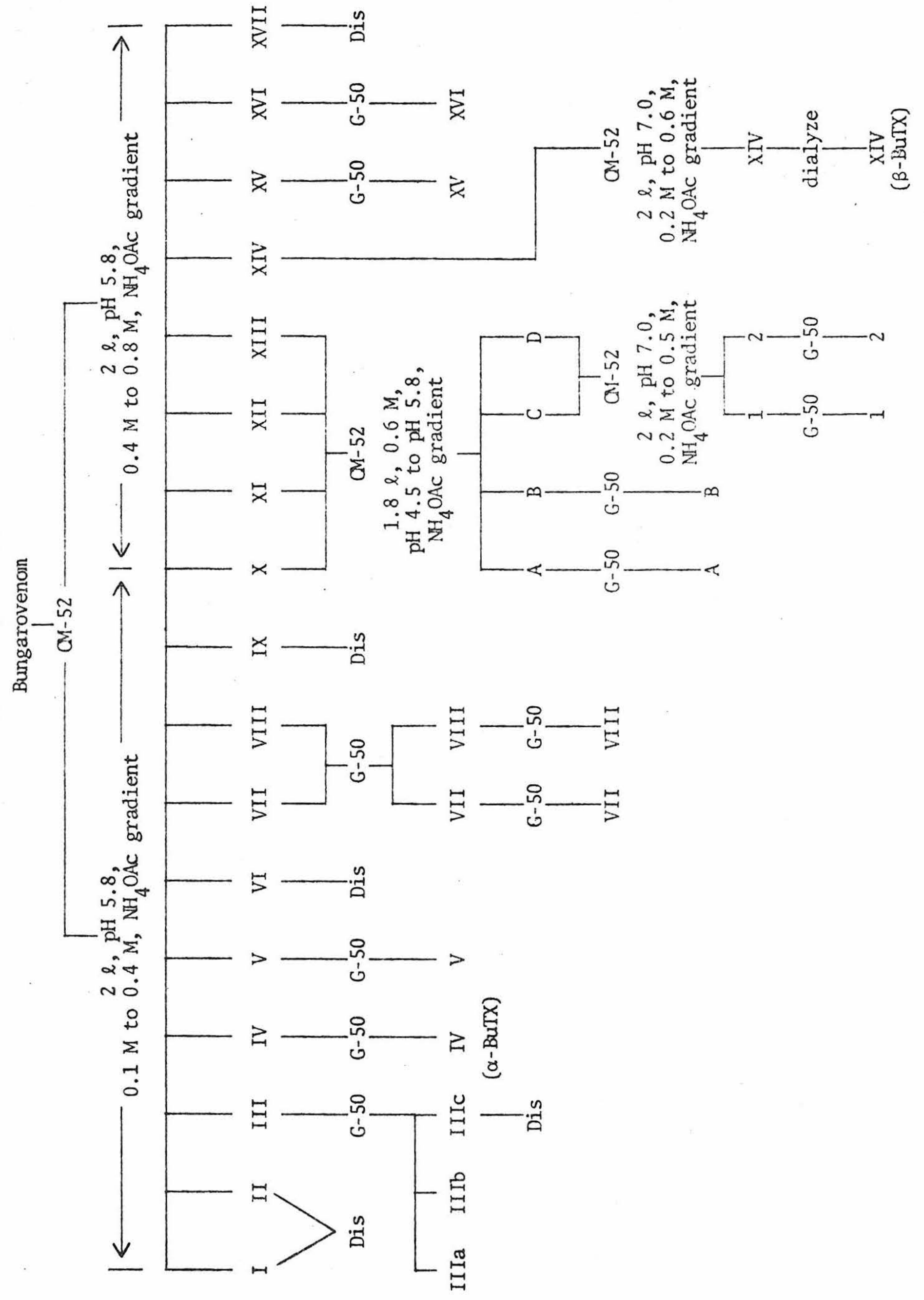

Bungarovenom
CM-52
2 ℓ, pH 5.8, 0.1 M to 0.4 M, NH_4OAc gradient
2 ℓ, pH 5.8, 0.4 M to 0.8 M, NH_4OAc gradient
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
XVII
Dis
G-50
IIIa
IIIb
IIIc
IV (α-BuTX)
CM-52
1.8 ℓ, 0.6 M, pH 4.5 to pH 5.8, NH_4OAc gradient
A
B
C
D
CM-52
2 ℓ, pH 7.0, 0.2 M to 0.5 M, NH_4OAc gradient
1
2
CM-52
2 ℓ, pH 7.0, 0.2 M to 0.6 M, NH_4OAc gradient
dialyze
XIV (β-BuTX)

Figure 2-2

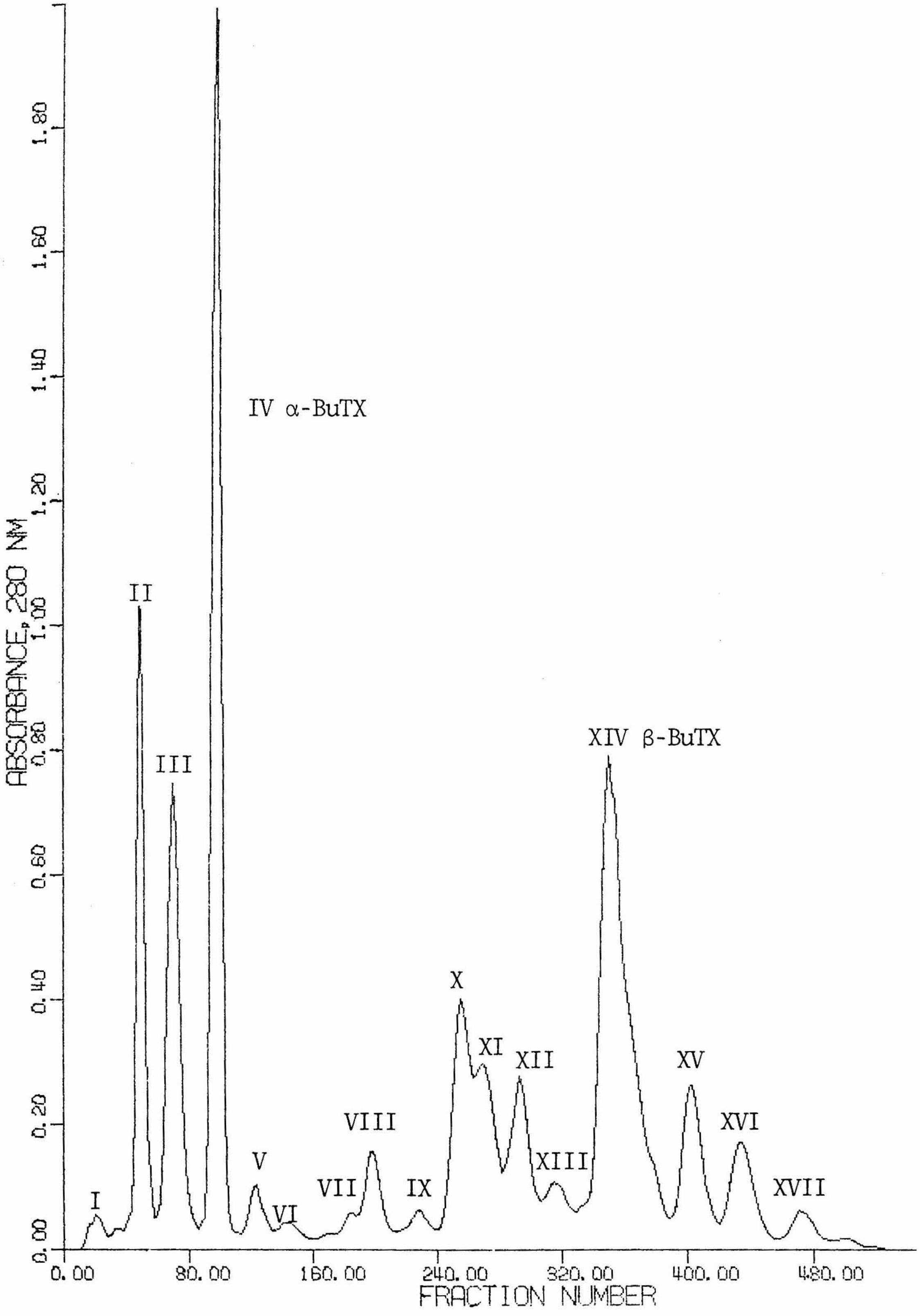

ABSORBANCE, 280 NM
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
FRACTION NUMBER
0.00
80.00
160.00
240.00
320.00
400.00
480.00
I
II
III
IV α-BuTX
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV β-BuTX
XV
XVI
XVII

Figure 2-3

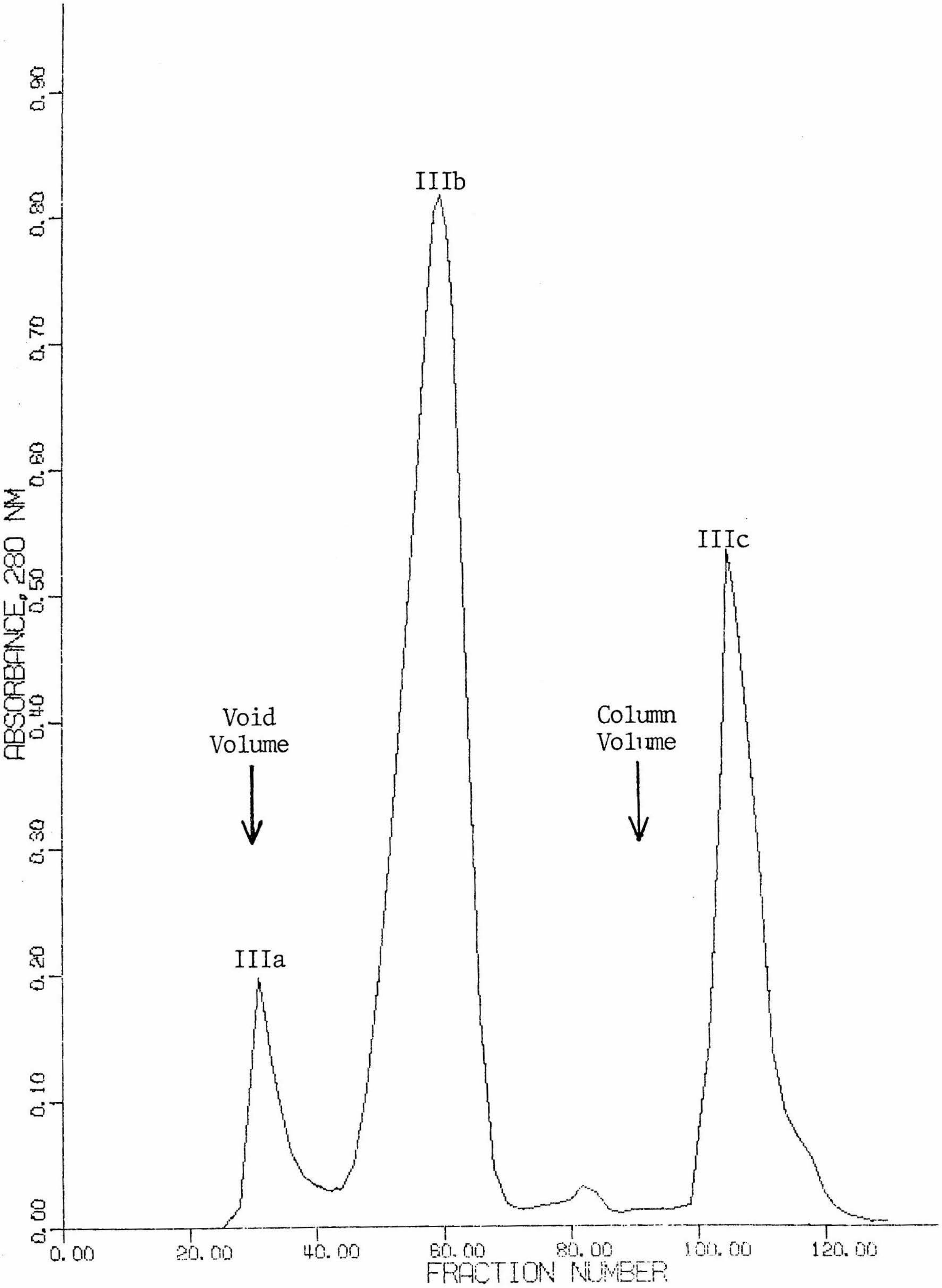

ABSORBANCE, 280 NM
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
FRACTION NUMBER
0.00
20.00
40.00
60.00
80.00
100.00
120.00
IIIa
IIIb
IIIc
Void Volume
Column Volume

Figure 2-4

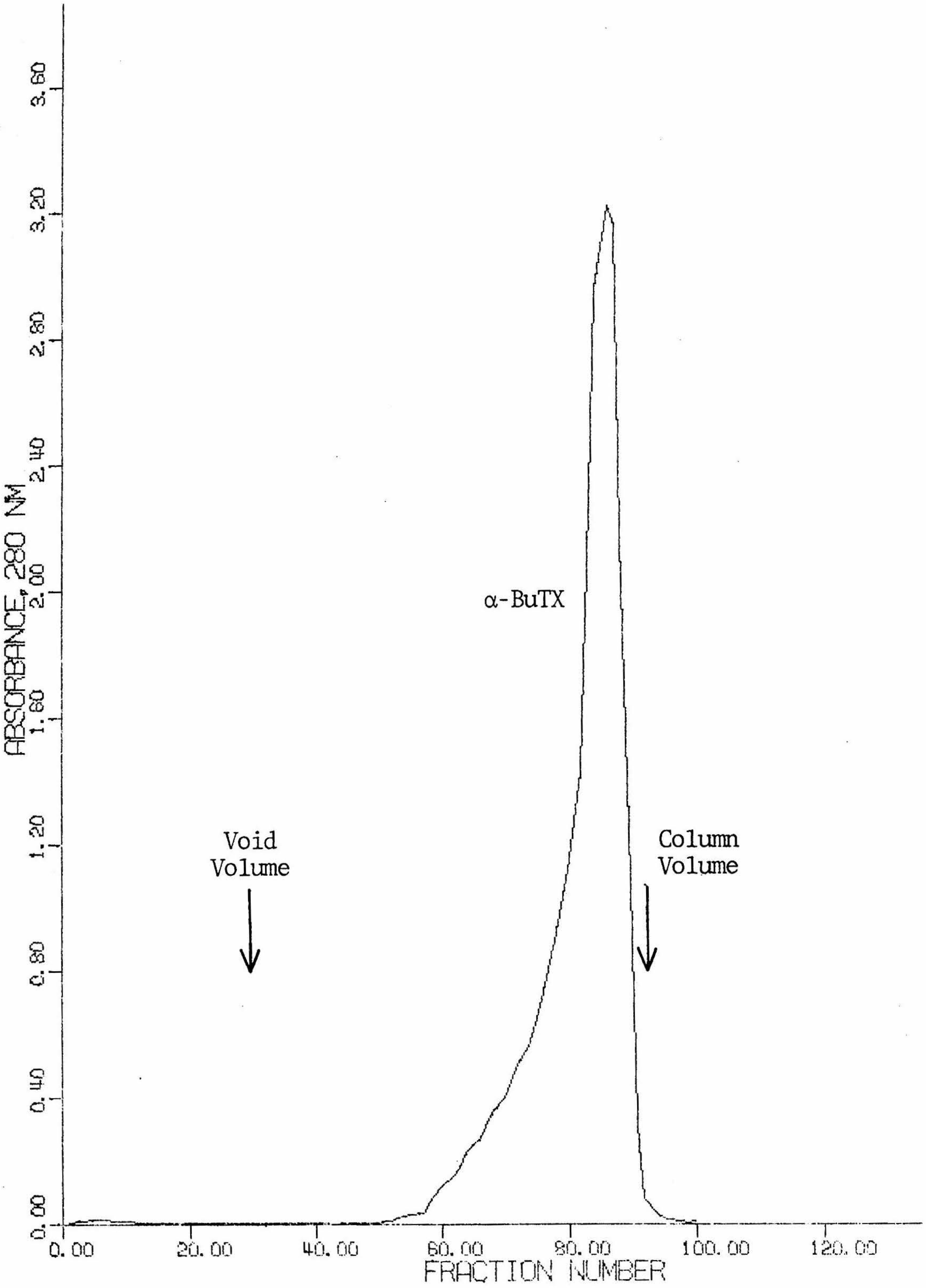

ABSORBANCE, 280 NM
0.00
0.40
0.80
1.20
1.60
2.00
2.40
2.80
3.20
3.60
FRACTION NUMBER
0.00
20.00
40.00
60.00
80.00
100.00
120.00
α-BuTX
Void Volume
Column Volume

Figure 2-5

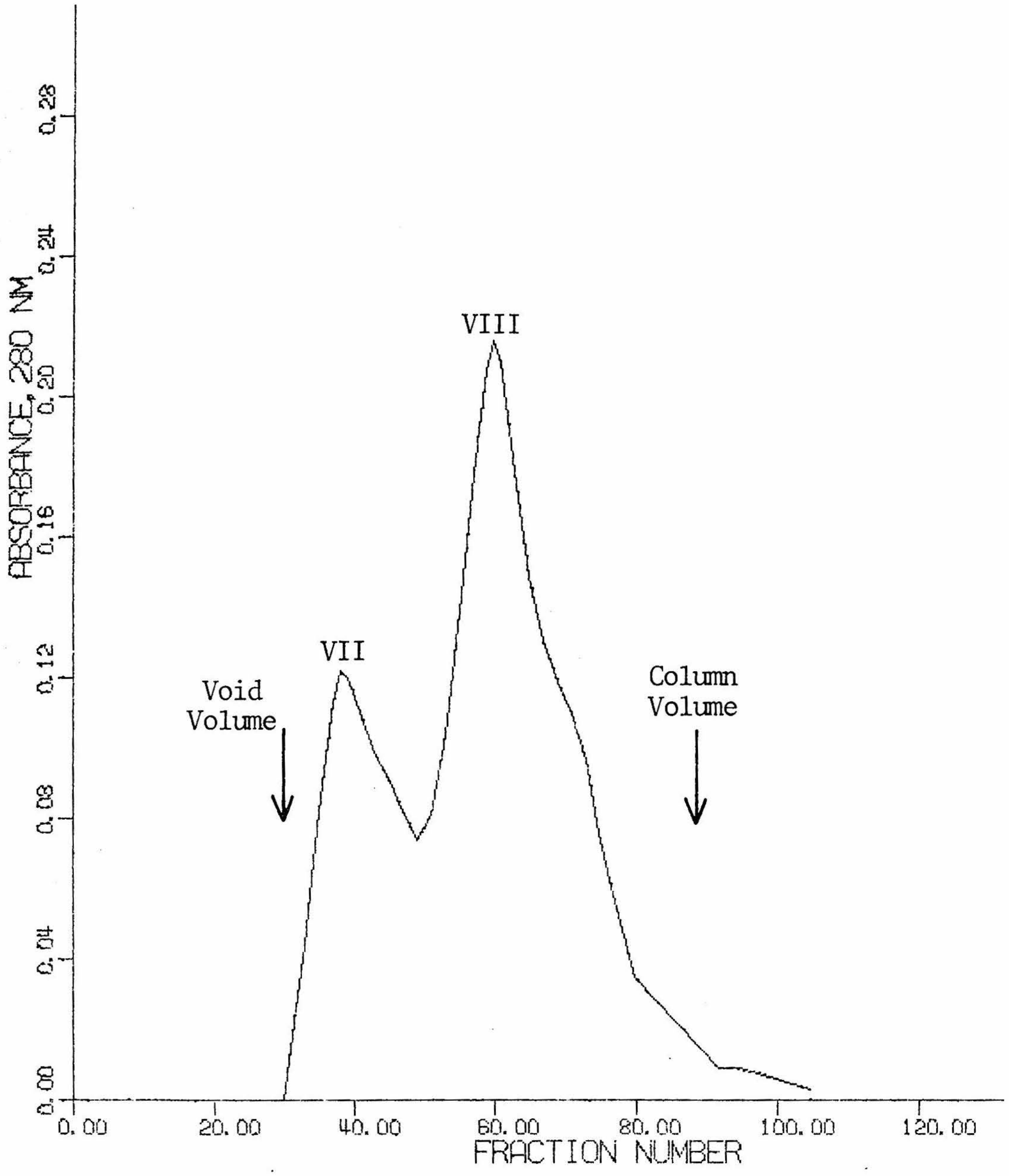

ABSORBANCE, 280 NM
0.00
0.04
0.08
0.12
0.16
0.20
0.24
0.28
FRACTION NUMBER
0.00
20.00
40.00
60.00
80.00
100.00
120.00
VII
VIII
Void Volume
Column Volume

Figure 2-6

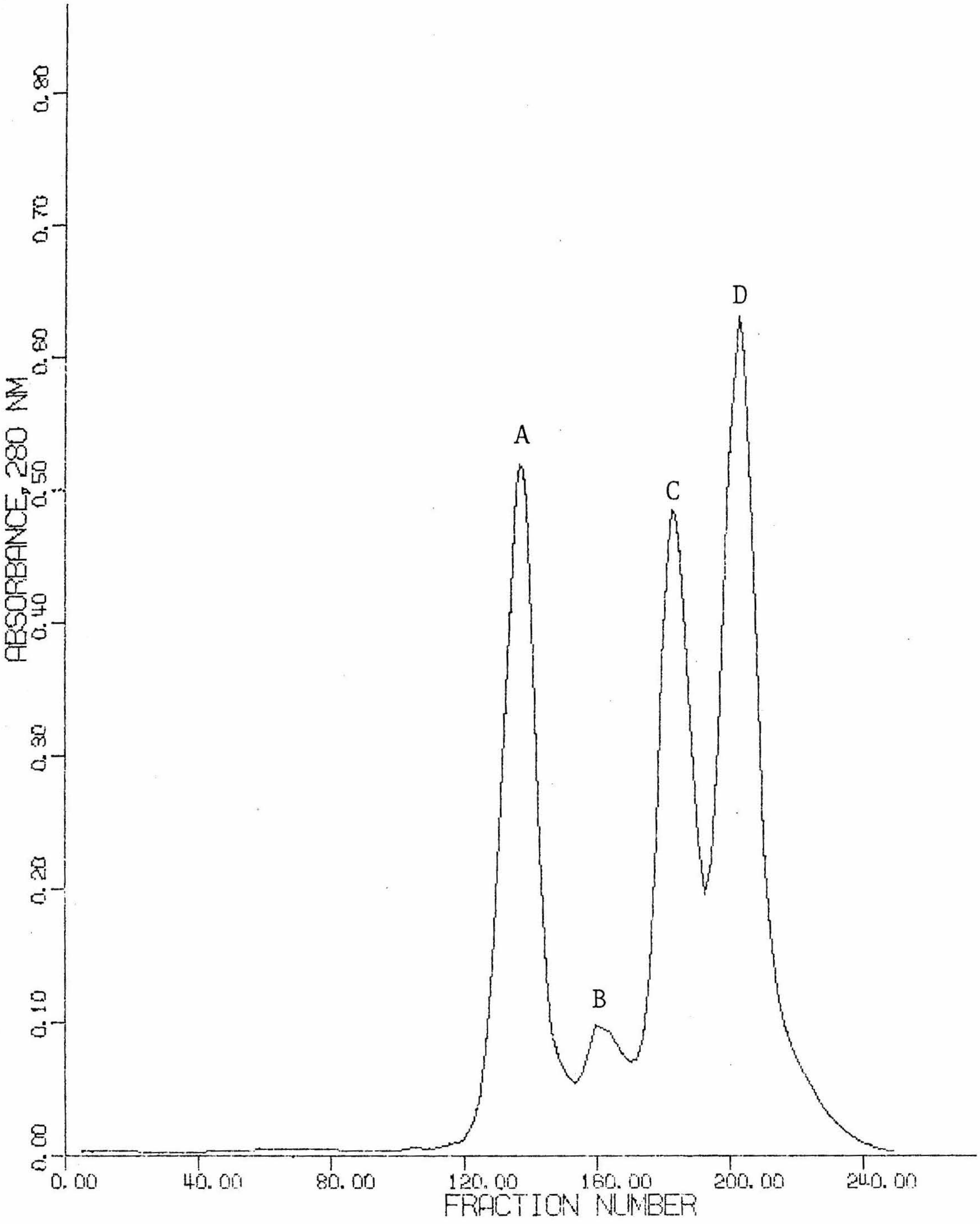

ABSORBANCE, 280 NM
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
FRACTION NUMBER
0.00
40.00
80.00
120.00
160.00
200.00
240.00
A
B
C
D

Figure 2-7

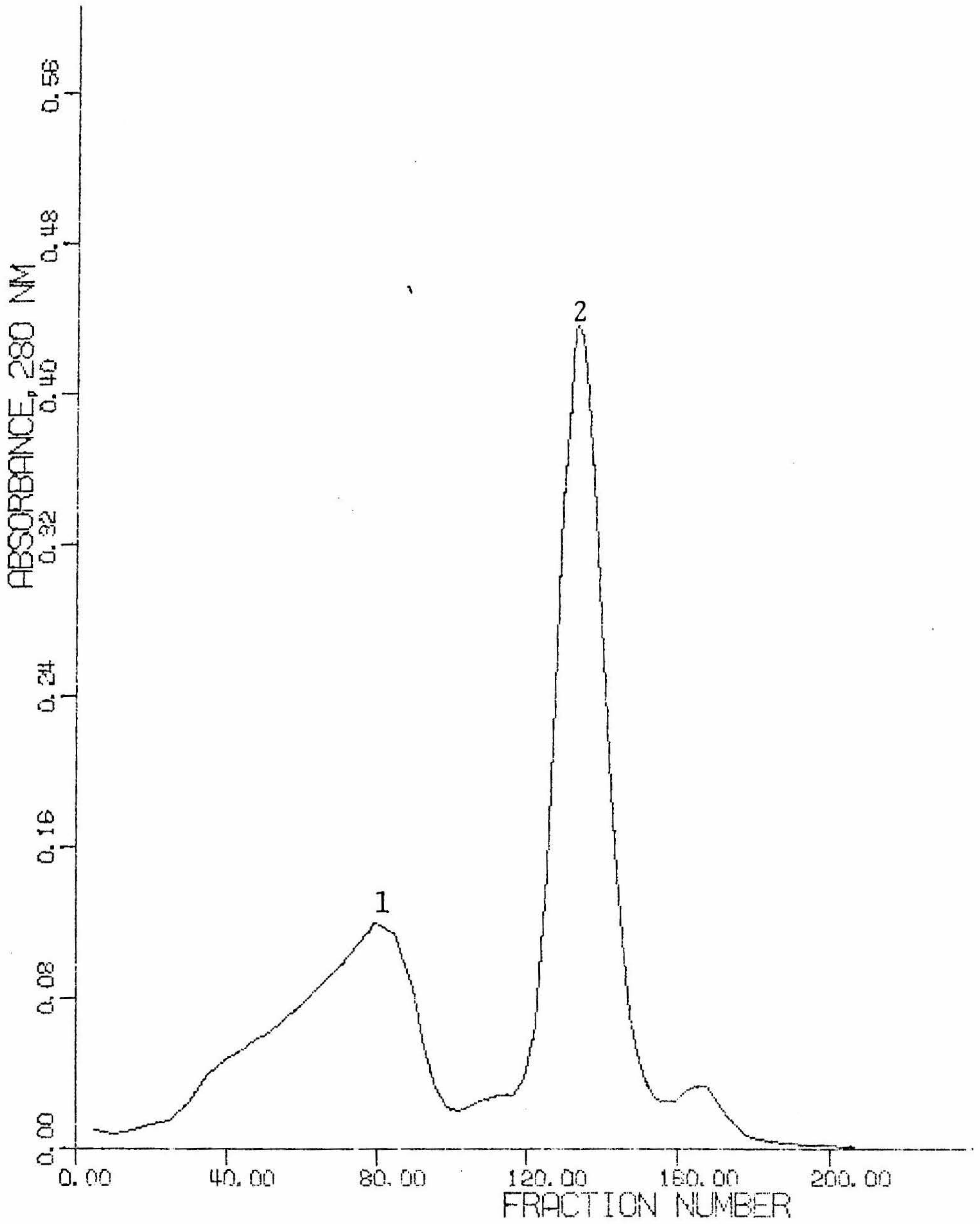

ABSORBANCE, 280 NM
0.00
0.08
0.16
0.24
0.32
0.40
0.48
0.56
0.00
40.00
80.00
120.00
160.00
200.00
FRACTION NUMBER
1
2

Figure 2-8

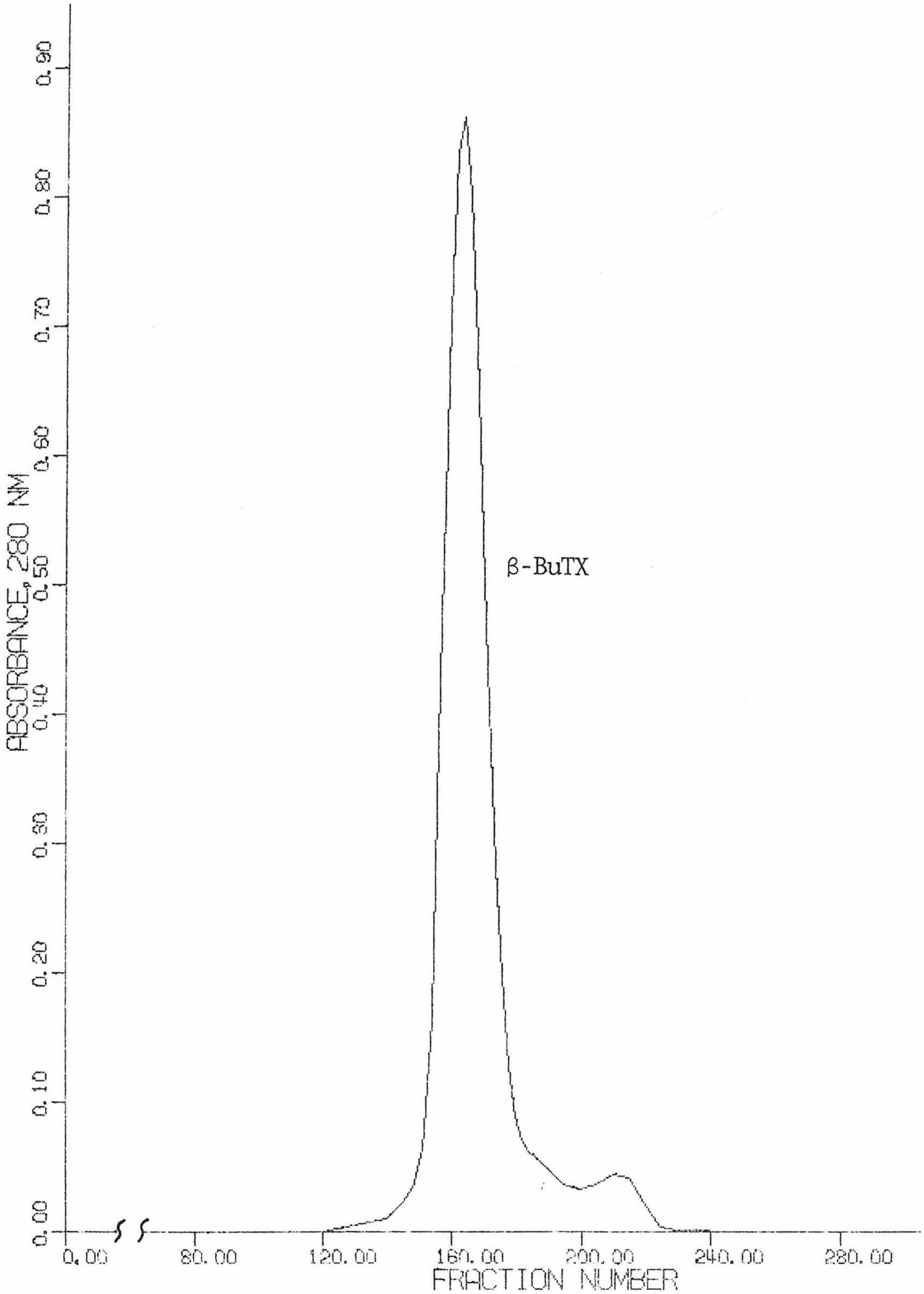

ABSORBANCE, 280 NM
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
β-BuTX
0.00
80.00
120.00
160.00
200.00
240.00
280.00
FRACTION NUMBER

Figure 2-9

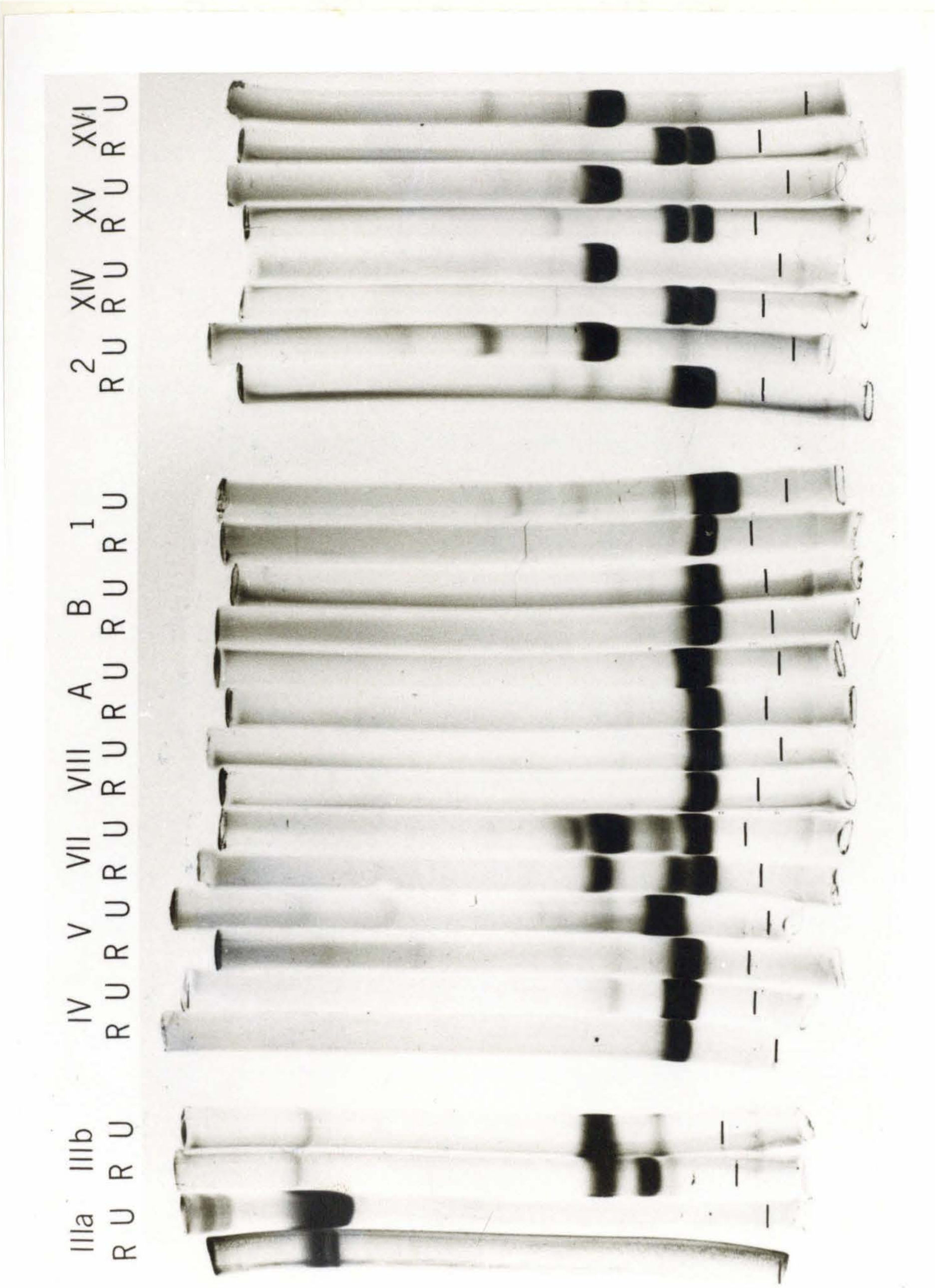
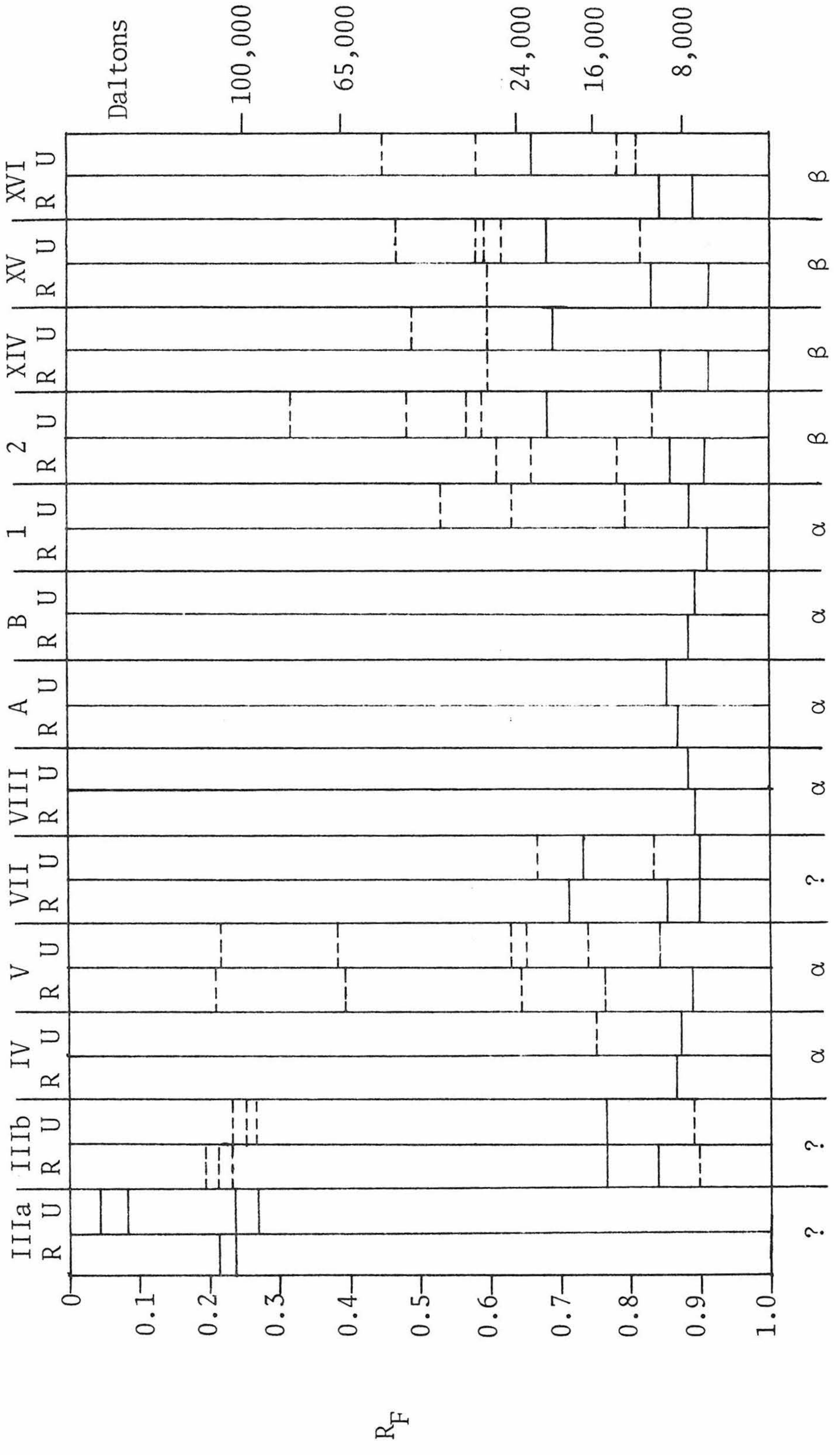

IIIa
R U
IIIb
R U
IV
R U
V
R U
VII
R U
VIII
R U
A
R U
B
R U
1
R U
2
R U
XIV
R U
XV
R U
XVI
R U

Figure 2-10


Daltons
100,000
65,000
24,000
16,000
8,000
IIIa R U
IIIb R U
IV R U
V R U
VII R U
VIII R U
A R U
B R U
1 R U
2 R U
XIV R U
XV R U
XVI R U
?
?
α
α
?
α
α
α
α
β
β
β
β
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
R_F

CHAPTER III

The Crystallization of α-Bungarotoxin and the Determination of its Crystalline Properties

"Ill habits gather by unseen degrees--
As brooks make rivers, rivers run into seas."

--John Dryden

Materials and Methods

Reagents

Ammonium sulfate was Schwarz/Mann 901946 "enzyme grade"; glycine and glutamic acid were obtained from Matheson Coleman and Bell; all other amino acids were from Sigma. Other chemicals were reagent grade.

Crystallizing techniques

Free diffusion: The free diffusion technique described by Salemme (1972) was used for the preliminary crystallizing experiments. In this method, a cell was made by cutting a disposable Pasteur pipet[1] approximately 5 cm from the tip and inserting a rubber serum cap in the large end (Figure 3-1). A hypodermic syringe equipped with a vernier control was inserted through the serum cap and 10 µl of α-BuTX solution (33 mg/ml) were drawn into the pipet, followed by 10 µl of the precipitant solution. An air bubble was then drawn into the tube and the bottom was sealed with sticky wax[2]. The syringe was removed and the tubes were stored upright, usually at 12°-14°C. The precipitate ring which formed at the protein-precipitant interface usually disappeared after several hours in the cold.

---

[1] The pipet was 5-3/4" x 7 mm O.D. VWR Scientific #14672-029.

[2] Kerr Manufacturing Company, Detroit, Michigan.

Equilibrium dialysis: The equilibrium dialysis cells described by Zeppezauer, Ekland, and Zeppezauer (1968) could not be used because the α-BuTX diffused through the dialysis membrane. Instead, the alternate method suggested by them, employing glass capillary tubes plugged by 20% polyacrylamide gel, was used. In place of the fragile glass capillaries used by Zeppezauer and coworkers, the free diffusion Pasteur pipet cells were employed (Figure 3-1).

The dialysis cells were prepared by dipping the pipet tip into a solution of 1.9 g acrylamide and 0.1 g BIS in 8 ml of buffer containing 0.4 saturated[3] sodium chloride and 0.2 saturated ammonium sulfate. Buffer containing a high salt concentration was used instead of water to prevent the plugs from shrinking and pulling free when the cells were placed in the precipitant solutions. The acrylamide plug was polymerized by heating at 80° to 100°C for about 2 minutes; then the tubes were filled with the desired precipitant solution and left to soak at least three days to remove any unreacted acrylamide. The acrylamide plug produced by this method was about 1 cm long.

The tubes were filled by first removing the buffer with a syringe and then placing 10 μl of α-BuTX solution (33 mg/ml) on top of the gel plug. Each tube was capped with a serum cap through which a hypodermic needle had been inserted to prevent the buildup of air pressure inside the cell. After the serum cap was inserted, the needle was removed, and the tube was placed upright in the desired

---

[3] "Saturation" refers to saturation at toom temperature (∿23°C).

buffer. The crystallizing trials were generally performed at 12°-14°C.

Batch crystallization: Once crystallizing conditions were established using the above techniques, large numbers of crystals were grown by batch crystallization. For this technique, a buffer of the proper precipitant concentration and pH was first made, and then lyophilized α-BuTX was added to produce a final protein concentration around 40 mg/ml. This solution was filtered through a 0.45 μm Millipore filter to remove solid contaminants, and then placed in a 5 x 50 mm test tube. The tube was sealed with Parafilm® and stored upright at 10°C.

Crystal density

Two methods were used to measure crystal density. In the first, several crystals were removed from their mother liquor, quickly dried by touching them with a Kimwipe, and then placed in a mixture of chlorobenzene and bromobenzene of approximately the correct density. The composition of this mixture was adjusted until the crystals were neutrally buoyant. The density of the mixture was then measured pycnometrically.

® American Can Company.

The second method involved placing crystals in cesium chloride solutions of known densities until a solution was found in which the crystals were neutrally buoyant. To guard against crystal damage and consequent density changes, the cesium chloride solutions were maintained at the same pH as the mother liquor, and the crystals were monitored under a microscope for visible changes. In addition, the flotation behavior of the crystals was determined within seconds after they were placed in the cesium chloride solution, thereby circumventing the problem of cesium chloride diffusing into the crystals and changing their density before an observation was made.

X-ray measurements

Initial measurements of the x-ray diffraction properties of the crystals were made using a precession camera. Precession data and diffractometer data were later collected by the methods described in detail in the next chapter.

Results and Discussion

Crystallization

Free diffusion: Initial crystallizing experiments were performed using the free diffusion method. The precipitant was ammonium sulfate buffered to pH 7.0 with sodium phosphate; the protein solution was unbuffered. These experiments produced small,irregular crystals or patches of polycrystalline material, usually within seven days after being set up. Varying the precipitant pH from 6.5 to 8.0 in 1/2 pH unit steps did not improve these crystals, nor did changing the buffer; and so other precipitants were tried.

Organic precipitants (ethanol, *p*-dioxane, 2-methyl-2,4-pentanediol) produced only amorphous precipitate, whereas several halide salts gave tiny needles. The best of these salts was sodium chloride, which produced needles 0.2 to 0.3 mm long, but only 0.02 mm or less in diameter. The pH of the sodium chloride solutions was varied in approximately 1/2 pH unit steps between pH 4.0 and 9.5 in an attempt to increase the size of these crystals. Below pH 7.5, fine needles formed; and at pH 4.0 clusters of needles 0.5 to 1.0 mm long (although still less than 0.05 mm wide) were produced.

Above pH 7.5, only amorphous precipitate formed except in the case of one pH 9.0 solution. In this case, ammonium sulfate solution had been added to the tube about a month after it had been set up, and after several months the tube began to leak. Long, broad needles formed (about 0.5 x 0.15 x 0.05 mm), and they were examined by x-ray diffraction (Spencer, 1973). The crystals produced from this cell are referred to later as "high pH" crystals.

Because the diffusion cell leaked, the internal pH was unknown, thereby making it difficult to duplicate the crystallizing conditions. All attempts to duplicate the crystals at pH 9.0 were unsuccessful. Around pH 7.0, similar crystals were produced, but they never grew as large as the original ones. These efforts were finally abandoned. The mixture of sodium chloride and ammonium sulfate was retained in later crystallizing experiments, because the ammonium sulfate seemed to make the thin needles produced by the sodium chloride grow larger in one direction.

Equilibrium dialysis: The limitations imposed by the free diffusion method upon the degree to which crystallizing conditions could be varied proved to be too great an inconvenience, and so most crystallizing experiments were performed in the acrylamide-plugged equilibrium-dialysis cells. A mixture of ammonium sulfate and sodium chloride was used as the precipitant, and cells were set up at approximately every 1/2 pH unit from pH 2.5 to 9.0. The cells equilibrated with the outside solution after about two weeks, and usually produced crystals or precipitate within this time period.

The crystallization of α-BuTX occurred in several distinct zones. Above pH 7.5, the protein tended to form clear, oil-like droplets and occasionally very small crystals; seed crystals did not induce these solutions to crystallize. Between pH 7.5 and 6.0, needles formed. The solubility of α-BuTX in the precipitant solution decreased markedly below pH 6.0, although needles continued to form down to pH 3.5. Below pH 3.5, the crystals changed shape and began to form as short, two-dimensional barrels or long, thin laths (Figure 3-2a). Crystals were produced most readily around pH 4.0 and pH 7.0.

The morphological change at pH 3.5 occurred at the pH where the buffer changed from Tris-acetate to diglycine-HCl. To see whether the change in buffer was responsible for this change in morphology, diglycine was added to some of the pH 4.0 crystallizing solutions. Although needles were still produced from these solutions, they were of significantly larger diameter than those produced previously. Various monomeric amino acids were also tried, but only glycine produced crystals of a size comparable to those produced by adding diglycine.

Several attempts were made to grow crystals at room temperature as well as in the cold. The pH 4.0 solutions which produced large needles at 12°C produced only very thin plates or thick clumps of tiny needles at room temperature. These crystals tended to fuse into one large mass, and so further experiments at room temperature were abandoned.

Batch crystallization: Once the crystallizing conditions were determined by equilibrium dialysis, batch crystallization experiments were set up to grow the large number of crystals needed for x-ray structure determination. These crystals were produced as follows: 8.7 mg of α-BuTX were dissolved in 200 μl of a solution containing 0.25 saturated sodium chloride, 0.1 saturated ammonium sulfate, 5 mM EDTA, 0.05 M Tris, 0.05 M glycine, and sufficient glacial acetic acid to adjust the pH to 4.25. This solution was filtered and put into a small test tube which was sealed and immediately placed in a refrigerator at 10-12°C. If the solution was not cooled before precipitate started to form, only long, thin crystals were produced.

Crystals appeared in two to four weeks and continued to grow for several months (Figure 3-2b). New crystals also continued to form, usually producing thick clusters from which the individual crystals had to be broken. Eventually the mother liquor was removed and replaced by a solution of 0.3 saturated sodium chloride, 0.15 saturated ammonium sulfate, 0.05 M glycine, 1.0 M acetic acid and sufficient saturated sodium hydroxide to raise the pH to 4.0. The crystals were stable indefinitely in this solution at 8°C.

Crystal properties

Table 3-1 contains a summary of some properties of several of the α-BuTX crystal types. These properties are discussed below.

Density: The two methods for determining crystal density gave slightly different results. The bromobenzene-chlorobenzene flotation gave a value of 1.25 ± 0.02 g/cm^3 for the pH 4.0 crystals, while the cesium chloride flotation gave a value of 1.29 ± 0.01 g/cm^3. The latter value was probably more nearly correct because, unlike the bromobenzene-chlorobenzene flotation result, this measurement was not appreciably affected by mother liquor clinging to the crystals.

The density of the artificial mother liquor was measured pycnometrically as 1.07 g/cm^3.

TABLE 3-1

α-Bungarotoxin crystal properties

Crystal form		pH 2.6	High pH[1]	pH 4.0[2]	pH 4.0 (small cell)[3]
Space group		$P2_12_12_1$[4]	$P2_12_12_1$	--[3]	$P2_12_12_1$
Cell dimensions, Å:	a =	69.6	69.4	69.86±0.05[5]	67.8
	b =	75.9	77.9	76.71±0.05	78.4
	c =	44.7	22.4	44.75±0.01	22.4
Density, g/cm^3		--	--	1.29	--
Molecules/asymmetric unit		4	2	4	2
V_M, Å^3/dalton		1.84	1.89	1.88	1.86
% Protein		65%	63%	64%	64%

---

[1] The pH was probably around 7.0 (see text).

[2] This crystal form was used for the structural study reported in this thesis.

[3] Refer to the X-ray Diffraction section of this chapter for an explanation of the small-cell crystal form and the large-cell space group ambiguity. These problems are discussed in detail in Chapter V.

[4] The space group and cell dimensions may be in error due to the weak, low-resolution diffraction produced by these small crystals.

[5] Six crystals were used to calculate these cell dimensions and standard deviations. Cell dimensions were measured on a Syntex $P\bar{1}$ diffractometer.

Stability: The pH 4.0 crystals were unstable with respect to increases in temperature or pH. At room temperature or above pH 4.5, the crystals usually cracked; and crystals stored at room temperature had a significantly different diffraction pattern from those stored at 10°-12°C (see the next section). Reducing the pH to 3.5 had no effect, nor did lowering the storage temperature of the crystals to 4°C. For this reason, the crystals were stored at 5°-8°C at a pH slightly below that at which they were grown. A high buffer concentration (1.0 M) was used to guard against pH changes.

The sensitivity of the crystals to a pH above 4.5 pointed to one or more carboxylic acids as the most probable groups responsible for altering the crystal packing. Since raising either the pH or the temperature would favor ionization of carboxyl groups, it appeared that the stability of the crystal packing depended on these groups remaining neutral. This metastable packing arrangement may have been the factor which allowed these crystals to grow larger than any of the other crystal forms produced.

X-ray diffraction: The x-ray diffraction properties of several of the α-BuTX crystal forms are summarized in Table 3-1. The "high pH" crystals were those produced by the leaking free diffusion cell, as discussed previously. Neither these crystals nor the pH 2.6 crystals were examined further.

Figure 3-3 shows the relationship between the crystal axes and crystal morphology for the large-cell pH 4 crystals used in the structure determination. 15° precession photographs of the $\underline{h0\ell}$ and $\underline{0k\ell}$ levels are shown in Figure 3-4. Measurable diffraction from these crystals extended to a Bragg spacing of about 2 Å.

The diffraction pattern had $\underline{mmm}$ symmetry, indicating an orthorhombic space group. The $\underline{h}00$ and $0\underline{k}0$ reflections were present only when $\underline{h}$ or $\underline{k}$ was even. The $00\underline{\ell}$ reflections were more difficult to assess, because the $\underline{\ell}$ = odd reflections were weak in any case. With some crystals, the 0,0,5 reflection appeared to be weakly present, and so the space group was tentatively assigned as $P2_12_12$. (See Chapter V for further discussion of the space group problem.)

The most striking feature of the diffraction pattern was that the $\underline{\ell}$ = odd reflections were generally weak throughout reciprocal space, thus indicating that the unit cell contained two similar (but not identical) layers of molecules stacked along the $\underline{c}$ axis. The average intensity of the $\underline{\ell}$ = odd reflections increased at high diffraction angles, with the $\underline{\ell}$ = odd and $\underline{\ell}$ = even reflections reaching the same average intensity at a Bragg spacing of about 3 Å (Figure 3-5). This result indicated that at 3 Å resolution the differences between the top and bottom halves of the cell (in $\underline{z}$) were fully resolved. Thus, the doubled cell was probably caused by a relative shift of about 3 Å between the molecules in these two halves of the cell.

The diffraction pattern also showed evidence of crystal packing disorder. Diffuse scattering occurred on the $\underline{0k\ell}$ projection as streaks extending from the discrete reflections in a direction parallel to the b axis and confined to the ℓ = even rows. The diffuse scattering became more prominent at higher angles on the projection, but was not present at all on the h0ℓ projection. The nature of the disorder which produced this diffuse scattering has not been determined.

An unusual property of these crystals was that they slowly decayed into another crystal form when kept at room temperature. This process was hastened by exposure to x-ray radiation. The change involved a conversion from the "large-cell" form, with a c axis length of 44.7 Å, to a "small-cell" form with c = 22.4 Å; and was marked by the disappearance of the ℓ = odd reflections from the large-cell diffraction pattern (Figure 3-6). As can be seen in Table 3-1, the other two cell dimensions changed substantially as well.

In addition to the above changes, the strong 0,0,10 reflection totally disappeared upon conversion of the crystals to the small-cell form (in which it became the 0,0,5 reflection). This change strongly indicated that the c axis in the small-cell crystals was a true 2_1 screw axis. The systematic absences for this crystal form were consistent with space group $P2_12_12_1$.

A 3.5 Å set of three-dimensional data was collected on one small-cell crystal, and was then scaled to the ℓ = even reflections in the large-cell data set. The small-cell data fell in intensity (as a function of $(2 \sin\theta)^2$) much more rapidly than the large-cell data and had to be multiplied by $\exp\{1.57\ (2 \sin\theta)^2\}$ to compensate for this difference. A comparison of the structure-factor magnitudes from the two crystal forms, although of dubious significance because of the large differences in lattice parameters between the crystals, showed a difference[4] of 20% between the two forms at 10 Å resolution, with this difference rising to 36% at 3.5 Å resolution. This increase in the difference as a function of resolution may have been due to the large relative temperature factor between the two data sets, because these values were computed before application of the exponential scaling term. Final scaling parameters were not available at the time of this writing.

Molecules per cell: The approximate number of molecules in the crystallographic unit cell was calculated from the measured density and unit cell volume, using the relation (Eisenberg, 1970)

$$n = NV\rho X_p/M$$

where n = number of molecules per cell, N = Avogadro's number, V = cell volume (in cm^3), ρ = crystal density, X_p = weight fraction of protein in the crystal, and M = molecular weight of the protein.

---

[4] $\langle |F_s - F_{\ell,\mathrm{even}}|/F_{\ell,\mathrm{even}} \rangle$, where s refers to the small cell, ℓ refers to the large cell, and "even" refers to the ℓ = even reflections. F is used to represent the magnitude of the structure factors.

A value for X_p was calculated from the expression (Eisenberg, 1970):

$$X_p = (\bar{v}_\ell - 1/\rho)/(\bar{v}_\ell - \bar{v}_p)$$

where $\bar{v}$ is the partial specific volume, ℓ and p refer to the mother liquor and protein respectively, and ρ is the crystal density. $\bar{v}_p$ was calculated from the amino acid composition (Schachman, 1961), and $\bar{v}_\ell$ was approximated by $1/\rho_\ell$. Thus, for the large-cell pH 4 crystals,

$$X_p = (1/1.07 - 1/1.29)/(1/1.07 - 0.72) = 0.74.$$

With the above value for X_p, a measured cell volume for the pH 4 crystals of 2.40×10^5 Å^3, and a molecular weight of 8,000 for α-BuTX, the value 17.2 molecules per cell was computed for n. Since the crystals were orthorhombic, n had to be a multiple of 4. Also, as described above, the crystals were able to decay into another orthorhombic cell about half the size of the original one. This required that the number of molecules in the original cell be divisible by eight as well as by four, thus indicating that sixteen was the correct value for the number of molecules in the large cell.

Using the value n = 16, V_M, the volume per dalton of protein, was calculated as 1.88 Å^3/dalton. This value was below the mean value for protein crystals (Matthews, 1968), but well within the "normal" range. From this number, the protein content of the crystals was calculated as 64% by volume.

Based on a solvent content of 36%, the presence of sixteen molecules in the cell and the measured density of 1.07 g/cm^3 for the mother liquor, the crystal density was calculated as 1.27 g/cm^3. This calculated value agrees reasonably well with the measured value of 1.29 g/cm^3.

Further Discussion

The occurrence of a protein crystal containing a supercell was, at the very least, quite surprising. Since the crystals contained four molecules in the asymmetric unit, and since two of these were apparently related to the other two by a translation of approximately c/2, it seemed probable that a dimer pair was the basic crystallizing element, and that the supercell resulted from the uneven stacking of these dimer pairs. This belief was further strengthened by the room-temperature behavior of the crystals, which was characterized by the conversion of the crystals to a small-cell form with two molecules in the asymmetric unit. Presumably this change resulted from slight relative movements between the dimer pairs.

The fact that this conversion from one crystal form to the other occurred without destroying the crystal lattice (although the crystals often cracked in the process) indicated that the movement involved was not very large. This required that any model for the molecular packing arrangement be able to explain this conversion by some simple mechanism. This restriction later proved to be useful for choosing one heavy-atom derivative solution over several other possibilities (see Chapter V).

References

Eisenberg, D., in "The Enzymes" (P. Boyer, ed.), 3rd edition, Vol. 1, Academic Press, New York, pp. 13-14, 1970.

Matthews, B.W. (1968), J. Mol. Biol. 33, 491-497.

Salemme, F.R. (1972), Arch. Biochem. Biophys. 151, 533-539.

Schachman, H.K. (1961), Methods in Enzymology 4, 70-71.

Spencer, S.A. (1973), Candidacy Report, California Institute of Technology, Department of Chemistry.

Zeppezauer, M., Eklund, H., and Zeppezauer, E.S. (1968), Arch. Biochem. Biophys. 126, 564-573.

Figure Captions

Figure 3-1: Crystallizing cells. The cell on the left is a free-diffusion cell, and the one on the right is an equilibrium dialysis cell. Normally the tip of the equilibrium dialysis cell would be submerged in some precipitant solution. Each cell contains about 10 μl of protein solution (colored for visibility) and the free-diffusion cell also contains about 10 μl of precipitant solution below the protein solution. Each scale division is 1 mm.

Figure 3-2: α-Bungarotoxin crystals. a) Thin laths produced at pH 2.6; b) Fat needles produced at pH 4.0. These crystals were used for structure determination. The crystals in both photographs are magnified about 60x. The bar in the upper right corner of each picture corresponds to 0.1 mm.

Figure 3-3: An idealized drawing of the pH 4 α-bungarotoxin crystals showing the relationship between the direct-space crystal axes and the crystal morphology.

Figure 3-4: Diffraction patterns[5] from the pH 4 α-bungarotoxin crystals. a) $\underline{h0\ell}$ precession photograph. b) $\underline{0k\ell}$ precession photograph. Both pictures were taken at a precession angle of 15° (Bragg spacing = 3 Å) using nickel-filtered, CuK_α radiation from an Elliott GX-6 rotating anode x-ray generator.

---

[5] The density at the centers of these photographs was reduced during printing (Courtesy of the Artful Dodger) to make the inner reflections more readily visible.

Figure 3-5: A graph of the change in the ratio $<F_{odd}>/<F_{even}>$ as a function of resolution. Odd and even refer to the $\underline{\ell}$ = odd and $\underline{\ell}$ = even reflections respectively in the large-cell data set. The point at $(2 \sin\theta)^2 = 0.292$ was omitted because an unusually small number of $\underline{\ell}$ = odd reflections in this zone (65 versus approximately 180 for the surrounding zones) resulted in a ratio of 1.98.

Figure 3-6: An $0\underline{k}\underline{\ell}$ diffraction pattern[5] for the small-cell, pH 4, α-bungarotoxin crystals. This photograph was taken under the same conditions as those in Figure 3-4, except that the precession angle was 17° (Bragg spacing = 2.64 Å) in this case. Conversion to the small cell was not complete when this photograph was taken, and so some of the reflections on the $\underline{\ell}$ = 9 rows are still visible.

Figure 3-1

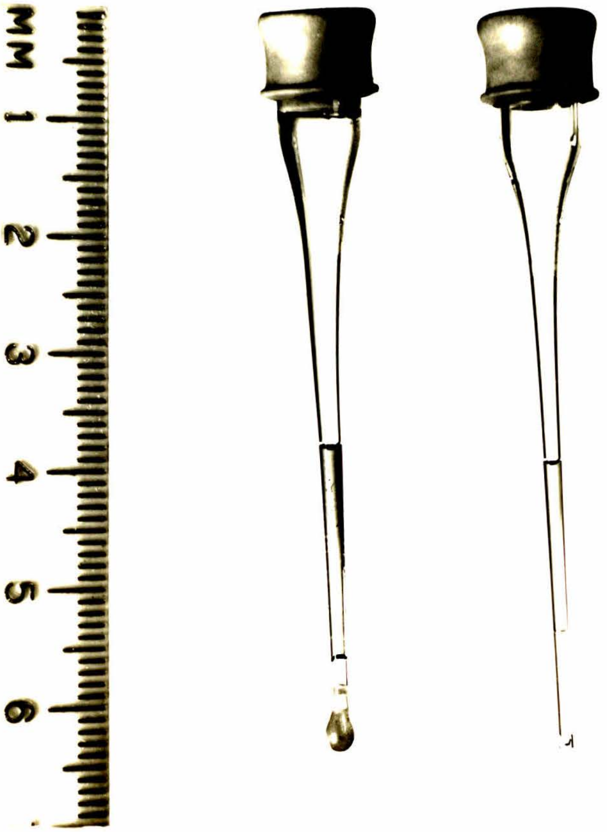


Figure 3-2a

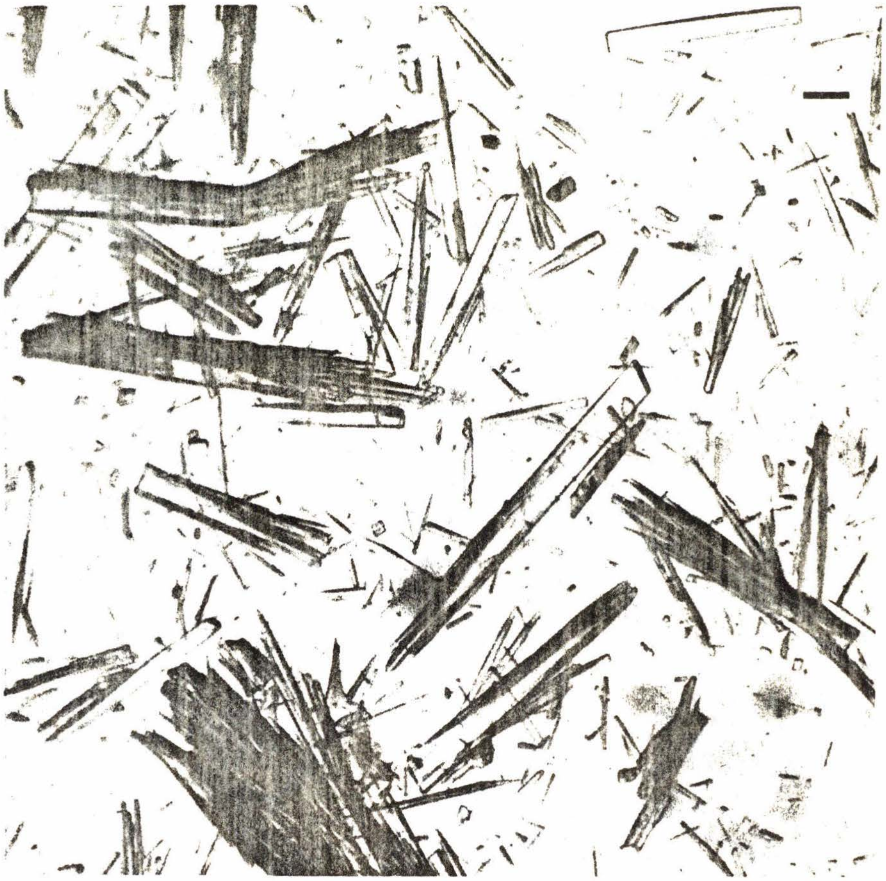


Figure 3-2b

Figure 3-3

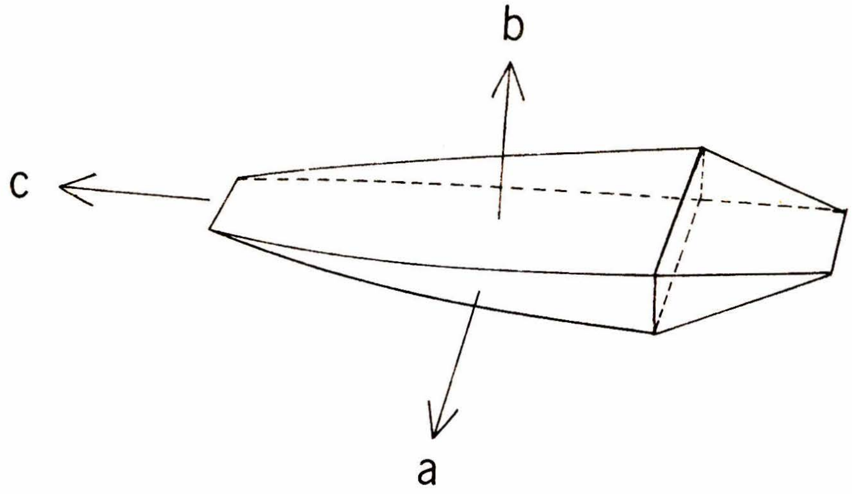


Figure 3-4a

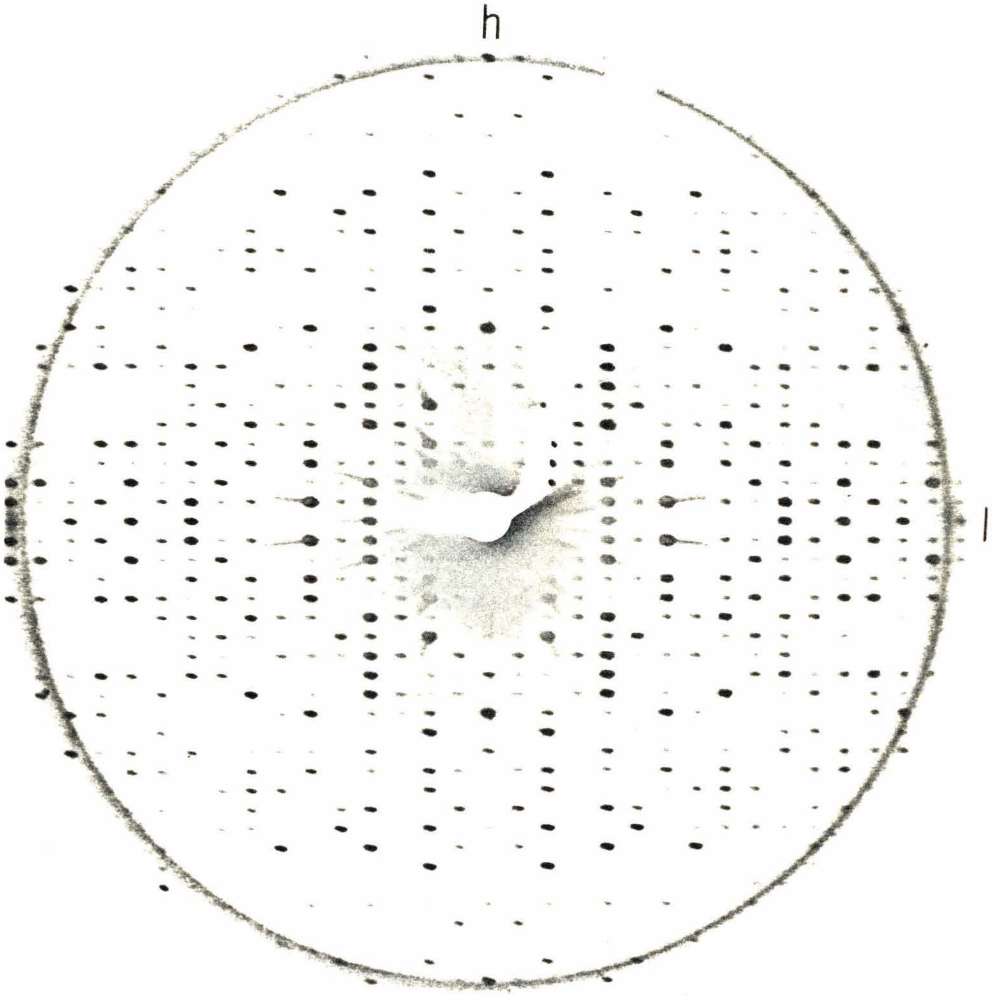


Figure 3-4b

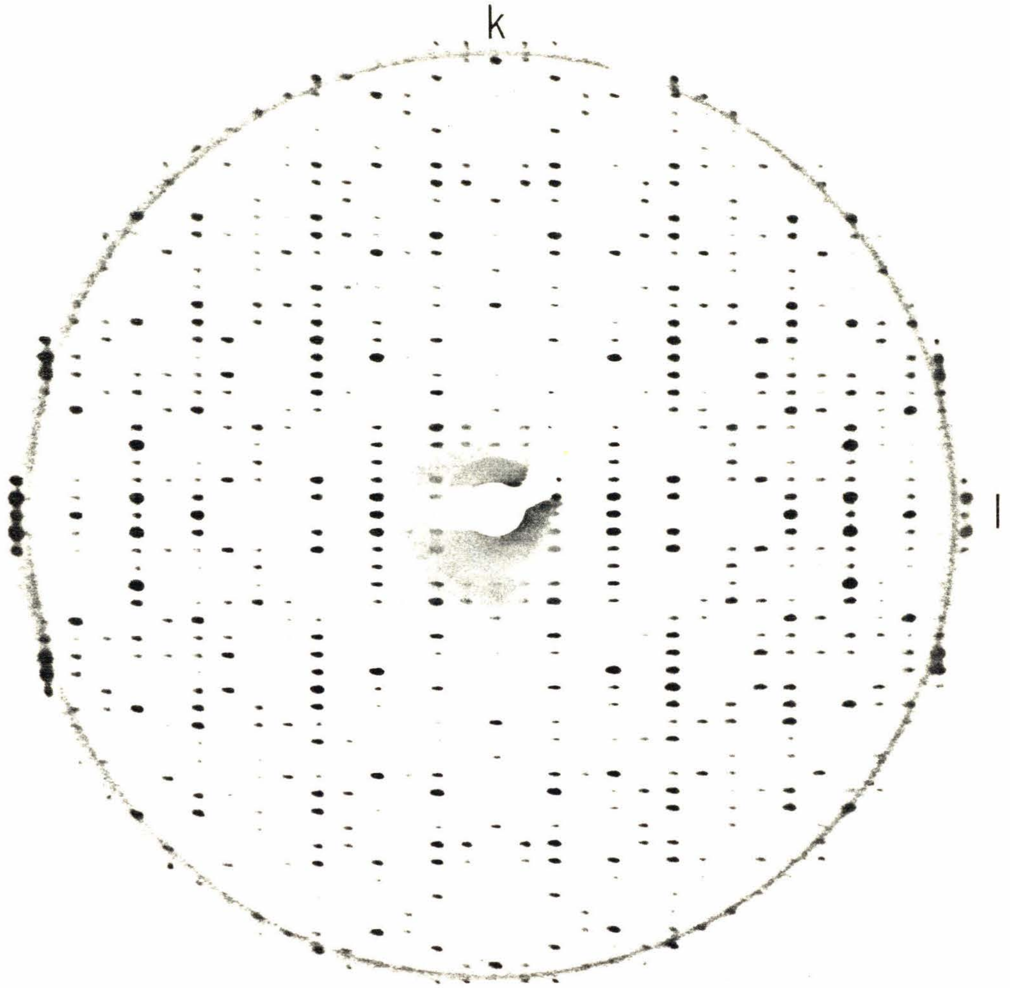


Figure 3-5

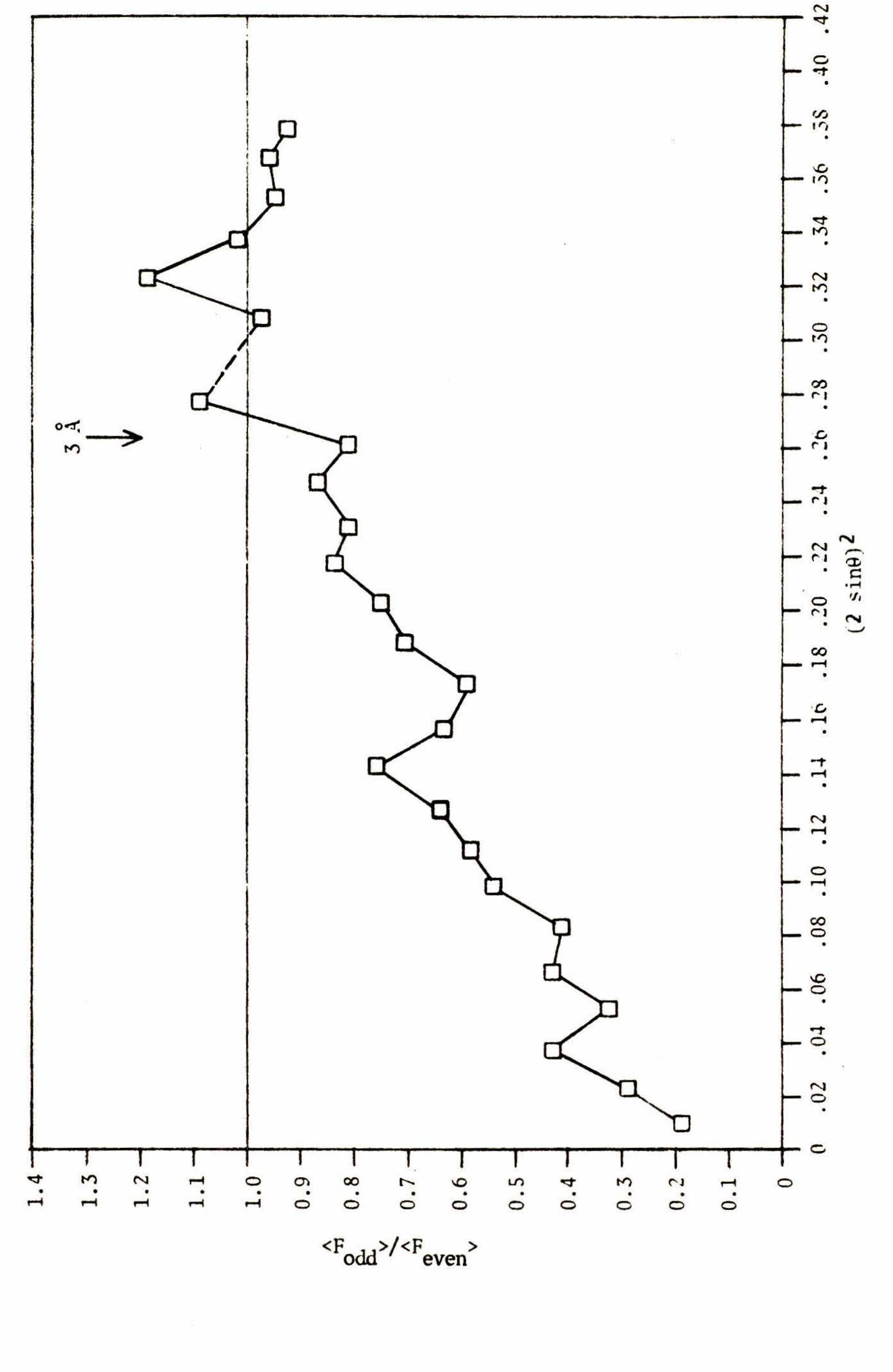
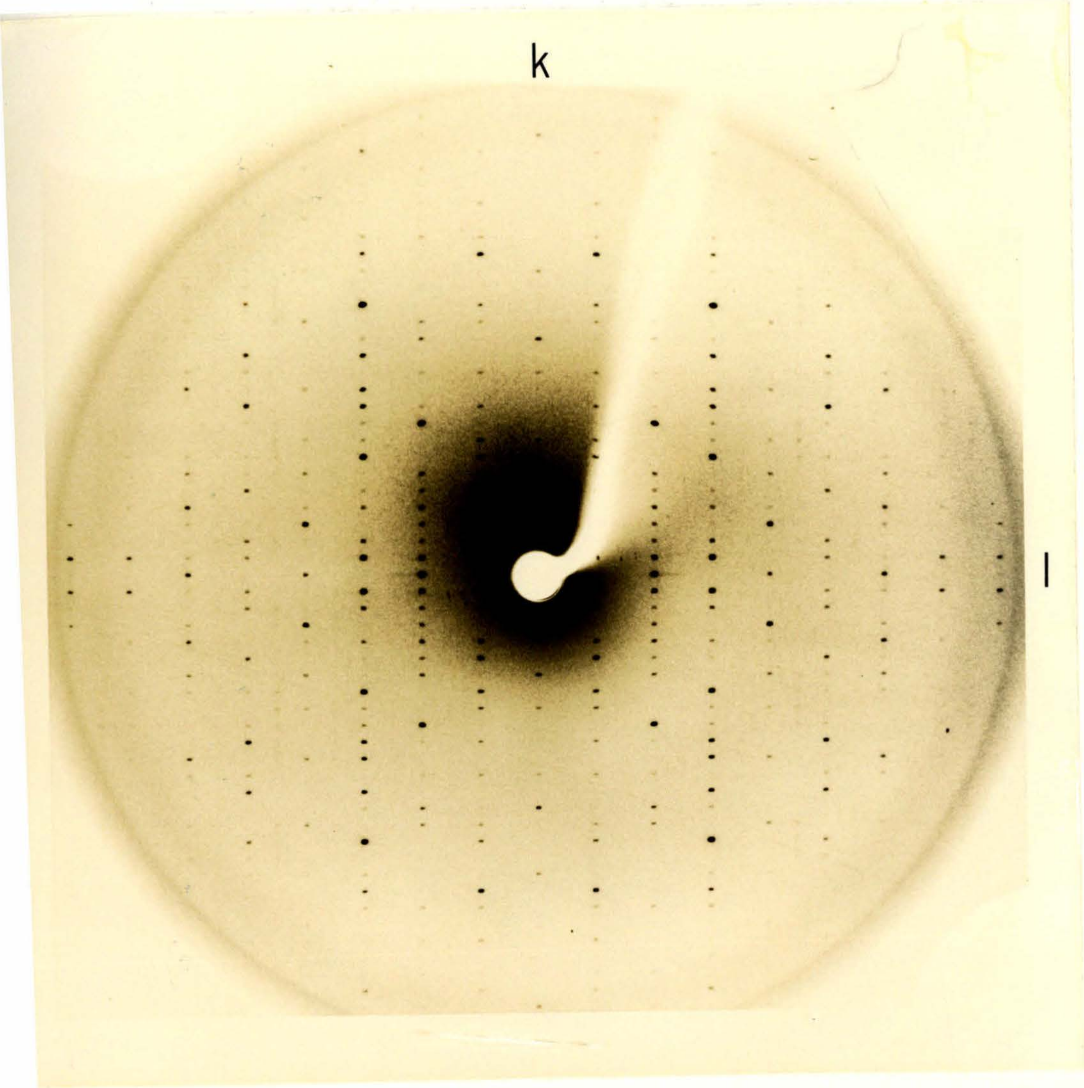

3 Å
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
$<F_{odd}>/<F_{even}>$
0
.02
.04
.06
.08
.10
.12
.14
.16
.18
.20
.22
.24
.26
.28
.30
.32
.34
.36
.38
.40
.42
(2 sinθ)²

Figure 3-6



CHAPTER IV

Data Collection Methods and Data Reduction

"It is a capital mistake to theorise before one has data."

--Sherlock Holmes
(Scandal in Bohemia)

Crystal Mounting

The crystal mounting procedure was similar to the technique described by Timkovich (1975). In this procedure, 0.5 mm thin-wall glass capillary tubes were first rinsed several times to remove any glass powder inside them, then they were oven dried and coated with the hydrophobic coating, Desicote®. The latter was applied by allowing a drop of the liquid to run through the tubes and then blotting the excess. This coating helped to protect the crystals from increases in pH caused by base leaching from the glass, and, in addition, it allowed the liquid around the crystals to be more completely removed.

To mount a crystal, one of these capillary tubes was first sealed at the tip with a low-melting sticky wax[1], then mounted vertically with the same wax in a hollow brass pin which fit into the goniometer head, and filled, using a fine capillary pipet, with the crystal's mother liquor. The height of the crystal above the goniometer head was determined by the distance between the top of the pin and the bottom of the liquid column.

The crystal was then placed in the large, funnel-shaped top of the capillary tube and allowed to settle to the bottom of the liquid. Excess mother liquor was removed from the crystal with the capillary

---

® Beckman Instruments.

[1] Obtained from Kerr Manufacturing Company, Detroit, Michigan.

pipet and a fine (0.2 mm) capillary tube. A small drop of liquid was left above the crystal to keep it from drying out. The top of the tube was then broken off 1-2 cm above the crystal and sealed with sticky wax.

Crystals for the diffractometer were mounted in the manner described above, except that the crystals frequently had to be cut to keep them within the roughly 0.7 mm diameter region of uniform x-ray intensity produced by the diffractometer. The crystals, which were often over 1 mm long, were cut to a length of 0.5 to 0.7 mm with a sharp scalpel blade. This procedure produced a relatively clean cleavage, with little crumbling of the cleaved face, and did not appear to increase crystal disorder in comparison with uncut crystals. These crystals (as well as the uncut ones) were invariably mounted with the needle axis ($\underline{c}$ axis) parallel to the axis of the capillary tube.

Film Data

Data collection

Precession photographs were taken on a Nonius precession camera (crystal-to-film distance = 100 mm) using nickel-filtered copper K_{α} radiation from an Elliott GX-6 rotating anode x-ray generator. The Elliott was equipped with a 200 μm focus cup, and was run at a power setting of 40 kV and 40 mA. The x-ray film used was Kodak No-Screen.

Multiple films were used when needed to record intense reflections.

Although the crystals were unstable at room temperature, no attempt was made to cool them. Since 11° precession pictures of both the h0ℓ and 0kℓ levels could be completed within 24 hours, it was assumed that the crystals would change very little during the exposures. In fact, the cell dimensions did shift slightly towards those of the small-cell form, a process which was accelerated by the addition of some heavy atom salts. The photographs were always taken in the same order using approximately the same exposure times in order to minimize intensity differences between films which could result from this instability.

Alignment of these crystals proved to be another problem. Since the crystals were always mounted with the short c axis parallel to the camera's spindle axis, reflections were widely spaced on the films. Still photographs did not contain enough spots to allow alignment of the crystals, and small angle precession photographs were not readily interpretable. This problem was compounded by the inherent cleanliness of the Elliott x-ray beam, which produced sharp copper K_α and K_β reflections with essentially no white radiation streaks.

The alignment problem was finally solved by using the cone-axis oscillation method described by Climenson and Eisenberg (1974). This technique was the only reliable method by which the crystals could be oriented initially with sufficient accuracy to allow final, fine adjustment with small-angle precession pictures.

Data reduction

The precession films were densitometered on a Syntex AD-1 autodensitometer using the standard Syntex programs. A rectangular aperture 109 x 54 μm was used, and the scans were made large enough to ensure that the whole reflection and its surrounding background were observed.

The data were reduced using programs written for a Data General NOVA 800 minicomputer by Michael Ross, with some modifications by the author and others. With these programs, the measured optical densities were corrected for film non-linearity (Morimoto and Uyeda, 1963) and background scattering, then integrated. Excessively dark reflections were discarded. Lorentz and polarization corrections (Waser, 1951) were then applied, and equivalent reflections averaged to give the final observed intensity.

Structure factor amplitudes, F,[2] from different films were scaled together using the scaling program described in the diffractometer "Data Reduction" section. The R-factor[3] between films from different native crystals was around 5%, not including very weak reflections. This was true even when the two films had large differences in reflection size and shape.

---

[2] Throughout this thesis, F is used for the structure factor amplitude. Where the vector quantity is meant, it is so labeled.

[3] $R = \Sigma|F_1 - F_2|/\Sigma F_1$, where the sums are over all equivalent reflection pairs between films 1 and 2.

Diffractometer Data

Temperature control

The considerable time required to collect diffractometer data made it necessary to cool the crystals during data collection, and for this purpose a cooling device was built which was partially based on the one described by Marsh and Petsko (1973). In this device, air was filtered, dried with anhydrous calcium sulfate (Drierite), and cooled to -20°C by passing it through a 20' x 1/4" O.D. copper coil immersed in a refrigerated bath. The cooled dry air was conducted via an insulated tube to a Mylar chamber mounted on the goniometer head, where it was first passed through the chamber to cool the crystal, then circulated over the outside of the Mylar window to prevent water from condensing on it. The construction of this chamber has been described in detail elsewhere (Krieger and Stroud, 1976).

The temperature inside the chamber was monitored by means of a thermistor positioned about 2 cm above the crystal. No effort was made to regulate the temperature, except by adjustment of the air-flow rate. The diffractometer was in a room maintained at 20°C, and this arrangement was adequate to keep the crystal at a constant temperature of 8 ± 1°C.

Condensation was not a problem except during humid weather, when water condensed around the top of the chamber and dripped down across the window despite the stream of dry air over the outside. A second source of cool, dry air for drying the window might have solved this problem, as a higher air velocity could then have been maintained.

Data collection

Three-dimensional data were collected on a Syntex P$\bar{1}$ four-circle diffractometer. Graphite monochromatized copper K_α radiation was used, with the monochromator mounted on the x-ray tube in the orientation which gave maximum beam divergence in the horizontal (counter) plane. The x-ray beam was collimated, with a final pinhole diameter of 1.0 mm. The beam was tested for uniformity at the crystal position by translating a tiny beryllium acetate crystal across it and observing the diffracted intensity. The beam was found to be uniform ($\pm$5%) over a diameter of about 0.7 mm in both the vertical and horizontal directions. The tube power was set at 40 kV, 20 mA.

The scintillation counter was mounted on an extension arm (crystal-to-counter distance = 40.6 cm), with a helium-filled tube inserted between the crystal and counter to reduce air absorption. Background scatter entering the counter was further reduced by passing the diffracted beam through a variable front aperture (maximum diameter of 4 mm) near the beam stop and a fixed rear aperture (5 mm diameter) near the counter.

Intensity data were collected using fifteen step ω-scans to cover the total peak. The reflections were asymmetric and varied in width both between crystals and at different χ settings for a given crystal, and so it was necessary to select a scan width for each crystal which covered the widest peaks. This resulted in a considerable over-scan of the narrower peaks, sometimes by as much as a factor of two. In general, the reflections from heavy-atom derivative crystals were broader than those from the native crystals. Some derivative crystals had reflections over 2° wide in $\underline{\omega}$. Normally, a scan range of 1° to 1.5° was used, although occasionally a crystal had reflections as narrow as 0.75°. Crystals with reflections over 1.5° wide were usually discarded.

Data were collected in shells of approximately 1000 reflections, with the shells initially increasing in 2θ. Once the decay properties of the crystals were determined, shells of decreasing 2θ were collected in order to minimize the effect of radiation damage on the high angle reflections. A rapid, preliminary scan of each reflection was made, and those reflections below a given minimum intensity were skipped. Counting times on the other reflections depended on the quality of the crystal, but were always in the range 1-5 minutes. Most scan times were around 2 minutes per reflection.

A set of five standard reflections (usually the 0,1,16; 6,15,2; 15,9,4; 11,7,0 and 10,0,0 reflections) distributed through the 2θ range 13° to 32° were collected every 100 to 150 reflections as a check on crystal decay and alignment. They were also used to apply a time-decay correction to the data during data reduction. Data were collected until the high-angle check reflection had dropped 10% in intensity, at which point low-angle reflections (usually the centric planes to 4 Å) were collected for use in scaling the new data to previous data sets. Except in unusual circumstances, no data were used where the check reflections of the same or lower 2θ had fallen by more than 15% from their initial values. By this criterion, the native crystals lasted for about 80 hours of x-ray exposure; derivative crystals were normally usable for somewhat less time.

To increase the amount of useful data collection time, no backgrounds were collected with the reflections. After the intensity data were collected, individual backgrounds were collected by counting for approximately 30 seconds at a position 1° either side (in $\underline{\omega}$) of each reflection. An attempt was made to use interpolated backgrounds (Krieger et at., 1974), but the presence of large, diffuse peaks in the background, probably resulting from crystal disorder, made this technique difficult to apply.

After data collection was completed, the intensity of a reflection at $\chi \simeq 90°$ was measured at intervals of 10° in ϕ, and these values were used to apply an absorption correction to the data. The 0,0,4 reflection was normally used for low-angle data and the 0,0,10 reflection for high-angle data. In one case, both reflections were measured on the same crystal and gave similar results. Normally, the peak-to-trough intensity change was around 30% for these absorption curves.

Data reduction

Data reduction was performed on a Data General NOVA 800 minicomputer with programs written initially by Monty Krieger and modified by several other people. Many of these programs were based on ones written for an IBM 370 computer by Dr. Robert Stroud.

Integrated intensities were computed by summing the 15 contiguous scan steps which gave the highest total[4], and backgrounds were taken as the sum of the background counts on each side of the peaks. Both the integrated intensities and backgrounds were normalized to a one-minute counting time. Any reflection which had:

$$|BG1 - BG2| > 4\ (BG1 + BG2)^{1/2}$$

where BG1 and BG2 are the background counts on each side of the reflection, was discarded. The limit of four standard deviations in the total background count was arbitrary, and was chosen to eliminate only gross errors.

The data were partially compensated for radiation damage by applying a correction factor computed from the observed decay properties of the standard reflections and based on the assumption that the diffracted intensity decreased with time at a rate proportional to $\sin\theta/\lambda$. The multiple measurements of each standard were

---

[4] The diffractometer was able to take up to two extra steps on each side of the peak if it was not centered in the scan range.

fitted with linear least squares by the function:

$$100\ I/I_o = Mt + B$$

where M and B are the slope and intercept[5] respectively, I is the intensity measured at time t and I_o is the initial intensity measurement (at time t_o). Except in very rare cases, the slope (M) of this function was negative. The least-squares values of M and B were then used to calculate the percentage of the original intensity of each standard reflection remaining at the end of data collection. This was done by inserting the total exposure time into the above equation. These five values (one for each standard reflection) were then used to calculate a least-squares slope and intercept for the function:

$$\text{total percent remaining (TPR)} = M'\ (\sin\theta/\lambda) + B'$$

The data were corrected for time-dependent decay by first calculating TPR at the value of $\sin\theta/\lambda$ for each reflection, using the values of M' and B' derived from the standards, and then multiplying the observed intensity by:

$$\left[1 - \left(\frac{100\text{-TPR}}{100}\right)\left(\frac{\text{time of observation}}{\text{total exposure time}}\right)\right]^{-1}$$

In addition to the above calculations, the values of M for the five standards were plotted as a function of $\sin\theta/\lambda$. If one value of

---

[5] The intercept was not forced to be 100% at t = 0 because a better fit to the data was obtained.

M deviated significantly more than the others from the line determined by the five standards, it was assumed to be in error and was omitted from the above time-decay correction calculation. Otherwise, all five standards were used.

The data were corrected for absorption by the method of Stroud, Kay and Dickerson (1974). Transmission factors of the form:

$$T(\phi) = (I_{av}/I(\phi))^{1/2}$$

where ϕ is the measured ϕ angle, I_{av} is the average intensity over the 360° rotation, and $I(\phi)$ is the measured intensity at angle ϕ , were calculated from the intensities measured during the $\chi = 90°$ ϕ-scan mentioned in the "Data Collection" section. The data were then multiplied by the factor T(Inc) T(Ref), where Inc and Ref refer to the ϕ angles of the incident and reflected x-ray beams respectively.

Lorentz and polarization corrections suitable for a four-circle diffractometer using monochromatized radiation (Levy and Ellison, 1960) were applied to the intensity data. Standard deviations for the intensities, normalized to one-minute counting time, were calculated by the formula (Stroud, Kay and Dickerson, 1974):

$$\sigma(I) = [S_c/t_s^2 + (BG1 + BG2)/t_b^2 + (0.02\ S_c/t_s)^2]^{1/2}$$

where S_c is the total scan count and t_s and t_b are the scan and background counting times respectively. The last term in this expression set a lower limit for the error.

The values of $F_{hk\ell}$ from different crystals were scaled together by multiplying each new data set by a scale factor of the form Ke^{BS^2}, where K and B were determined from the least squares fit of the data to a Wilson-type plot of the form (Dickerson et al., 1967):

$$\ln\ [\langle F_1\rangle/\langle F_2\rangle] = \ln K + B\langle S^2\rangle$$

where $S = 2\sin\theta$ and F_1 and F_2 are the structure factor magnitudes for the old and new data sets respectively. The values for F and S were averaged in zones approximately 0.015 wide in S^2. Negative values of F were set to zero for operational reasons, except when they were more than 1.5σ below zero, in which case the reflection was discarded.

After native data to 2.5 Å had been collected, the data sets for the six native crystals were scaled together a second time. Where four or more measurements were made of the same reflection, the Q-test (See Waser, 1966) was used to eliminate any measurements with gross errors. Equivalent reflections were then averaged to give the final value for each $F_{hk\ell}$.

Each derivative data set was first scaled to the native data, then checked to make sure that the heavy-atom occupancy in the crystal used to collect this data set was the same as in previous crystals of this same derivative. This check consisted of calculating [001] projection difference Fourier maps, using as coefficients: $\Delta F = F_1 - F_2$, where 1 and 2 refer to data from two different crystals of a given heavy-atom derivative. Only derivative data sets which did not produce significant peaks (positive or negative) at the known heavy-atom positions (or elsewhere) were averaged; the other sets were discarded or treated separately. The Q-test was applied in those cases where three or more measurements were made for the same reflection. The averaged data set for each derivative was then scaled to the native data once again.

The average R factor[6] between data sets from different native crystals was 5.0% for reflections more than 3σ above background and 7.2% for all reflections. The derivative data were, in general, somewhat worse. These rather high values for R were probably due to the instability of the crystals and the varying extent of disordering among them. The presence of heavy atoms further accentuated these differences between crystals, although except in a few cases this did not seem to be due to different degrees of heavy-atom substitution.

Absolute Scale

A Wilson plot (Wilson, 1942) for the 2.5 Å native α-BuTX data is shown in Figure 4-1. The values of F^2 were averaged in zones 0.015 wide in $(2 \sin\theta)^2$ and unobserved reflections were included at 1.5 times the mean value of σ for each zone. The points above 5 Å resolution (0.1 in $(2 \sin\theta)^2$) were fitted by linear least squares, from

---

[6] See Footnote 3.

which an overall temperature factor of 2.89 (27.5 Å^2) and an approximate absolute scale of 4.53 were derived. Only points above $(2\ \sin\theta)^2 = 0.1$ were used based on the assumption that disordered solvent scattering made no contribution to these intensities, and therefore only protein scattering was present. This value for the absolute scale is probably an underestimate, because the crystals undoubtedly contained ordered water molecules and ions. The absolute scale was not applied to the data.

References

Climenson, C.S. and Eisenberg, D. (1974), J. Appl. Crystallogr. 7, 268-274.

Dickerson, R.E., Kopka, M.L., Borders, C.L. Jr., Varnum, J., and Weinzierl, J.E. (1967), J. Mol. Biol. 29, 77-95.

Forsyth, J.B. and Wells, M. (1959), Acta Crystallogr. 12, 412-415.

Krieger, M., Chambers, J.L., Christoph, G.G., Stroud, R.M., and Trus, B.L. (1974), Acta Crystallogr. A30, 740-748.

Krieger, M. and Stroud, R.M. (1976), Acta Crystallogr. A32, 653-656.

Levy, H.A. and Ellison, R.D. (1960), Acta Crystallogr. 13, 270-271.

Marsh, D.J. and Petsko, G.A. (1973), Acta Crystallogr. 6, 76-80.

Morimoto, H. and Uyeda, R. (1963), Acta Crystallogr. 16, 1107-1119.

Stroud, R.M., Kay, L.M., and Dickerson, R.E. (1974), J. Mol. Biol. 83, 185-208.

Timkovich, R. (1975), Ph.D. Thesis, California Institute of Technology, pp. 23-24.

Waser, J. (1951), Rev. Sci. Instrum. 22, 563-566.

Waser, J. (1966), in Quantitative Chemistry, W. A. Benjamin, Inc., New York, pp. 83-84.

Wilson, A.J.C. (1942), Nature 150, 152.

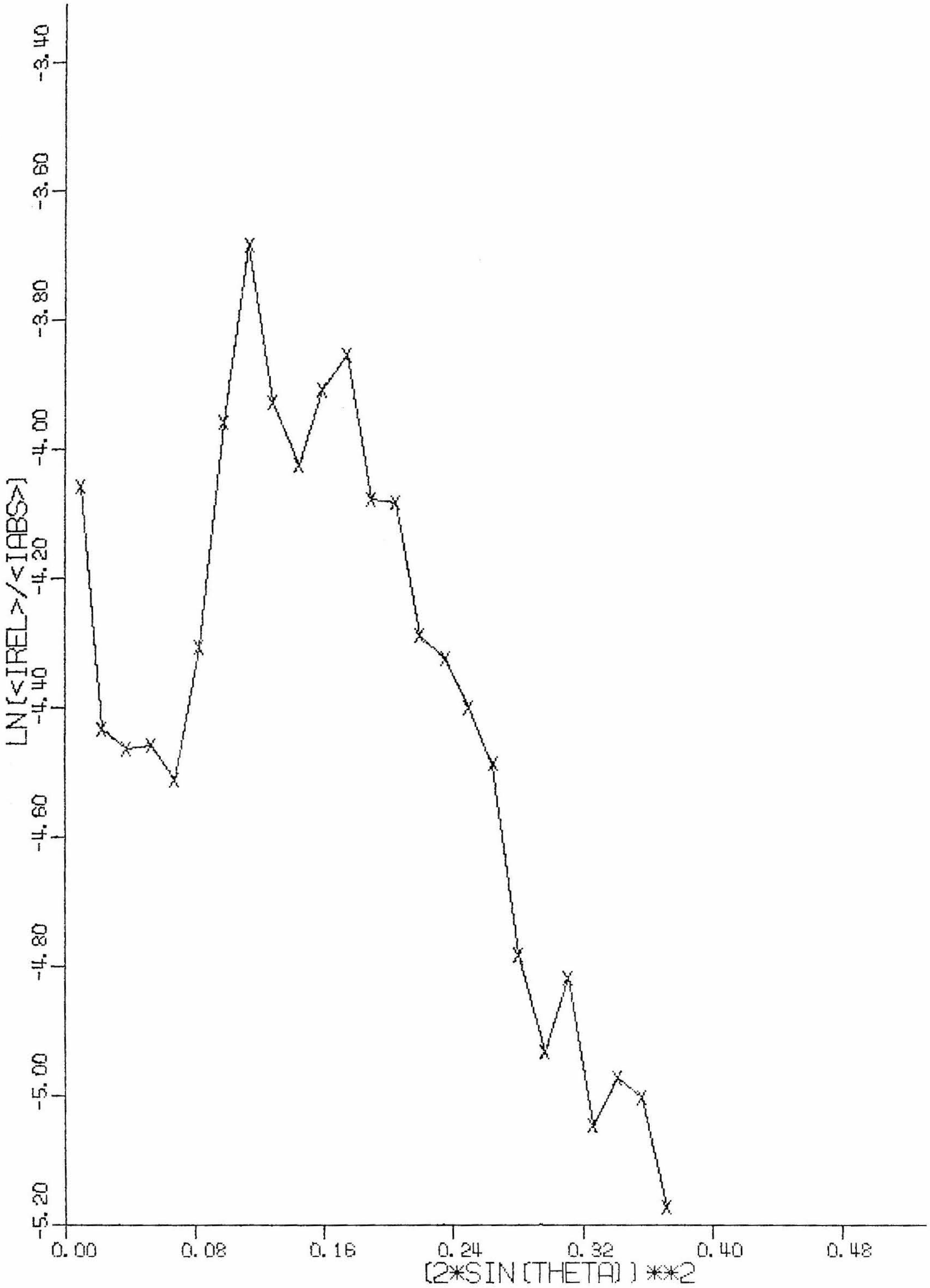
Figure Caption

Figure 4-1: Wilson plot for the 2.5 Å, large-cell, native α-bungarotoxin data. Intensities were averaged in zones 0.015 wide in $(2 \sin\theta)^2$. $<I_{rel}>$ is the mean value of the relative intensity measured from the crystals. $<I_{abs}>$ is the absolute intensity, which was calculated as follows:

$$<I_{abs}> = \Sigma f_i^2$$

where f_i is the atomic scattering factor for atom i in the cell. The sum was computed for the protein atoms only. The values for f_i were computed from the five-parameter function derived by Forsyth and Wells (1959).

Figure 4-1


LN[<IREL>/<IABS>]
-3.40
-3.60
-3.80
-4.00
-4.20
-4.40
-4.60
-4.80
-5.00
-5.20
0.00
0.08
0.16
0.24
0.32
0.40
0.48
(2*SIN(THETA))**2

CHAPTER V

Heavy-Atom Derivatives and Their Solutions

"There are no rules in a knife fight."

--from Butch Cassidy and the Sundance Kid

"It is quite a three-pipe problem."

--Sherlock Holmes

(The Red-Headed League)

Materials and Methods

Materials

Most heavy-atom salts were obtained from either K and K Laboratories, Plainview, New York, or Mallinkrodt Chemical Works. Potassium uranyl pentafluoride was prepared by Mrs. Olga Kallai, who used the procedure described by Mellor (1932).

Derivative screening process

The normal procedure for screening potential heavy-atom derivatives was to prepare a fresh, 10 mM (approximately) solution of the heavy-atom salt in 2 ml of artificial mother liquor, and then to add several small crystals to the solution. These crystals were observed over a period of about seven days, and if they did not crack, 11° $\underline{h0\ell}$ and $\underline{0k\ell}$ precession photographs were taken to determine the unit cell dimensions and to check for any intensity changes. If the cell dimensions were unchanged but no intensity changes were visible, the heavy-atom concentration was increased and the crystals allowed to soak another seven days and then checked again. This process was repeated until the heavy-atom concentration reached 0.1 M, at which point the compound was not tested further.

If the crystals cracked in the 10 mM solution, a fresh solution was prepared with the heavy-atom concentration reduced by a factor of ten, and new crystals were placed in this solution. This procedure was repeated until no cracking was observed, at which point the heavy-atom concentration was increased in small steps until the crystals cracked. New crystals were then placed in the most concentrated heavy-atom solution which did not crack the crystals, allowed to soak for seven days, and precession pictures taken as before. The $\underline{h}0\underline{\ell}$ photographs for all potential derivatives were scanned on a Syntex AD-1 autodensitometer to quantitate any intensity changes.

A spot test was also invented as a possible method for quickly screening potential derivatives. This test, which was based on the observation that heavy atoms frequently bound to proteins in solution and caused them to precipitate, involved placing 1 μl of α-BuTX solution (30 mg/ml in 1.0 M, pH 4.0 acetate buffer) on a glass microscope slide and adding to it 1 μl of a concentrated solution of the heavy-atom salt in water. A blank was also run using acetate buffer containing no protein. The formation of precipitate in the test solution but not the blank was taken as an indication that the heavy atoms bound to the toxin and that the compound was therefore worthy of further testing. A negative result was interpreted as indicating that either no heavy atoms bound, or that they bound to too few sites to precipitate the protein.

Direct Methods

Two different direct-methods programs were tried in attempting to solve for the heavy-atom positions. For two-dimensional data, the program MAGIC, written by Robert Dewar and Allen Stone at the University of Chicago, was used to perform symbolic addition (Karle and Karle, 1966); whereas for three-dimensional data, the tangent refinement program MULTAN74 (Germain, Main and Woolfson, 1971) was used.

Normalized structure factor amplitudes for the heavy-atom structures were calculated by first taking the values of $|\Delta F|$ (= $|F_{PH}-F_P|$, where P and PH refer to native and derivative data, respectively) as approximations to the true magnitudes of the heavy-atom structure factors (Steitz, 1968; Steitz, Fletterick and Hwang, 1973). The $|\Delta F|$'s were then placed on an approximate absolute scale and corrected for thermal motion by means of a Wilson plot. Since the absolute scale could not be determined accurately because the number of heavy-atom sites and their occupancies were not known (sixteen fully occupied sites per unit cell were assumed for the Wilson plot calculation), the values of the normalized structure factor magnitudes, E, were adjusted by a scale factor so that the mean value of E^2 was 1.0. For MAGIC, a single scale was applied to all the data; whereas for MULTAN74, a separate scale was applied to each parity group (i.e., $\underline{h}$ = even, $\underline{k}$ = even, $\underline{\ell}$ = even; $\underline{h}$ = even, $\underline{k}$ = even, $\underline{\ell}$ = odd; etc.).

MAGIC: The starting set of reflections for MAGIC was chosen from a list of the Σ_2 interactions (Karle and Karle, 1966) among the reflections. Those reflections with both high E values and large numbers of Σ_2 interactions with other reflections were assigned symbols and used to initiate the symbolic addition process. Origin-fixing reflections were not chosen until the end of the symbol assignment process so that a reflection whose symbol did not appear in the final list of assigned symbols was not picked as an origin-fixing reflection.

A probability for the symbolic phase assigned to each reflection was calculated from the relation (Cochran and Woolfson, 1955):

$$P = \frac{1}{2} + \frac{1}{2} \tanh [\sigma_3/\sigma_2^{3/2} |E_{hk} \ E_{h'k'\ell'} \ E_{h-h',k-k',\ell-\ell'}|]$$

where $\sigma_3/\sigma_2^{3/2}$ = (number of heavy atoms in the cell)$^{-1/2}$ for the equal-atom structures assumed here. To minimize the possibility of accepting an incorrect symbolic phase, the minimum probability for accepting a given phase was kept as high as possible while still allowing new symbolic phases to be assigned.

Once no more symbolic phases could be assigned, two origin-fixing reflections were chosen (MAGIC was only used for two-dimensional data) and their symbols given phases of (+). E maps were then calculated by assigning various combinations of (+) and (-) to the other symbols and then using the resulting phases for the reflections to calculate the maps. Normally, there were strong indications of the phase relationships among several of the symbols, and so only a small number (usually four or five) of E maps actually had to be calculated.

MULTAN74: MULTAN74 was a much more automatic program than MAGIC, and so it required very little operator intervention. The three origin-fixing reflections were chosen by the program, which used the convergence method of Germain, Main and Woolfson (1970,1971) to make this choice. It also chose three other starting reflections. One of these, the enantiomorph-fixing reflection, was assigned an initial phase of either $\pi/4$ or $3\pi/4$, while the other two, which were chosen to be centric reflections, were given initial phases of either 0 or π, or $\pm\pi/2$ (depending on the centric plane to which they belonged). It was from the various combinations of these signs that the multiple starts for the tangent refinement process were initiated.

The tangent refinement process was fully automatic and resulted in eight sets of phases produced from the two possible initial phases for each of the three starting reflections. E maps were usually calculated only for the four or five phase sets with the highest "combined figure of merit" (as defined in the program write-up).

Heavy-atom refinement

Heavy-atom positions and occupancies were refined on a minicomputer (Data General NOVA 800) by alternate cycles of structure-factor calculation and least-squares refinement (Dickerson *et al.*, 1961 and

1968). Because of the low resolution (5 Å) at which most of this refinement was performed, the temperature and scale factors were not normally allowed to refine.

Isotropic temperature factors, B, were determined for each derivative from the slope of a plot of $\ln \{<|\Delta F|>/<f_H>\}$ versus $(2 \sin\theta)^2$, where $|\Delta F|$ is $|F_{PH}-F_P|$ and f_H is the calculated magnitude of the heavy-atom structure factor. It was assumed in calculating f_H that the atomic scattering factors were constant at all values of 2θ, and so the decrease in f_H caused by the drop in the atomic scattering factors with increasing 2θ was included in the temperature factor. The 4 Å three-dimensional data were used, when available, for this calculation. The values of $|\Delta F|$ and f_H were averaged in zones approximately 0.015 wide in $(2 \sin\theta)^2$

The overall scale factor for each derivative was also calculated from the 4 Å data. This scale factor (Kraut scale) helped to compensate for the additional scattering caused by the presence of heavy atoms in the crystals (Kraut et al., 1962).

The calculated "best" phases (Blow and Crick, 1958), ϕ_B, for the protein structure factor amplitudes, F_P, were used throughout the refinement process. Figures of merit (Dickerson, Kendrew and Strandberg, 1961) were assigned to each F_P and used in subsequent electron density map calculations. The F_P were not weighted for the least-squares refinement. Values of F_P or F_{PH} less than 20 on an arbitrary scale (approximately 0.22 x absolute scale) were not included in the data set used for refinement.

Fourier and Patterson maps

Three-dimensional electron density (Fourier) and Patterson maps (Patterson, 1935) were calculated using a program written for an IBM 370 computer by Dr. George Reeke. Two-dimensional maps were calculated on a Data General NOVA 800 minicomputer using a program written by Stephen Skedzeleski and modified by the author.

Unless otherwise noted, figures of merit and the "best" protein phases were used in calculating all Fourier maps. Difference Fouriers were calculated using the coefficients: $\Delta F = kF_{PH} - F_P$, where $\underline{k}$ is the Kraut scale factor. Difference Patterson maps were calculated using the values of $|\Delta F|^2$ as coefficients.

Results and Discussion

Terminology

For the purposes of this discussion, the term "large cell" refers to native α-BuTX crystals with a 44.7 Å $\underline{c}$ axis; "small cell" refers to these same crystals after they have decayed at room temperature to a cell with a 22.5 Å $\underline{c}$ axis; and "small subcell" refers to the use of only the $\underline{\ell}$ = even reflections (reindexed by dividing $\underline{\ell}$ by two) from the large-cell data set. The latter change helped to eliminate, in part, the problems caused by the presence of the doubled cell along the $\underline{c}$ axis. Refer to Chapter III for further information on the first two crystal forms.

Derivative screening

The results of some of the heavy-atom soaking experiments are shown in Table 5-1. Many of the heavy-atom salts caused the crystals to convert to the small-cell form or hastened their conversion to this form when they were exposed to x-ray radiation. The crystals frequently cracked during this conversion. Large crystals (0.5 x 0.15 x 0.1 mm or larger) were more sensitive to damage than the smaller crystals.

The heavy-atom salts were divided into three categories based on the intensity changes observed on the 11° *h0ℓ* photographs:
1) R[1] factor $\leq 5\%$ -- not a derivative; 2) $5\% < R$ factor $< 10\%$ -- possibly a weak derivative; and 3) $R \geq 10\%$ -- a definite derivative. This classification was, of course, applied only to those compounds which did not cause changes in the unit cell dimensions. Three-dimensional data were normally collected only on derivatives from category 3).

---

[1] $R = \Sigma|F_{PH}-F_P|/\Sigma F_P$, where the sums are taken over all equivalent reflections. Weak reflections were not included in this calculation.

TABLE 5-1

Results of the heavy-atom soaking experiments

Compound	Concentration	Spot Test[1]	Crystal Stability	R Factor[2]	Number of Binding Sites[3]
$K_2Pt(SCN)_4$	10^{-6} M	+	Severe cracking	-	-
$[Pt(NH_3)_4]Cl_2$	0.15 M	-	No change	7%	0
$K_2Pt(NO_2)_4$	10^{-6} M	NT	Severe cracking	-	-
K_2PtCl_4	10^{-6} M	-	Severe cracking	-	-
$(NH_4)_2Pt(CN)_4$	14 mM	+	No change	11%	4
"	35 mM		Unstable[4]	18%	-
$Pt(SO_4)_2$	0.5 mM	-	Crack above 0.5 mM	11%[5]	2
K_2PdCl_6	10^{-4} M	+	Severe cracking	-	-
PdC_2O_4	10^{-5} M	NT	Cracking	-	-
$K_3Ir(CN)_6$	50 mM	-	No change	9%[6]	8
Mersalyl	Saturated	-	No change	9%	-
$HgBr_2$	10 mM	-	Unstable	12%	5
PCMBS[7]	2 mM	+	Slight cracking	9%	-
K_2HgI_4	10^{-4} M	+	Crack over 0.5 mM	6.5%	-
$NaAuCl_4$	0.16 mM	+	Unstable over 10^{-4} M	9.5%	-
$NaAu(CN)_2$	11 mM	+	No change	Visible changes	-
K_3TlCl_6	10 mM	+	No change	13%	10
"	0.1 M		Unstable	35%	-
$K_3UO_2F_5$	22 mM	-	No change	11%	3
Na_2OsO_4	1 mM	NT	Crack over 1 mM	10%[5]	3

---

[1] (-) means no precipitate, (+) means precipitate, NT means not tested. See text for details.

[2] Calculated from the 11° <u>h</u>0<u>ℓ</u> precession data, as described in the text.

[3] Number of sites in the small-subcell asymmetric unit (containing two α-BuTX molecules). This number should be doubled for the large cell. A dash means the number of sites is not known. The process by which these sites were found is described later in this chapter.

[4] Unstable means that the crystals either tended to crack, convert to the small cell, or both.

[5] Determined from 3.5 Å three-dimensional diffractometer data.

[6] Determined from 3.8 Å three-dimensional diffractometer data.

[7] <u>p</u>-chloromercuribenzene sulfonate.

Spot test

The results of the spot test are shown in Table 5-1. In all cases where the spot test gave a positive result (i.e., those cases where the heavy atoms precipitated the α-BuTX from solution), binding of the heavy atoms to the crystals was indicated either by crystal cracking or by significant intensity changes in the diffraction pattern. Three-dimensional data were collected on two of these heavy-atom derivatives (K_3TlCl_6 and $(NH_4)_2Pt(CN)_4$), and both were found to bind to more than one major site per molecule in the crystals. This result tended to confirm the belief those heavy atoms which were able to precipitate the toxin from solution were capable of binding to multiple sites.

Negative results (i.e., those results in which the heavy atoms failed to precipitate the toxin from solution) were more difficult to interpret. Some members of this group ($[Pt(NH_3)_4]Cl_2$, for example) did not bind to the crystals at all, while others (such as $K_3Ir(CN)_6$) bound very weakly to multiple sites, and still others ($Pt(SO_4)_2$, for example) bound tightly to a single site on the toxin molecule. Therefore, it was necessary to use a second assay, such as the ability of these compounds to crack the toxin crystals at moderate concentrations (~ 10 mM) or to produce changes in the diffraction pattern intensities, in order to separate the compounds in this group which bound to the toxin from those which did not. As it turned out, most of the useful derivatives and all the derivatives with a single major site per molecule were produced by compounds from this group.

Thus, the spot test divided the heavy-atom salts into two groups. Based on the limited results presented here, those compounds which gave a positive result tended to form multiple-site derivatives, whereas those compounds which gave a negative result tended to bind to two or fewer major sites per molecule. Compounds which fell in this second group, and which also produced significant intensity changes in the diffraction pattern of the crystals, had a high probability of forming derivatives with only one or two major sites per molecule. The spot test did not, of course, give any information about the stability of the crystals at heavy-atom concentrations high enough to produce usable derivatives.

Derivative crystal properties

All the useful heavy atoms, with the possible exceptions of $Pt(SO_4)_2$ and $HgBr_2$, formed negative ions in solution. This preference for anions probably resulted from the presence at pH 4 of a considerable positive charge on the basic α-BuTX molecules. Despite this charge difference, however, most of the heavy atoms had to be present in relatively high concentrations before they produced significant intensity changes in the diffraction pattern. This could mean that most of the compounds which were capable of binding tightly to the toxin molecules were eliminated because they tended to destroy the crystals.

Those derivatives for which three-dimensional data were collected are listed in Table 5-2, along with a summary of the conditions under which they were produced and a comparison of their cell dimensions to those of the native crystals. As described later in this chapter, all of these derivatives were eventually solved and used to determine the native α-BuTX structure. The properties of each derivative are discussed in detail below.

$\underline{Pt(SO_4)_2 \cdot 4H_2O}$: Platinum sulfate (PtS) was one of the two platinum compounds which still bound to the crystals when used at concentrations low enough to avoid cracking them. It was the only Pt^{IV} complex tried, and this could explain why its behavior differed from that of the other platinum complexes.

The stability of the PtS crystals depended on their soaking time. Large crystals soaked seven days in a 0.5 mM solution remained intact when cold (< 10°C), but began to crack when raised to room temperature for mounting. One small crystal, listed in Table 5-2, survived an eight day soak, but it decayed so rapidly in the x-ray beam that only 5 Å data could be collected. Because of this stability problem, the crystals were normally allowed to soak for only five days. They remained useful for approximately 60 hours of x-ray exposure on the diffractometer.

The variation in the mean value of $|\Delta F|$ as a function of resolution is shown in Figure 5-1a. It is apparent from this graph that the platinum scattering fell rapidly until about 5 Å resolution, then remained at a low, constant level until around 3 Å, at which point it

TABLE 5-2

Three-dimensional derivative data sets

Compound	Concentration	Soaking Time	Resolution[1]	Number of Crystals[2]	Unit Cell Dimensions	Used for Phase Determination
Native	--	--	2.5 Å	6	69.9 x 76.7 x 44.75	
$Pt(SO_4)_2$	0.5 mM	5 days	3.0	2	69.8 x 76.8 x 44.9	*
$Pt(SO_4)_2$	0.5 mM	8 days	5.0	1	69.7 x 76.8 x 44.9	
$(NH_4)_2Pt(CN)_4$	20 mM	2-7 weeks[3]	3.0	5	69.7 x 76.6 x 44.8	*
$K_3Ir(CN)_6$	50 mM	5 weeks	3.8	1	69.9 x 76.7 x 44.7	*
K_3TlCl_6	32 mM	9-12 weeks	3.8	2	69.8 x 76.8 x 44.7	*
K_3TlCl_6	100 mM	13 days	4.1	2	69.8 x 76.8 x 44.7	
$HgBr_2$	10 mM	9 days	5.0	1	69.8 x 76.7 x 44.8	*
Na_2OsO_4	1 mM	~ 30 hours	3.5	1	69.8 x 76.6 x 44.8	*
Na_2OsO_4	1 mM	2 days[4]	3.2	1	69.8 x 76.6 x 44.8	*
$K_3UO_2F_5$	22 mM	4-6 weeks	2.9	4	69.9 x 76.9 x 44.8	*

---

[1] Resolution to which data were collected.

[2] Number of crystals used to collect data to the indicated resultion.

[3] Crystals were assumed to have equilibrated with the heavy-atom solutions after two weeks.

[4] Crystal was soaked 48 hours in plain buffer before data collection was begun.

began to rise sharply. This rise at high resolution indicated an increasing lack of isomorphism between the native and derivative crystals. The rise was confirmed by data from a second crystal, and indicated that the PtS derivative was not useful beyond this point. In fact, the intensity differences beyond 5 Å were so small as to make this derivative nearly useless past 5 Å resolution in any case.

$\underline{(NH_4)_2Pt(CN)_4}$: In contrast to the other Pt^{II} compounds, ammonium tetracyanoplatinate(II) (PtC) bound quite weakly to the crystals. The crystals were often left in solutions of moderate concentration (20 mM) for several weeks without apparent damage, although eventually they began to break apart.

Unfortunately, this complex increased the disorder in the crystals. This made data collection more difficult because the reflections became broader and their intensities weaker. Many crystals had reflections over 2° wide (in $\underline{\omega}$) and were therefore unusable. In addition, the diffraction pattern faded around 3 Å resolution, and so despite the apparent isomorphism between this derivative and the native crystals (Figure 5-1a), no useful data could be collected beyond this point.

The PtC crystals were useful for about 70 hours of x-ray exposure on the diffractometer. The broad peaks and low intensities made long counting times necessary, however, and so a large number of crystals were needed for data collection.

$\underline{K_3Ir(CN)_6}$: Potassium hexacyanoiridate(III) (Ir) was unusual in that it gave off a considerable quantity of gas--possibly HCN--when dissolved in the artificial mother liquor. Since the compound obviously decomposed in solution, the nature of the complex which actually bound to the crystals was even less certain for this compound than for the other heavy-atom compounds.

The crystals were stable for long periods of time in the Ir solutions, but, unlike the PtC crystals, their diffraction patterns did not show increased disordering. The degree of substitution was quite low, however, and so this compound was a rather poor derivative. The crystal used for data collection lasted approximately 60 hours in the diffractometer x-ray beam.

The mean value of $|\Delta F|$ for this derivative showed a sharp rise between 4.5 and 4 Å resolution (Figure 5-1a). No data were collected on a second crystal, so it is uncertain whether this rise resulted from the derivative becoming non-isomorphous at this point or was just a property of this particular crystal.

$\underline{K_3TlCl_6}$: Potassium hexachlorothallate(III) (Tl) seemed at first to be an ideal derivative. It produced large intensity changes in the diffraction pattern, but did not crack or disorder the crystals, even with long soaking times at high concentrations. Surprisingly, these crystals also had an $\underline{h0\ell}$ diffraction pattern which looked very much like the one produced by the small-cell crystal form, except that the cell dimensions were unchanged and the $\underline{\ell}$ = odd rows were still present.

Unfortunately, two major problems arose during data collection. The first was that these crystals were prone to slip in the capillary tubes, thereby necessitating frequent recentering. Attempts to stop this slippage by removing as much liquid as possible from around the crystals were largely unsuccessful, although mounting the crystals in contact with a piece of absorbent paper reduced this problem somewhat.

The second problem was the rapid and erratic intensity changes observed for some reflections. The 10,0,0 check reflection, for example, frequently dropped 10% in intensity during the first ten hours of x-ray exposure, while other check reflections increased in intensity. Since the thallium ions were obviously moving during data collection, the data were of low quality.

This problem was compounded by the fact that the check reflections could not readily be used to indicate the extent of radiation damage to the crystals during data collection or to correct for it afterwards. Because of this, data were arbitrarily collected for 60 hours of x-ray exposure. Application of a time-dependent decay correction was difficult because the check reflections either showed a higher rate of decay at low 2θ than at high 2θ or else the decay rates were virtually random. In some cases, no decay correction was applied.

As shown in Figure 5-1a, the thallium scattering decreased rapidly with increasing resolution. This indicated that the thallium ions were disordered, as would be expected if they were indeed free to move within the crystals. To the resolution of the data, there was no evidence to indicate that the Tl crystals were not isomorphous with the native crystals.

$HgBr_2$: Mercuric bromide (Hg) frequently cracked the crystals and significantly increased their rate of decomposition in the x-ray beam. The crystal on which the three-dimensional data were collected lasted 51 hours in the x-ray beam.

The rapid drop in scattering from the mercury as a function of resolution is shown in Figure 5-1b. Because of this apparently high degree of disorder, as well as the instability of these crystals, higher-resolution data were not collected.

Na_2OsO_4: Sodium osmiate (Os) released a sweet-smelling gas when dissolved in artificial mother liquor. It formed a light-purple solution which darkened over a period of several days and eventually became opaque black. Crystals which were allowed to soak more than one or two days also became opaque black and converted to the small cell form. The black material could not be diffused out of these crystals, although soaking them in osmiate-free buffer did significantly reduce the intensity changes in the diffraction pattern which were produced by the presence of osmium.

The Os crystals which were used for data collection were removed from the osmiate solution before they started to turn black. Unfortunately, the use of short soaking times meant that the crystals did not have time to equilibrate with the surrounding Os solution, and it was found that each Os crystal was different despite all attempts to use exactly the same soaking conditions each time. Difference maps[2]

---

[2] $\Delta F = F_{Os1} - F_{Os2}$, where Os1 and Os2 refer to data from two different Os crystals. Phases for the native data were used to calculate these maps.

calculated between data sets from these various osmiate-soaked crystals showed that the variations were caused by differing degrees of substitution at the same sites.

The osmiate-soaked crystals were different from the other derivative crystals in that they were more stable to x-ray radiation than the native crystals. Their average lifetime in the x-ray beam was around 100 hours. There was some evidence in the difference map[3] for this derivative of changes in the protein structure relative to the native crystals. This movement, which will be discussed in more detail in Chapter VI, may have been responsible for the increased stability of these crystals, and may also have resulted in the observed steady increase in $|\Delta F|$ as a function of $(2 \sin\theta)^2$ (Figure 5-1b).

---

[3] The difference map was calculated using native protein phases and coefficients of $\Delta F = F_{Os} - F_P$. It contained two negative peaks of height (-) 120 and (-) 132, both of which were located 10 Å (in different directions) from the two major Os sites. These negative peaks were at least twice the average for negative values on the map and were the largest negative peaks present. The largest positive peak was 469. It corresponded to an Os site which had been included in the phase calculation.

$K_3UO_2F_5$: Potassium uranyl pentafluoride (uranyl) formed a pale yellow solution from which yellow crystals appeared after about two weeks. These crystals were similar to ones produced by uranyl nitrate solutions within one to two days, and probably indicated that the uranyl pentafluoride was decomposing. Before being mounted, the protein crystals had to be washed in a fresh uranyl pentafluoride solution in order to remove these small crystals.

The protein crystals were stable for several months in uranyl pentafluoride solutions. The derivatized crystals lasted for about 60 hours of x-ray irradiation. The mean value of $|\Delta F|$ for these crystals tended to increase slowly as a function of resolution (Figure 5-1b). The reason for this was not apparent.

Native Patterson[4]

Certain features of the native Patterson proved to be valuable for finding the heavy-atom positions; therefore, a description of the main features of the native Patterson is included at this point. For the purpose of this discussion, the three centric projections, calculated from the 2.5 Å large-cell native data, are reproduced in Figure 5-2. Unless otherwise noted, the peak heights and positions given

---

[4] The reader may find it valuable to read the summary at the end of this and the following sections, as well as the overall summary at the end of this chapter, before tackling the detailed information presented here. Such a procedure may help to prevent confusion caused by the complexity which resulted from the unusual α-BuTX crystal form.

below were determined from the 2.5 Å three-dimensional Patterson map and not from the projections.

Since the desired peak heights were those corresponding to the unusual features of the crystals and not the absolute heights above zero, the F_{000} term was omitted from the Patterson summation. This had the effect of setting the general background of interatomic vectors approximately to zero. Naturally, any fluctuation in this background could have introduced an error into the relative peak heights, which were calculated based on the assumption that the background level under all the peaks was zero.

Peak A: The largest non-origin peak in the native Patterson resulted from the presence within the large unit cell of a subcell. The large cell was apparently produced by the alternation of two similar, but not identical, layers of molecules stacked one above the other along the $\underline{c}$ axis. This alternation of layers, which was first mentioned in Chapter III, resulted in a peak at $\underline{u} = 0$, $\underline{v} = 0$, $\underline{w} = 1/2$ with a height 0.34 that of the origin peak (Peak A in Figures 5-2a and 5-2b).

If the alternating layers of molecules had been identical and separated by exactly c/2 (i.e., if the true unit cell had been half as long in the $\underline{c}$ axis direction), they would have produced a peak equal in height to the origin. Since the observed peak was only one-third the height of the origin, only about one-third of the sixteen molecules in the cell (or, more likely, three of the eight pairs of molecules) were actually separated by exactly c/2. No other definite peaks were present around Peak A, and so the molecules in the remaining five pairs could not have been related to each other by a simple translation. An alternate possibility is that none of the molecules in the eight pairs were separated by an exact translation of c/2 and that Peak A resulted from the overlap of vectors between nonequivalent atoms within these molecule pairs.

Peaks B and C: The second major Patterson peak occurred at $\underline{u} = 0.41$, $\underline{v} = 0.34$, $\underline{w} = 0.28$, and had a height 0.13 that of the origin peak (Peak B in Figures 5-2a-c). It was slightly elongated in $\underline{w}$ and had a second maximum (Peak C in Figure 5-2a) at $\underline{w} = 0.21$ with a height 0.04 that of the origin peak. This double peak can best be explained by first describing how a large peak such as this could arise from a relatively simple molecular arrangement, such as the one illustrated in Figure 5-3.

For the purpose of this discussion, a unit cell with $P2_12_12_1$ symmetry was chosen. In addition, an asymmetric unit was chosen which contained a molecular dimer pair consisting of molecules A1 and B1, which were related to each other by a noncrystallographic twofold axis parallel to the $\underline{z}$ axis. It can be seen from Figure 5-3 that such an arrangement of molecules results in a situation where molecule A1, for example, is rotated 180° by the noncrystallographic axis to give B1, then B1 is rotated another 180° by a crystallographic 2_1 axis parallel to the $\underline{z}$ axis to give B3. Since B3 is just A1 rotated 360°, A1 and B3 are in identically the same orientation (except for a translation) with respect to the cell axes. Similarly, the molecules in the pairs B1-A3, A2-B4, and B2-A4 are in identical orientations.

A Patterson map calculated for this molecular arrangement would ideally have eight peaks on the $\underline{w} = \frac{1}{2}$ section, corresponding to the vectors between the molecules in each of these four pairs. Since each of these peaks would have a height 1/8 that of the origin, and since the vectors between A1 and B3 are identical to those between B1 and A3 (similarly, the vectors from A2 to B4 are identical to those from B2 to A4), the Patterson would in reality contain only four peaks on the $\underline{w} = \frac{1}{2}$ section, each with a height 1/4 that of the origin. This arrangement of molecules is similar to the one found by Steitz (1971) in his crystals of yeast hexokinase B.

It should also be noted in Figure 5-3 that just as the crystallographic 2_1 axis at $\underline{x} = \frac{1}{4}$, $\underline{y} = \frac{1}{2}$ must be midway between the centers of molecules A1 and A3 (or B1 and B3), so must it be midway between the two noncrystallographic axes relating A1 to B1 and A3 to B3. Furthermore, if the centers of molecules A1 and B3 are separated by -0.41 in $\underline{x}$ and 0.66 in $\underline{y}$ (which would result in native Patterson peaks at $\underline{u} = \pm 0.41$, $\underline{v} = \pm 0.34$, $\underline{w} = \frac{1}{2}$), then the noncrystallographic axis between molecules A3 and B3 must also be -0.41 in $\underline{x}$ and 0.66 in $\underline{y}$ from the noncrystallographic axis between molecules A1 and B1. These two conditions require that the noncrystallographic axis relating A1 to B1 be located at $\underline{x} = \frac{1}{4} - (-0.41)/2$, $\underline{y} = \frac{1}{2} - 0.66/2$, and the one relating A3 to B3 be at $\underline{x} = \frac{1}{4} + (-0.41)/2$, $\underline{y} = \frac{1}{2} + 0.66/2$.

Normally, of course, the positions of the molecules are not known and the only information comes from the presence of a large peak at

$\underline{u} = 0.41$, $\underline{v} = 0.34$, $\underline{w} = \frac{1}{2}$ in the native Patterson. If, as in the above example, the space group is $P2_12_12_1$ and the noncrystallographic axes are parallel to the crystallographic $\underline{z}$ axis, then the positions of these axes are limited to four possibilities (1, 2, 3, and 4 in Figure 5-4): $\underline{x} \pm 0.41/2$, $\underline{y} \pm 0.34/2$, where $\underline{x}$ and $\underline{y}$ are the positions of the crystallographic 2_1 axes parallel to the $\underline{z}$ axis. These four possibilities simply correspond to shifts of half a cell in the position of the unit cell origin, and so once the position of one noncrystallographic axis is fixed with respect to the cell origin, the positions of the other three axes are fixed by the space group symmetry (the set of axes labeled 1 in Figure 5-4, for example).

The importance of this result is that all the possible positions for the noncrystallographic axes can be found from the position of the Patterson peak and are limited to a very small number. Therefore, although the Patterson does not contain the information necessary for determining the unit cell axis to which the noncrystallographic axes are parallel, it does allow all the possible positions to be examined systematically. Often, considerations of molecular size and packing prevent the noncrystallographic axes from being parallel to two of the three cell axes. Further, it should be noted that the above discussion applies whether the noncrystallographic axes are parallel to twofold or 2_1 crystallographic axes, and so these results could be applied equally well to noncentrosymmetric orthorhombic space groups other than $P2_12_12_1$.

In the case of α-BuTX, the native Patterson contained a large peak at $\underline{u} = 0.41$, $\underline{v} = 0.34$, $\underline{w} = 0.28$ (Peak B in Figures 5-2a-c), as well as the seven positions related to this one by the symmetry (mmm) of the Patterson. This meant that the unit cell contained pairs of identically oriented molecules separated by the vectors represented by these eight peaks.

The first problem was to determine the crystallographic axis to which the noncrystallographic axes were parallel. Since the large cell space group appeared to be either $P2_12_12_1$ or $P2_12_12$, attempts were made to pack sixteen 26 Å diameter spheres (about the size of an 8000 molecular weight globular protein) into a 69.9 x 76.7 x 44.7 Å unit cell while maintaining the symmetry of one or the other of these space groups. In addition, the molecules had to obey noncrystallographic axes at the positions which were derived from the native Patterson, as described in the previous paragraphs. The results, although not entirely conclusive, strongly suggested that the noncrystallographic axes were parallel to the crystallographic $\underline{c}$ axis of the cell. This being the case, they had to be located ±0.21 in $\underline{x}$ and ±0.17 in $\underline{y}$ from the 2_1 (in $P2_12_12_1$) or twofold (in $P2_12_12$) axes parallel to the $\underline{c}$ axis. Figures 5-5 and 5-6 show one of the four possible locations for the noncrystallographic axes in space groups $P2_12_12_1$ and $P2_12_12$, respectively.

Once the positions and orientations of the noncrystallographic axes were known, it was possible to determine an approximate arrangement for the molecules in the cell. Since the Patterson peak occurred at $\underline{w} = 0.28$, the identically oriented molecules had to be separated in the z direction by 0.28 of the cell. In order for this to be true and still allow the molecules to obey the symmetry of the large-cell space group, the molecules in each dimer pair had to be separated in z by either $\frac{1}{2}$ - 0.28 for space group $P2_12_12_1$ (Figure 5-5) or 0.28 for space group $P2_12_12$ (Figure 5-6). Although the large separation between the two molecules in the dimer pair was somewhat surprising, it was assumed during attempts to solve the heavy-atom Pattersons in three dimensions that one or the other of these molecular arrangements corresponded to the one in the α-BuTX crystals. Unfortunately, this assumption was not entirely correct.

It was discovered by other means[5] that the large unit cell did not belong to either space group $P2_12_12_1$ or $P2_12_12$, but consisted instead of two subcells, each with approximately $P2_12_12_1$ symmetry, stacked above one another along the c axis. Thus, although the x and y positions determined above for the noncrystallographic axes were correct, the relative z positions were not. Because the subcells had 2_1 axes parallel to c which translated the molecules by 1/4 of the large cell in the z direction, not by 1/2 the cell as was assumed in the previous paragraph, the molecules in the dimer pair actually had to be separated in z by only 0.03 of the cell. Such a molecular arrangement is shown in Figure 5-7.

---

[5] See the "Correct three-dimensional derivative solutions" section of this chapter for a description of how this problem was solved.

An interesting side result of the molecular arrangement in Figure 5-7 (as well as those in Figures 5-5 and 5-6) is that the native Patterson would actually contain two unique peaks at $\underline{u} = 0.41$, $\underline{v} = 0.34$. One of these would be at $\underline{w} = 0.28$ and have a height 1/8 the origin height; and the second, which would also be 1/8 the origin height, would be at $\underline{w} = 0.22$. As was mentioned at the very beginning of this discussion, the native Patterson did contain two peaks at this position (Peak B in Figure 5-2a-c). The first was at $\underline{w} = 0.28$ with the expected height (0.13 of the origin height), while the second peak (Peak C in Figure 5-2a) was at $\underline{w} = 0.21$ with a height only 0.04 that of the origin. The small size and shifted position of this second peak may have been a result of the same rotational disorder which reduced the height of Peak A (Figures 5-2a and 5-2b), since both were reduced by about two-thirds from their expected heights. Thus, the molecular arrangement shown in Figure 5-7 is consistent with the native α-BuTX Patterson. It is probably close to the one actually present in the crystals.

It is unfortunate that the native Patterson for the small-cell crystal form was not studied more closely, because the correct packing arrangement might have been discovered much sooner. This three-dimensional Patterson was calculated with 3.5 Å data; and, just as with the large-cell native Patterson, it too contained a large peak at $\underline{u} = 0.41$, $\underline{v} = 0.34$. In this case, however, the peak was on the $\underline{w} = \frac{1}{2}$ section and had a height 0.26 that of the origin peak. It arose from a molecular arrangement similar to the one shown in Figure 5-3, which can be seen to be similar to the arrangement in Figure 5-7, once the factor of two difference between the c axis lengths is taken into account. From this result, it is apparent that one of the major changes involved in the conversion from the large cell to the small cell is a movement in the z direction which changes the spacing between the identically oriented molecules from 0.28 in the large cell to 0.25 (referred to the large-cell c axis) in the small cell.

Peaks D, E, and F: In addition to the above peaks, the [010] projection Patterson (Figure 5-2b) contained three moderately large peaks at $u = 0.46$, $w = 0.25$ (Peak D); $u = 0.50$, $w = 0.08$ (Peak E); and $u = 0.50$, $w = 0.42$ (Peak F). No large peaks appeared at these positions in the other projections or in the three-dimensional Patterson, and so these peaks had to result from the overlap in projection of several small peaks spread out along the v direction.

The first peak (Peak D) may have been related to the peak at $u = 0.414$, $w = 0.277$ (Peak B) which was produced by the noncrystallographic symmetry in the cell. The reason for the existence of this peak was not determined.

The other two peaks (E and F) appeared to be related to each other, since they were separated by a distance of c/2 (after taking into account the mirror symmetry of the Patterson). This meant that they probably arose from the two similar molecular arrangements in the top and bottom halves (along the c axis) of the cell. Since they were at $u = \frac{1}{2}$, they were most likely produced by vectors between molecules related to each other by the crystallographic 2_1 axes parallel to the a axis of the cell. Similar peaks appeared in two of the heavy-atom Pattersons (K_3TlCl_6 and $(NH_4)_2Pt(CN)_4$) as a result of the coincidental overlap of vectors between some of the heavy atoms and their symmetry-related positions. The significance of these peaks in the native Patterson is not known.

Summary: The native Patterson provided two useful pieces of information. The first came from a peak at u, v, w = 0, 0, $\frac{1}{2}$ (Peak A in Figures 5-2a and 5-2b), which indicated that the large unit cell consisted of two subcells which were stacked one above the other along the c axis. From the height of this peak (0.34 that of the origin) and from the lack of any other significant peaks near this one, it was concluded that the molecules in these two halves of the cell were not only related by a translation of c/2, but that some of the molecules in the top half of the cell were also rotated with respect to the equivalent molecules in the bottom half of the cell.

The second important piece of information was deduced from a peak at u = 0.41, v = 0.34, w = 0.28 (Peak B in Figures 5-2a-c). This peak indicated that the asymmetric unit contained pairs of molecules related to each other by a noncrystallographic twofold axis (possibly coupled with a translation) parallel to one of the cell axes. Molecular packing studies indicated that this noncrystallographic axis was probably parallel to the crystallographic c axis. The position of the peak in the Patterson determined where the noncrystallographic axis had to be in relation to the cell symmetry axes, and so it was possible to deduce a potential arrangement for the molecules in the large unit cell (Figure 5-7).

Derivative Pattersons

Considerable effort was expended in trying to solve the derivative difference Pattersons, with little success. The unusual symmetry properties of the crystals considerably complicated these efforts, as did the ambiguity in the large-cell space group. It was possible to find "solutions" in both $P2_12_12_1$ and $P2_12_12$, and it was unclear which answer was correct.

In hindsight, the major difficulties were: 1) the initial derivatives ($Pt(CN)_4^{-2}$, $TlCl_6^{-3}$) had more than one binding site per molecule (i.e., more than four sites per asymmetric unit) and therefore had very complicated Patterson functions; 2) the doubled cell along the c axis produced a Harker section (Harker, 1936) at w = 0.25 instead of w = 0 or w = 0.5 as would be expected for space groups $P2_12_12$ and $P2_12_12_1$, respectively; and 3) the molecules related by the noncrystallographic twofold axis had heavy-atom binding sites at approximately the same z height, which produced double-weight cross-vector peaks on the w = 0 section. These latter peaks were originally interpreted as Harker peaks, and this resulted in several incorrect solutions.

Two-dimensional direct methods

$Pt(CN)_4^{-2}$: The failure to derive an acceptable solution for the heavy-atom positions from the difference Patterson maps necessitated the use of less conventional techniques. It was therefore decided to try to solve the $Pt(CN)_4^{-2}$ derivative (PtC) in two dimensions by calculating PtC phases *a priori* with the symbolic addition program MAGIC. To simplify the problem, the [001] projection was chosen. In this projection, equivalent heavy-atom sites from the top and bottom halves of the cell (along the *c* axis) would appear at approximately the same position. Thus, the problem of the doubled cell was eliminated. Also, both space groups $P2_12_12_1$ and $P2_12_12$ have *pgg* symmetry in this projection, and so the ambiguity in the large-cell space group was eliminated by using plane space group *pgg*.

Normalized structure factor amplitudes, E, were calculated for the 3 Å, PtC *hk*0 data, as described in the Materials and Methods section of this chapter. The results of this normalization procedure are shown in Table 5-3. From the list of Σ_2 relationships (Karle and Karle, 1966) among the 92 reflections with $E > 1$, eight reflections were chosen for the starting set (Table 5-4). Seven of these were assigned symbols, and the eighth was given a phase of (+) by use of the Σ_1 formula (Hauptman and Karle, 1953).

From this starting set, symbolic phases were calculated for other reflections in the data set. Probabilities were calculated for each symbolic phase by the formula given in the Materials and Methods

TABLE 5-3

Distribution of normalized structure factors for the two platinum derivatives[1]

	3 Å PtC Data	4 Å PtS Data	Theoretical[2]	
			Centric	Acentric
$<E>$	0.788	0.786	0.798	0.886
$<E^2>$	1.000[3]	1.000[3]	1.000	1.000
$<E^2 - 1>$	0.997	1.011	0.968	0.736
$E > 1.0$	30.0%	28.2%	32.0%	36.8%
$E > 2.0$	4.0%	5.2%	5.0%	1.8%
$E > 3.0$	0.3%	0.5%	0.3%	0.01%

---

[1] $(NH_4)_2Pt(CN)_4$ (PtC) and $Pt(SO_4)_2$ (PtS). Only hk0 data were used.

[2] Obtained from Stout and Jensen (1972).

[3] These values were set to 1.000 by adjustment of the scale factor.

TABLE 5-4

Starting reflections for the $Pt(CN)_4^{-2}$ MAGIC run

Reflection	Symbol	E
14,9,0	A	3.39
15,6,0	B	2.39
3,4,0[1]	C	2.25
3,5,0	D	2.16
6,13,0	E	2.11
9,3,0[1]	F	2.07
10,0,0[2]	G	2.91
20,0,0[2]	+	1.14

---

[1] Origin-fixing reflection

[2] Structure invariant.

section of this chapter. For the purpose of probability calculation, the unit cell was assumed to contain sixteen platinum atoms. No symbolic phase was accepted if it had a probability less than 0.986.

As new reflections were given symbolic phases, they were used along with the initial set of reflections to calculate symbolic phases for the remaining unphased reflections. This process was continued until no more reflections could be given symbolic phases. In all, 75 of the 92 reflections were given either signs or symbolic phases.

The two starting reflections whose symbols appeared most often in the final list of symbolic phases (excluding structure invariants)--the 3,4,0 and 9,3,0 reflections--were chosen as origin-fixing reflections and assigned signs of (+). The 3,5,0 reflection was also given an arbitrary sign of (+) because its symbol appeared in only two of the final symbolic-phase combinations. In addition, there were strong indications during the symbolic phasing process that G was (-), and so only symbols A, B, and E were left without a fixed sign.

Symbol A was involved in only nine of the final symbol combinations, so the effect of its sign on the calculated E maps was relatively small. Also, there were indications during the symbolic-phasing process that the combination AE was (+), so the number of E maps which had to be calculated was effectively reduced to four. Five were actually calculated.

Most of the E maps contained three strong peaks at approximately the same positions on all the maps. One map, shown in Figure 5-8, contained only these three major peaks. The possible positions for the noncrystallographic twofold axes, as determined from the native Patterson, are also indicated in Figure 5-8; and it can be seen that either the position midway between sites A and B (at $x = 0.21$, $y = 0.33$) or the one between sites A and C (at $x = 0.21$, $y = 0.17$) is an acceptable choice for the location of the noncrystallographic axis. The fact that the PtC solution contained sites which obeyed the noncrystallographic symmetry was encouraging, although the presence of two pairs of sites which we related by different noncrystallographic axes was somewhat disturbing.

The positions and occupancies of the three major PtC sites (A, B, and C in Figure 5-8) were refined[6] in two dimensions in space group $P2_12_12$ (which is, in [001] projection, the same as *pgg*); and from these refined sites, single isomorphous replacement (SIR) phases were calculated for the structure factor amplitudes from the native crystals. The derivative refined poorly ($R_c = 0.64$,[7] overall figure of merit = 0.56 at 5 Å resolution), and the PtC difference Fourier map which was calculated with the resulting SIR phases was quite noisy and gave no definite indication of the existence of other PtC sites.

---

[6] Refer to the Materials and Methods section of this chapter for the details of the heavy-atom refinement technique.

[7] $R_c = \Sigma|F_{PH} - |\vec{F}_P+\vec{f}_H||/\Sigma|F_{PH}-F_P|$, where f_H is the heavy-atom structure factor. The sums are taken over all the centric reflections (Cullis et al., 1961).

$\underline{Pt(SO_4)_2}$: A difference map was also calculated for the $Pt(SO_4)_2$ derivative (PtS) using the SIR phases derived from the PtC solution. This map contained several peaks, and PtS sites corresponding to the five largest were put into the phase-refinement program. Three of the five PtS sites refined to a low occupancy and were discarded. The other two sites, which corresponded to the two largest peaks from the difference map, were at $\underline{x} = 0.26$, $\underline{y} = 0.22$ and $\underline{x} = 0.15$, $\underline{y} = 0.44$. They refined extremely well and gave an R_c of 0.44 and an overall figure of merit of 0.73 at 5 Å. An encouraging result was that these two sites obeyed the same noncrystallographic axis as PtC sites A and B (Figure 5-8).

To test the PtS solution for internal consistency, each of the two PtS sites was refined by itself and used to calculate SIR protein phases which were untainted by the presence of the other site. These phases were then used to calculate PtS difference maps. In each case, the site which was not included in the phase calculation appeared in these maps at about half its original height. This site and the one used for phasing produced the only major peaks present in these maps. This result provided further evidence that the PtS solution was correct.

It was decided to try to provide independent evidence for the correctness of the PtS solution, and so MAGIC was tried again--this time using the 4 Å PtS data ($\underline{hk}0$ reflections only). Normalized structure factors were calculated from the PtS data in the manner described in the Materials and Methods section of this chapter. The distribution of these E values is shown in Table 5-3. These data were input to MAGIC, and symbolic phases were calculated in essentially the same manner as was described for the PtC data. The only two significant differences were that the starting set of data contained only five reflections (Table 5-5) and the minimum probability for accepting a symbolic phase was kept above 0.996. Symbolic phase assignment proceeded smoothly, and 45 of the 75 reflections in the input data set were eventually given either signs or symbolic phases.

The 1,4,0 and 2,5,0 reflections were chosen to fix the origin and were given signs of (+). During the symbolic addition process, it became apparent that the symbol combination ACE had to be (+) as well. Since A was already (+), C and E had to have the same sign. Thus, only four E maps had to be calculated in order to cover all the possible combinations of signs for the three symbols (C, D, E) without fixed signs.

The results of calculating these four E maps were as follows: map 1 contained two peaks and a large, broad area of positive density; map 2 contained only a single peak; map 3 contained two peaks (A and B in Figure 5-9) which corresponded, except for an origin shift, to the PtS sites which were found in the PtS difference map that had been calculated using the SIR phases derived from the PtC solution; and map 4 contained a single peak at nearly the same position as one of the two peaks in map 3. In order to avoid biasing the results in favor of the expected solution (map 3), the sites from the three unique

TABLE 5-5

Starting reflections for the $Pt(SO_4)_2$ MAGIC run

Reflection	Symbol	E
1,4,0[1]	A	2.62
2,5,0[1]	B	2.11
6,2,0[2]	C	2.44
4,2,0[2]	D	2.57
13,12,0	E	3.54

---

[1] Origin-fixing reflection.

[2] Structure invariant.

solutions[8] were input to the phase-refinement program and their positions and occupancies refined for several cycles. PtS difference Fourier maps were then calculated for each solution, and any new sites which appeared on these maps were included in the refinement process. The three sets of sites were then refined to convergence.

It was apparent from the results of this refinement (Table 5-6) that the sites from map 3 were significantly better than the sites from the other two maps. Thus, the PtC MAGIC run and the PtS MAGIC run independently resulted in the same two sites for the PtS derivative. It seemed unlikely that this result could be a coincidence, especially since the two PtS sites refined quite well and gave a good agreement with the PtS [001] projection Patterson map (Figure 5-10).

Summary: The $Pt(SO_4)_2$ derivative (PtS) was solved in the [001] projection in two independent ways. In the first, the symbolic addition program MAGIC was used to solve the $Pt(CN)_4^{-2}$ derivative (PtC) in [001] projection (Figure 5-8), and the resulting three PtC sites were refined and used to calculate single isomorphous replacement (SIR) protein phases. These phases were then used to calculate a PtS difference Fourier map. From this map, two PtS sites were found which refined quite well ($R_c = 0.44$, overall figure of merit = 0.73 at 5 Å).

---

[8] The site in map 4 was just a subset of the sites in map 3, and was therefore not treated as a separate solution.

TABLE 5-6

Results of refining at 4 Å resolution[1]
the three unique $Pt(SO_4)_2$ MAGIC solutions

Map	Number of Sites	R_c	$R_{K(c)}$ [2]	FM[3]
1	3	0.57	0.066	0.59
2	2	0.63	0.073	0.53
3	2	0.50	0.059	0.67

---

[1] Only <u>hk</u>0 data were used (172 reflections).

[2] $R_{K(c)} = \Sigma|F_{PH}-D_H|/\Sigma|F_{PH}|$, where $D_H = |\vec{F}_P+\vec{f}_H|$ and the sums are taken over the centric reflections.

[3] Mean figure of merit for all reflections used in the refinement process.

The second independent solution came from solving the PtS derivative directly using the program MAGIC. Three possible solutions were found and all three were put into the refinement program. The results of this refinement indicated that one solution (Figure 5-9) was clearly better than the other two. The two PtS sites in this solution were the same as the ones that were found in the PtS difference map calculated with SIR phases derived from the PtC solution.

In addition, the PtS sites agreed with the [001] projection Patterson map (Figure 5-10). Thus, four separate results (the PtS and PtC MAGIC solutions, the heavy-atom refinement and the Patterson map) all indicated that the PtS solution was correct in the [001] projection.

Solution of the α-bungarotoxin structure in [001] projection

Derivative solutions: As described in the previous section, there was convincing evidence that the solution of the PtS derivative was correct in the [001] projection, so this derivative was used to solve the other heavy-atom derivatives in this projection by means of difference Fourier maps. To guard against the possibility that the peaks in these maps were simply artifacts of the PtS phasing, no derivative solution was initially accepted unless the sites derived from the difference maps obeyed two conditions. First, each major site had to be related to another major site by the noncrystallographic twofold axis. Second, no major site could coincide with, or be close to, a PtS site. Minor sites were ignored.

The PtC solution violated the first condition since its difference map showed three major sites, one of which had no mate. The $K_3UO_2F_5$ (uranyl) and $HgBr_2$ (Hg) solutions violated the second condition, since both the uranyl sites and two of the four Hg sites were near the positions of the PtS sites. Therefore, these three derivatives were temporarily set aside. The K_3TlCl_6 (Tl) and Na_2OsO_4 (Os) difference maps also contained sites close to the PtS sites, but they were small in comparison to the two major sites for each of these derivatives and were ignored.

As was done with the PtS solution, the Tl and Os solutions were tested for internal consistency by separately refining each site of each derivative and then using the refined position and occupancy to calculate SIR protein phases. These phases were then used to calculate a difference map for the corresponding derivative. In all cases, both the site used for phasing and the site which was not used for phasing appeared in the difference maps.

This process was also tried with the Hg derivative, only in this case the two Hg sites which did not coincide with the PtS sites were used for phasing instead of just a single site. The difference map had peaks corresponding to all four sites--the two which were input plus the two which coincided with the PtS sites. This result indicated that the two sites which coincided with the PtS sites were not artifacts of the PtS phasing.

Finally, cross-difference Fourier maps (Dickerson et al, 1967) were calculated by refining the Os, Tl, and PtS solutions individually and using the resulting SIR protein phases to calculate difference maps for the other two derivatives. In all three cases, the major sites for the other two derivatives returned at or very near their original positions. False sites or ghosts of the sites used for phasing occasionally appeared, but they were considerably weaker than the correct sites.

The above tests indicated that the Os, Tl, and PtS solutions were correct and that the Hg solution was probably correct. Therefore, these four derivatives were refined jointly at 5Å resolution. As can be seen in Table 5-7, the derivatives refined reasonably well, with the exception of Tl, which was missing a large number of minor sites. The mean figure of merit for all the reflections used in the refinement (115 reflections) was 0.94.

The α-bungarotoxin structure: With the multiple isomorphous replacement (MIR) phases derived from the four derivatives, a 5 Å electron density map was calculated in [001] projection for the native α-BuTX molecule (Figure 5-11). As would be expected, there were no identifiable structures present in this map, but there were indications that the map truly represented the toxin structure and was not simply random noise.

First of all, there were regions of positive density, probably representing the toxin molecules, which were clearly related to each other by the noncrystallographic axis. Also, there were regions of

TABLE 5-7

Results of heavy-atom refinement in the [001] projection at 5 Å resolution[1]

Derivative	Site	x	y	A[2]	B[3]	$R_{K(c)}$	R_c
$Pt(SO_4)_2$	1	0.265	0.220	4.3	6.15	0.059	0.40
	2	0.151	0.438	8.3	6.15		
OsO_4^{-2}	1	0.054	0.210	6.8	0.63	0.117	0.53
	2	0.372	0.451	5.4	0.63		
$TlCl_6^{-3}$	1	0.473	0.315	12.1	5.46	0.196	0.72
	2	0.443	0.175	8.9	5.46		
$HgBr_2$	1	0.049	0.454	6.3	5.68	0.115	0.57
	2	0.368	0.224	5.6	5.68		
	3	0.148	0.460	7.3	5.68		
	4	0.263	0.177	2.9	5.68		

---

[1] Refinement was performed in space group $P2_12_12$.

[2] Occupancy on an arbitrary scale.

[3] Isotropic temperature factor (as a function of $(2 \sin\theta)^2$).

negative density, most likely solvent, which separated the regions of positive density from other positive regions related to them by crystallographic symmetry. Finally, the apparent size of the molecules (approximately 37 x 26 Å) was reasonable for an elliptically shaped protein of 8000 daltons.

Thus, the picture this map provided of elliptical toxin molecules related to each other by the noncrystallographic axis and separated from crystallographically related molecules by regions of solvent was entirely consistent with known information about the toxin and about protein crystals in general. This result inspired confidence that the derivative solutions were correct in this projection and, from this firm foundation, the problem of solving the derivatives in three dimensions was attacked.

<u>Summary</u>: The PtS solution from the previous section was used to solve five other heavy-atom derivatives in [001] projection by means of difference Fourier maps. These solutions were checked in several ways to avoid any bias introduced by use of the PtS phases. Four of the derivatives (PtS, Os, Tl, and Hg) were then refined together (Table 5-7), and the resulting MIR phases were used to calculate a 5 Å Fourier map of the native toxin (Figure 5-11). This map contained approximately elliptical regions (37 x 26 Å) of positive density (toxin molecules) which obeyed the noncrystallographic symmetry, and which were separated from crystallographically related positive regions by regions of negative density (solvent regions). The results were consistent with known information about the toxin and about the

structure of protein crystals. Thus, there was strong evidence that the derivative solutions were correct in this projection.

Incorrect three-dimensional derivative solutions

The successful solution of the derivatives in two dimensions instilled confidence that they could readily be solved in three dimensions. This was especially true for PtS, because both the two-dimensional results and the three-dimensional PtS difference Patterson map indicated that there was only a single PtS site per molecule. Therefore, most of the effort in three dimensions was directed toward solving this derivative.

Large-cell, $Pt(SO_4)_2$ "solutions" in $P2_12_12$ and $P2_12_12_1$: At this point, it became necessary to choose either $P2_12_12_1$ or $P2_12_12$ as the large-cell space group, and the latter was initially chosen. As was discussed in connection with the native Patterson, the choice of this space group required that the two toxin molecules related by the non-crystallographic axis (i.e., the dimer pair) be separated by 0.277 of the cell in $\underline{z}$. (See Figure 5-6 for a diagram of this arrangement.) Therefore, a solution to the PtS three-dimensional difference Patterson was sought which both obeyed this condition and agreed with the two-dimensional PtS solution.

A solution was indeed found in which the four PtS sites (one per molecule in the asymmetric unit) agreed reasonably well with the three-dimensional PtS difference Patterson map (Figure 5-12) and obeyed the two conditions mentioned above. Their unrefined fractional coordinates were as follows:

$\underline{x}$	$\underline{y}$	$\underline{z}$
0.15	0.45	0.11
0.15	0.45	0.61
0.26	0.22	0.39
0.26	0.22	0.89

These four sites were refined, using the 5 Å, three-dimensional large-cell data. The results of this refinement ($R_c = 0.52$, overall figure of merit = 0.43) were not as good as the two-dimensional results, but they were sufficiently good that there was no reason to suspect that the solution was incorrect.

The "best" SIR protein phases, ϕ_B, which were derived from refinement of the PtS derivative, were used to solve the other derivatives by means of difference Fourier maps. The peaks on these maps were small--less than about three times the mean noise level--and usually came in pairs separated by approximately one-fourth of the cell in $\underline{z}$. These ghost peaks made it difficult to choose the correct sites, since both peaks in the pairs had approximately the same height.

The presence of these ghost peaks was disturbing, and therefore the PtS solution was tested further for correctness. In the first test, the PtS solution was checked for internal consistency by refining each site independently of the other site and then using the resulting ϕ_B to calculate a PtS difference map. In both cases, the site used for phasing and the site not used for phasing appeared in these maps. Unfortunately, the ghost sites appeared as well, although at a lower level than the "correct" sites.

Despite this somewhat equivocal result. a "round-robin" was tried among three of the derivatives (PtC, PtS, and uranyl). In this round-robin, two of the derivatives were refined together--using the 5 Å large-cell data--and the resulting MIR phases were used to calculate a difference map for the third derivative. In all three cases, the "correct" sites for the third derivative appeared in the difference maps. The ghost sites also appeared, but once again they were weaker than the expected sites. In hindsight, it was probably a mistake to include both the PtS and uranyl derivatives in this round-robin, since the close proximity of the sites for the two derivatives most likely biased the results toward those obtained with the PtS phases alone.

Despite the appearance of the ghost sites in the round-robin difference maps, the fact that the "correct" sites were stronger than the ghosts was sufficiently encouraging to warrant further effort. Therefore, the three derivatives were refined together. The resulting set of MIR phases was then used to calculate a 5 Å electron density map for the native toxin.

The 5 Å map was not encouraging. The positive peaks were at most three times the r.m.s. error level of the map (Dickerson, Kendrew, and Strandberg, 1961), and there was only limited continuity in the positive density. No distinct polypeptide chain was present. There were also regions of negative density which appeared to be solvent, but they did not form a clear outline of either a single molecule or a dimer pair. In addition, the noncrystallographic symmetry was only weakly present.

It was apparent from this map that the derivative solutions were not correct. The most obvious potential source of error was an incorrect choice of the large-cell space group, so the above process was repeated in space group $P2_12_12_1$. Except for slight variations, the refinement results were the same in this space group as in $P2_12_12$, so no native map was calculated. Instead, other means were tried in order to solve the PtS derivative.

Direct methods: The success of direct methods in two dimensions inspired an attempt to try to solve the PtS derivative in three dimensions by means of the tangent refinement program MULTAN74. To circumvent potential problems resulting from the weak $\underline{\ell}$ = odd reflections, the calculations were performed with the small-subcell data. These data were created by taking only the $\underline{\ell}$ = even reflections from the large-cell data and dividing the $\underline{\ell}$ indices by two. Since the 0,0,10 and 0,0,14 reflections in the large-cell data set (which became the 0,0,5 and 0,0,7 reflections in the small-subcell data set) were quite strong, the small-subcell space group was chosen as $P2_12_12$.

Normalized structure factors were calculated from the 4 Å, three-dimensional PtS data by the method described in the Materials and Methods section of this chapter. Surprisingly, their distribution (Table 5-8) agreed much better with the theoretical distribution for a

TABLE 5-8

Distribution of normalized structure factors for the 4 Å $Pt(SO_4)_2$ derivative

		Theoretical	
	Experimental	Centric	Acentric
$<E^2>$	1.000[1]	1.000	1.000
$<E^2-1>$	1.033	0.968	0.736
$<E>$	0.771	0.798	0.886

---

[1] Set to 1.000 by adjustment of the scale factor.

centric space group than an acentric one.

A set of origin-and-enantiomorph-fixing reflections were chosen by the program, as were two other reflections with which to initiate the multiple tangent refinement cycles (Table 5-9). In addition, two reflections (10,0,0 and 0,0,4) whose phases were determined from the Σ_1 formula were included in the starting set of reflections. For the purpose of probability calculation, the unit cell was assumed to contain sixteen platinum atoms.

Phase determination proceeded without incident, and produced eight sets of phases. The four phase sets with the highest "combined figure of merit" (as defined in the program write-up) were used to calculate E maps. Of these, the maps produced from the phase sets with the highest and second highest combined figures of merit were the same except for a shift by half a cell in the position of the origin. Referenced to a common origin, they both contained two sites with the following fractional coordinates:

$\underline{x}$	$\underline{y}$	$\underline{z}$	Height
0.15	0.48	0.23	1016
0.25	0.19	0.78	911

The next highest peak in the map had a height of 391, which was about twice the average noise level.

If the above $\underline{z}$ coordinates are divided by two to convert from the small subcell to the large cell, it can be seen that these sites are virtually identical to the ones which were found by solving the

TABLE 5-9

Starting reflections for the 4 Å $Pt(SO_4)_2$ MULTAN74 run

Reflection	E	Initial Phase
0,5,1[1]	3.30	0
8,0,1[1]	2.52	0
5,6,0[1]	1.58	0
10,0,0[2]	2.50	π
0,0,4[2]	2.48	0
10,0,2[2]	3.28	0 or π
5,6,2[3]	2.09	π/4 or 3π/4
2,5,0	2.03	0 or π

---

[1] Origin-fixing reflection.

[2] Structure invariant.

[3] Enantiomorph-fixing reflection.

large-cell PtS difference Patterson in space group $P2_12_12$. Since the other two E maps did not produce PtS sites which agreed with the two-dimensional solution, the above sites were accepted as the correct solution. This result strongly indicated that the large-cell space group was $P2_12_12$ and that the PtS solution in this space group was correct.

Since the PtS solution seemed to be correct, considerable effort was expended to determine if an incorrect choice of sites for the other derivatives had created the previous problems. It eventually became apparent that no choice of the other sites would result in a usable set of phases, so an attempt was made to sort out the PtS solution in each centric projection.

Centric refinement of the $Pt(SO_4)_2$ sites: The PtS sites were refined separately in each centric plane until a minimum R_c was obtained. The results were quite good for each plane (R_c = 0.35 to 0.40 at 5 Å), but the refined positions were not mutually consistent. Attempts to refine the PtS positions in all three centric phases simultaneously gave unsatisfactory results. Since it was obvious that the PtS solution was still not correct, other means were used to find the correct solution.

Summary: Three-dimensional solutions to the PtS large-cell difference Patterson were found for both potential large-cell space groups ($P2_12_12_1$ and $P2_12_12$). Both of these solutions were consistent with the two-dimensional PtS solution and with the noncrystallographic symmetry of the cell; however, they both produced equivocal results when tested for correctness. An electron density map of the native toxin was calculated with MIR phases derived from the $P2_12_12$ solution, but it was uninterpretable. This result indicated that this solution was not correct. Since the $P2_12_12_1$ solution gave the same type of refinement results, this solution was most probably incorrect as well.

In an attempt to solve the PtS derivative, the tangent refinement program MULTAN74 was run with the small-subcell PtS data. Systematic absences in the small subcell data were consistent with space group $P2_12_12$, but not $P2_12_12_1$; therefore space group $P2_12_12$ was used. The result of this procedure was a solution which agreed with the $P2_12_12$ large-cell PtS solution derived from the PtS difference Patterson. This result reinforced the belief that the large-cell space group was $P2_12_12$. It also suggested that the PtS solution was correct.

Despite the tangent refinement results, all attempts to produce a usable set of protein phases from the $P2_12_12$ PtS solution were fruitless, which meant that the PtS solution was not correct. Therefore, it became necessary to find a method for solving the PtS derivative without knowing the large-cell space group. The procedure by which the PtS derivative was finally solved correctly is described in the next section.

Correct three-dimensional derivative solutions

The $Pt(SO_4)_2$-$UO_2F_5^{-3}$ cross-vector map: The problem of solving the derivatives in three dimensions without knowing the true space group was finally solved by calculating Patterson-type maps using the following coefficients (see Phillips, 1966):

$$(F_{PH_1} - F_P)\ (F_{PH_2} - F_P)$$

where PH_1 and PH_2 are two different heavy-atom derivatives. This type of synthesis gives positive peaks at the positions of the cross-vectors between the sites from the two heavy-atom derivatives, while eliminating the self-Patterson peaks.

Using the large-cell data, cross-vector maps were calculated among several of the derivatives, but most contained too many peaks to be readily interpretable. Fortunately, the PtS-uranyl cross-vectors produced a relatively simple map which contained peaks only on the $\underline{w}$ = 0, 1/4, and 1/2 levels. The $\underline{w}$ = 0 and $\underline{w}$ = 1/2 sections were virtually identical, so only the first two sections are shown in Figure 5-13.

The two-dimensional PtS solution was assumed to be correct, so superposition was performed by using each PtS site in turn as an origin. All the PtS-uranyl cross-vectors were plotted relative to each of these origins, and the points at which the most vectors intersected were taken as the true positions of the uranyl sites in their correct three-dimensional orientation relative to the PtS sites.

Since only the $\underline{x}$ and $\underline{y}$ coordinates of the PtS sites were used in this process, the results were not biased in favor of a particular three-dimensional arrangement of sites.

This procedure gave a clear indication that the uranyl sites were located at the same $\underline{x}$ and $\underline{y}$ positions as had been determined from the two-dimensional uranyl difference map. The surprising result was that all the uranyl and PtS sites in the left half of the cell(i.e., $\underline{y} < 1/2$) were confined to the $\underline{z} = 0$ and $\underline{z} = 1/2$ sections, while the sites in the right half of the cell ($\underline{y} > 1/2$) were confined to the $\underline{z} = 1/4$ and $\underline{z} = 3/4$ sections. This arrangement for the PtS sites is shown in Figure 5-14.

This highly unusual arrangement for the heavy-atom sites was somewhat difficult to accept at first, but it did explain both the PtS and uranyl difference Patterson maps quite well (Figures 5-15 and 5-16 respectively). It also explained the conversion of the crystals from the large-cell form to the small-cell form. As can be seen by doubling the $\underline{z}$ coordinates in Figure 5-14, the PtS sites (and the molecules to which they were bound) were close to obeying a small-cell symmetry of $P2_12_12_1$[9] (dashed unit cell in Figure 5-14). Thus, only small shifts were needed to produce a crystal with a small $P2_12_12_1$ cell. None of the other heavy-atom "solutions" had provided a

---

[9] The arrangement of sites shown in Figure 5-14 does obey $P2_12_12_1$ symmetry because the peaks were assumed to be at exactly $\underline{w} = 1/4$ in the cross-vector map. This was not strictly true, so the sites were not exactly at the $\underline{z}$ coordinates shown in the drawing.

reasonable explanation for the conversion, so this result gave added credibility to this unusual solution.

Tests of the $Pt(SO_4)_2$ solution: Two further checks were made to test this arrangement of the PtS sites. In the first check, the PtS sites in the left half of the cell ($y < 1/2$) (Figure 5-14) were refined in space group P1, using a "unique" hemisphere of data generated from the octant which had actually been collected. Besides the mmm symmetry of the diffraction data and two 2_1 axes parallel to the unit cell a axis, no symmetry was included in this calculation. Thus, the result was not biased with regard to either the positions or the symmetry (2_1 or twofold) of the axes parallel to the b and c cell axes.

The SIR protein phases which resulted from this refinement were used to calculate a PtS difference map in the P1 cell. This map showed only the input PtS sites (at $z = 0$ and $z = 1/2$) in the left half of the cell, while the right half of the cell contained sites on only the $z = 0.23$ and $z = 0.73$ sections. Except for a slight shift in z, these were the same sites as were found from superposition with the PtS-uranyl cross vectors (Figure 5-14).

The second check involved solving the PtS derivative by an independent method. The success of direct methods when using $|\Delta F|$ $(=|F_{PtS} - F_P|)$as an approximation to the true platinum structure factor magnitudes suggested that it might also be possible to use the approximate structure factor magnitudes to solve for the PtS sites by Fourier methods. Since it was desirable to avoid where possible the assumption of any symmetry, a "unique" hemisphere of 5 Å, large-cell,

$|\Delta F_{PtS}|$ data was generated from the unique octant actually collected, and all calculations were performed in space group P1.

In the Fourier procedure, one site[10] was initially placed at an arbitrary position in the P1 cell and phases were calculated from this site for the approximate platinum structure factors. (Note that these were platinum phases, not protein phases.) As expected, the Fourier map calculated with these phases looked much like the PtS difference Patterson map. Since it was known from both the native Patterson and the PtS difference Patterson that there was another site related to this one by ±0.414 in *x*, ±0.342 in *y*, and ±0.277 in *z*, one of the peaks corresponding to these positions was arbitrarily chosen and this site was included in the next cycle of phase calculation. The choice of the second site fixed the cell origin and so no further arbitrary choices could be made.

The introduction of the second site into the phase calculation broke most of the Patterson mirror symmetry, but introduced an inversion center between the two input sites. Since it was not possible to choose arbitrarily between pairs of sites related by this center, both possibilities had to be tested. Therefore, the largest pair of peaks related by the inversion center was chosen and one or the other site in this pair was included in turn in the phase calculation.

---

[10] Each "site" actually consisted of two sites at *z* and *z* + 1/2 because of the doubled cell along *z*. All sites were assumed to have the same occupancy, which was determined from the two-dimensional refinement.

Inclusion of either one of these sites in the phasing caused its twin (and the inversion center) to disappear from the Fourier maps.

Once the inversion symmetry was broken, no other symmetry elements appeared, so only the two solutions resulting from the inversion center had to be examined further. For both branches, new sites were included in the phasing one-at-a-time, and Fourier maps were calculated after each phase cycle until no significant new sites appeared. For each cycle, the largest new site from the Fourier map was included in the phasing, with the stipulations that it had to have a mate separated from it by half the cell in $\underline{z}$, and that it must not have decreased from its initial height as new sites were included in the phase calculation.

The result of this procedure was that, due to the centrosymmetric arrangement of the PtS sites, both branches gave the same solution (except for a shift in the cell origin). The solution found by this technique agreed with both the two-dimensional PtS solution and the three-dimensional arrangement of PtS sites which was derived from the PtS-uranyl cross-vector map.

Solution of the other heavy-atom derivatives: The favorable results from these two tests indicated that the solution for the PtS derivative was correct. The exact arrangement of molecules which resulted in the large cell was still not known, however, so it was decided to use the small-subcell data for the preliminary, low-resolution heavy-atom refinement described below. This decision was based on the assumption that, at low resolution, the molecular arrangement was nearly indistinguishable from true $P2_12_12_1$ symmetry. As was mentioned in Chapter III (see Figure 3-5), the differences between the two subcells from which the large cell was made were on the order of 3 Å. Therefore, the assumption of true $P2_12_12_1$ symmetry for the small subcell was probably reasonable at 5 Å resolution.

The two PtS sites (the small subcell contained two molecules in the asymmetric unit) were refined using the 5 Å small-subcell data. The preliminary results of this refinement (R_c = 0.47, overall figure of merit = 0.51) were significantly better than the refinement results for any of the incorrect solutions described previously. A PtS difference Fourier map calculated with the SIR protein phases derived from this refinement showed no sign of the ghost sites which had plagued the incorrect solutions.

The PtS solution was checked for internal consistency by refining each site by itself and using the resulting SIR phases to calculate a PtS difference map. In both cases, the input site and the site not used in the phasing appeared in the difference map. Once again, no ghost peaks were present.

The Os sites were then found from a difference map calculated with the SIR phases determined from the PtS refinement. This map had four sites related to each other by mirror planes at $\underline{z} = 0$ and $\underline{z} = 1/2$, instead of the expected two sites (one per molecule). Examination of the PtS solution showed the reason for this result. In the small subcell, space group $P2_12_12_1$, all the PtS sites were confined to the $\underline{z} = 0$ and $\underline{z} = 1/2$ sections, so any Fourier map calculated with phases from this derivative would have to be centrosymmetric with mirror planes at these same z levels. The effective space group for the PtS solution was P_{nam}.

To break this mirror symmetry, it was first necessary to determine which pair of the Os sites was correct. This was done by phasing on only one Os site and then calculating an Os difference map to see which sites reappeared. This procedure was tried with all four sites, and in each case only one of the two possibilities for the second site appeared in the difference maps. This internally consistent set of sites obeyed the symmetry of the noncrystallographic axis. These two sites did not fully explain the Os difference Patterson map (Figure 5-17), but they were consistent with it.[11]

The Os solution created another problem since all the sites were close to the z = 1/4 and z = 3/4 sections. Their arrangement was such that they did not produce a mirror symmetry, but they did create inversion centers at z = 0 and z = 1/2 on the 2_1 axes parallel to the z direction. Hence the PtS map calculated with the SIR phases derived from the Os sites contained two sets of PtS sites related by these inversion centers. Since the Os sites were not exactly at z = 1/4 and z = 3/4, one set of the sites was slightly stronger than the other.

---

[11] As was mentioned in the description of the Os derivative, the Os difference map contained two significant negative peaks along with the positive Os peaks. This probably indicated movement of the protein in the Os crystals. Hence, the Os difference Patterson probably contained significant protein-Os vectors as well as Os-Os vectors.

Fortunately, the solution to this problem was not difficult. Since the arrangement of the PtS sites and the Os sites was already known from the previous results, it was only necessary to choose between two mirror-related sets in order to get the Os and PtS sites in their proper relative positions. The choice was arbitrary, since it only involved choosing an enantiomorph. Despite this arbitrary choice, one set of Os sites was first chosen and used to calculate SIR protein phases. A PtS difference map was calculated with these phases, and from this map the stronger set of PtS sites was chosen. The sites from these two derivatives were then refined together.

The joint Os and PtS MIR phases removed most traces of the false centrosymmetry, and these phases were used to solve the Tl derivative by means of a difference map. Once the major Tl sites were found, an attempt was made to phase on just the major site to see if the other Tl sites would appear in a Tl difference map calculated with these phases. When this did not work, the four major Tl sites were refined together, and the resulting SIR phases were used to calculate difference maps for the Os and PtS derivatives. In this case, both PtS sites appeared in the PtS difference map, but only the Os site at $\underline{x} = 0.31$, $\underline{y} = 0.29$, $\underline{z} = 0.23$ appeared in the Os difference map.

The poor quality of the Tl phases did not inspire confidence in the solution, so a PtC difference map was calculated with the Os + PtS MIR phases. Three sites appeared in this map, two of which were related by the noncrystallographic axis and one without a mate. SIR phases were calculated with one of the two paired sites, then these phases were used to calculate a PtC difference map. The input site and its noncrystallographic mate appeared in this difference map, but the third site did not appear. When both members of the noncrystallographic pair were used for phase calculation, the third site appeared in the resulting PtC difference map, but with a height only about twice the average noise level of the map. The three PtC sites were then refined together and the resulting SIR phases were used to calculate difference maps for the PtS and Os derivatives. The expected sites appeared in these difference maps, although one of the two Os sites (originally at $\underline{x} = 0.12$, $\underline{y} = 0.45$, $\underline{z} = 0.74$) was shifted 2.2 Å in $\underline{z}$

These three derivatives (PtS, Os, PtC) were then refined in a "round-robin" fashion, in which two of the derivatives were refined jointly and the resulting MIR phases were used to calculate a difference map for the third derivative. For this procedure, only the two PtC sites related by the noncrystallographic axis were used, since the third site was somewhat suspect. In all cases, the expected sites returned in the difference maps. The PtC derivative was not strong enough to completely break the false symmetry produced by the other two derivatives, so maps calculated with MIR phases from PtC plus one of the other two derivatives contained false sites as well. These sites were always considerably weaker than the correct ones, however, and caused no difficulty.

These three derivatives were then refined together and the resulting MIR phases used to calculate difference maps for the other derivatives. With one exception, the x and Y coordinates for the derivative sites found in the three-dimensional difference maps agreed with those derived from the two-dimensional difference maps. This exception was interesting in that a large PtC peak appeared in the projection, but only a very weak corresponding peak appeared in the three-dimensional map. It turned out that the position of this weak peak corresponded to the center of a solvent channel between toxin molecules, so the large peak in the projection was probably due to the overlap of several weak binding sites distributed along this channel.

Final heavy-atom derivative refinement: The eight derivatives (including two Os derivatives with different occupancies) were refined together to convergence. Virtually all the reflections in the 5 Å small-subcell data set were used for this refinement (562 reflections were used). After the positions and occupancies of the heavy-atom sites had converged, the derivative scale factors were refined at 4 Å resolution. Higher-resolution data was used in this case because the refinement tended to diverge at lower resolution. The scale factors did not change significantly. Finally, the occupancies and positions of the heavy-atom sites were refined to convergence at 5 Å once again.

The final refinement statistics for each derivative are shown in Table 5-10. The refined parameters for each heavy-atom site are given in Appendix B, and the mean lack-of-closure as a function of resolution is shown in Appendix C. The figure-of-merit histogram is shown

TABLE 5-10

Final statistics for the 5 Å, small-subcell heavy-atom refinement in space group $P2_12_12_1$

Derivative	R_c	$R_{K(c)}$	$R_{K(all)}$[1]	E_{rms}[2]	E_{av}[3]	$<\|\Delta F\|>$[4]	Sites per Asymmetric Unit
$Pt(SO_4)_2$	0.45	0.08	0.07	7.4	5.7	12.1	2
$Pt(CN)_4^{-2}$	0.68	0.12	0.10	10.6	8.3	11.9	4
OsO_4^{-2}[5]	0.63	0.14	0.12	12.8	10.0	14.6	3
OsO_4^{-2}[5]	0.66	0.08	0.07	7.2	5.6	8.4	3
$UO_2F_5^{-3}$	0.53	0.07	0.06	6.4	4.9	9.2	3
$TlCl_6^{-3}$	0.53	0.15	0.12	12.8	9.6	17.5	10
$Ir(CN)_6^{-3}$	0.67	0.07	0.06	6.4	5.1	7.8	8
$HgBr_2$	0.62	0.12	0.10	10.9	8.7	13.2	5

TABLE 5-10 (continued)

---

[1] $R_{K(all)}$ is the same as $R_{K(c)}$ (Table 5-6, footnote 2), except the sums are over all reflections.

[2] Root mean square lack-of-closure for all reflections.

[3] Average lack-of-closure for all reflections.

[4] $|\Delta F| = |k \cdot F_{PH} - F_P|$, where k is the Kraut scale.

[5] Two OsO_4^{-2} data sets from crystals with different degrees of Os substitution were included in the refinement.

in Figure 5-18. The final mean figure of merit for all reflections was 0.88, while that for only the centric reflections was 0.96.

Despite the use of an approximate space group and the poor quality[12] of most of the derivatives, the 5 Å refinement results were quite reasonable. In addition, the heavy-atom sites for all the derivatives were consistent with their respective three-dimensional difference Patterson maps. As discussed in the next chapter, the electron density map of the native toxin molecule, which was calculated with MIR phases derived from the refined heavy-atom positions, made chemical and crystallographic sense. Thus, all indications were that the derivatives had been correctly solved.

Summary: The PtS derivative was finally solved in three dimensions by using the two-dimensional PtS sites as origins for superposition with the cross vectors between the PtS and uranyl sites. This procedure resulted in an unusual arrangement of PtS and uranyl sites in which all the sites (PtS and uranyl) in the left half of the cell (i.e., $y < 1/2$) were confined to the $z = 0$ and $z = 1/2$ sections, while the sites in the right half of the cell ($y > 1/2$) were confined to the $z = 1/4$ and $z = 3/4$ sections. The arrangement of these sites indicated that the large cell was actually composed of two subcells which had approximately $P2_12_12_1$ symmetry (Figure 5-14).

---

[12] All the derivatives were poor in the sense that either the data were not good, or the sites all had low occupancies or high temperature factors, or the heavy atoms disordered the toxin molecules, or a combination of these factors.

This result was checked by calculating SIR phases in space group P1 with the PtS sites in just the left half of the cell in order to see where the sites in the right side of the cell would appear. It was also checked by directly solving the PtS derivative in P1 by Fourier techniques. The symmetry of the final result was not forced in either case, but in both tests the results confirmed the solution from the PtS-uranyl cross-vector map.

The other derivatives were then solved by calculating difference Fourier maps with SIR protein phases derived from the PtS solution. These derivative solutions were consistent with both the two-dimensional solutions and the three-dimensional difference Pattersons. A round-robin refinement test was then performed among three of the derivatives (PtS, PtC, Os), and in all cases the correct sites appeared in the difference maps. These maps contained no ghost sites such as the ones which had plagued the incorrect solutions.

Finally, all the derivatives were refined together using the 5 Å small-subcell data. The small-subcell data were used because the arrangement of molecules which produced the large cell was still not precisely known. For the refinement, it was assumed that the small subcell obeyed $P2_12_12_1$ symmetry. This assumption appeared to be valid at 5 Å resolution.

The derivatives refined reasonably well (Table 5-10) and resulted in the refined heavy-atom parameters given in Appendix B. MIR phases were calculated from these refined sites and used to calculate an electron density map for the native toxin. As described in the next chapter, this map appeared to be correct.

Chapter Summary

Heavy-atom derivatives

Seven usable heavy-atom derivatives were found. These were: $Pt(SO_4)_2$ (PtS), $(NH_4)_2Pt(CN)_4$ (PtC), $K_3Ir(CN)_6$ (Ir), K_3TlCl_6 (Tl), $HgBr_2$ (Hg), Na_2OsO_4 (Os), and $K_3UO_2F_5$ (uranyl). There were indications that the PtS derivative bound to only a single site per toxin molecule, and so most efforts were directed toward solving it.

Native Patterson

The native Patterson contained the two important pieces of information. The first was that the large cell contained two nonidentical subcells which were stacked one above the other along the $\underline{c}$ axis. The second piece of information was that the unit cell contained a noncrystallographic axis parallel to one of the cell axes. Considerations of molecular packing indicated that it was parallel to the $\underline{c}$ axis. This result considerably restricted the possible arrangements and locations for the toxin molecules in the unit cell.

Two-dimensional $Pt(SO_4)_2$ solution

The PtS derivative was solved in the [001] projection by two independent methods. In the first, the symbolic addition program MAGIC was used to solve the PtC derivative in this projection, and then SIR protein phases from the PtC sites were used to calculate a difference Fourier map for PtS. In the second method, PtS was solved directly with MAGIC. Both methods gave the same two PtS sites. These sites explained the [001] projection PtS Patterson and obeyed the noncrystallographic symmetry determined from the native Patterson.

The two PtS sites were refined ($R_c = 0.44$, overall figure of merit = 0.73 at 5 Å) and used to calculate SIR protein phases. These phases were used to calculate difference maps for the other derivatives. Four of these derivatives were then refined together and an electron density map of the native toxin was calculated with the MIR phases which were computed from the refined derivative sites. The map was interpretable in terms of elliptical toxin molecules (about 37 x 26 Å) separated by solvent regions.

Thus, all the evidence (the two MAGIC solutions, the PtS difference Patterson, the PtS refinement and the native map calculated with MIR phases derived from the PtS solution) indicated that the PtS solution was correct in the [001] projection.

Incorrect three-dimensional $Pt(SO_4)_2$ solutions

Based on the two-dimensional PtS solution, the three-dimensional PtS difference Patterson was solved in both potential large-cell space groups ($P2_12_12_1$ and $P2_12_12$). The PtS sites from these solutions refined poorly, and difference maps calculated for both PtS and the other derivatives contained ghost sites which could not be eliminated. An electron density map of the native toxin was computed with MIR phases derived from the $P2_12_12$ solution, but it was uninterpretable.

PtS was also solved in the small subcell. The tangent refinement program MULTAN74 was run with the small-subcell data in space group $P2_12_12$, and gave a result which agreed with the solution that was found by solving the PtS difference Patterson in the large cell in space group $P2_12_12$. Despite this agreement, the solution did not give reasonable refinement results or usable phases. Thus, the obvious space groups for the large cell ($P2_12_12_1$ and $P2_12_12$) and the obvious space group for the small cell ($P2_12_12$) did not give acceptable solutions to the PtS derivative. It appeared, therefore, that none of these space groups was correct.

Correct three-dimensional $Pt(SO_4)_2$ solution

The PtS derivative was finally solved by performing superposition with the vectors between the PtS and uranyl sites. The two-dimensional PtS sites were used as origins for this procedure. By doing this, it was found that all the PtS (and uranyl) sites in the left half of the cell ($\underline{y}$ <1/2) were on the $\underline{z}$ = 0 and $\underline{z}$ = 1/2 sections, and all those in the right half of the cell ($\underline{y}$ >1/2) were on the $\underline{z}$ = 1/4 and $\underline{z}$ = 3/4 sections. Their arrangement indicated that the large cell was composed of two subcells, both of which had nearly $P2_12_12_1$ symmetry.

The PtS solution was checked by calculating SIR phases in space group P1 from just the PtS sites in the left half of the cell. A PtS difference map was calculated with these phases, and it contained PtS sites in both the left and right halves of the cell which were in the expected positions. PtS was also solved directly by Fourier techniques in space group P1, and this solution agreed with the one derived from the PtS-uranyl cross-vector map.

The PtS derivative was then refined in the small subcell in space group $P2_12_12_1$ (R_c = 0.47, overall figure of merit = 0.51 at 5 Å) and used to calculate SIR protein phases. With these phases, difference Fourier maps were calculated for the other heavy-atom derivatives. Three of these derivatives were then refined in a round-robin fashion, wherein two of the derivatives were refined together and used to calculate MIR phases. These phases were then used to calculate a difference map for the third derivative. In all cases, the expected sites returned, and there was no sign of the ghost sites which had plagued the incorrect solutions. All these derivative solutions were consistent with their respective two-dimensional solutions and three-dimensional difference Patterson maps.

Finally, all the derivatives were refined together in space group $P2_12_12_1$ using the 5 Å small-subcell data. The MIR phases derived from this refinement were used to calculate an electron density map for the native α-BuTX structure. As described in the next chapter, this map was interpretable in terms of the known properties of the toxin.

Thus, all the evidence indicated that the PtS solution, as well as the solutions to the other derivatives, was (and is!) correct.

References

Blow, D.M. and Crick, F.H.C. (1958), Acta Crystallogr. 12, 794-802.

Cochran, W. and Woolfson, M.M. (1955), Acta Crystallogr. 8, 1-12.

Cullis, A.F., Muirhead, H., Perutz, M.F., Rossmann, M.G., and North, A.C.T. (1961), Proc. Roy. Soc. A265, 15-38.

Dickerson, R.E., Kendrew, J.C., and Strandberg, B.E. (1961), Acta Crystallogr. 14, 1188-1195.

Dickerson, R.E., Kopka, M.L., Varnum, J.C., Weinzierl, J.E. (1967), Acta Crystallogr. 23, 511-522.

Dickerson, R.E., Palmer, R.A., and Weinzierl, J.E. (1968), Acta Crystallogr. B24, 997-1003.

Germain, G., Main, P., and Woolfson, M.M. (1970), Acta Crystallogr. B26, 274-285.

Germain, G., Main, P., and Woolfson, M.M. (1971), Acta Crystallogr. A27, 368-376.

Harker, D. (1936), J. Chem. Phys. 4, 381-390.

Hauptman, H. and Karle, J. (1953), Solution of the Phase Problem: I. The Centrosymmetric Crystal, A.C.A. Monograph number 3.

Karle, J. and Karle, I.L. (1966), Acta Crystallogr. 21, 849-859.

Kraut, J., Sieker, L.C., High, D.F., and Freer, S.T. (1962), Proc. Nat. Acad. Sci. USA 48, 1417-1424.

Mellor, J.W. (1932), A Comprehensive Treatise on Inorganic and Theoretical Chemistry, 12, 78.

Patterson, A.L. (1935), Z. Krist. A90, 517-542.

Phillips, D.C. (1966), Adv. Struc. Res. by Diffraction Meth. 2, 98-100.

Steitz, T.A. (1968), Acta Crystallogr. B24, 504-507.

Steitz, T.A. (1971), J. Mol. Biol. 61, 695-700.

Steitz, T.A., Fletterick, R.J., and Hwang, K.J. (1973), J. Mol. Biol. 78, 551-561.

Stout, G.H. and Jensen, L.H., X-ray Structure Determination: A Practical Guide, p. 321. Macmillan Company, London, 1972.

molecules, such as A1 and B3, which are in the same orientation (except for a translation) with respect to the cell axes. In this example, the noncrystallographic axes are at $\underline{x} = 0.46$, $\underline{y} = 0.17$ (between A1 and B1), as well as the three positions related to this one by $P2_12_12_1$ symmetry. As indicated in the drawing, the distance from A1 to B3 is $\Delta\underline{x} = -0.41$, $\Delta\underline{y} = 0.66$, $\Delta\underline{z} = 1/2$.

Figure 5-4: The four possible sets of noncrystallographic two-fold axes parallel to the $\underline{z}$ axis which would result in native Patterson peaks at $\underline{u} = \pm 0.41$, $\underline{v} = \pm 0.34$, $\underline{w} = 1/2$ for a cell with $P2_12_12_1$ symmetry. The noncrystallographic axes are located 0.41/2 in $\underline{x}$ and 0.34/2 in $\underline{y}$ from the crystallographic 2_1 axes parallel to the $\underline{z}$ axis. Each axis with a given number is related to the other axes with the same number by the space group symmetry.

Figure 5-5: An arrangement of α-bungarotoxin molecules in the large unit cell--space group $P2_12_12_1$--which would produce peaks in the native Patterson at $\underline{u} = \pm 0.41$, $\underline{v} = \pm 0.34$, $\underline{w} = \pm 0.28$. Noncrystallographic axes are located at $\underline{x} = 0.46$, $\underline{y} = 0.17$ plus the three positions related to this one by $P2_12_12_1$ symmetry. For the sake of clarity, only half the molecules are shown. The other half are related to these by a translation of approximately half the cell in $\underline{z}$. (Note that if the other molecules were located exactly half the cell in $\underline{z}$ away from the ones shown, the native Patterson would also contain peaks at $\underline{u} = \pm 0.41$, $\underline{v} = \pm 0.34$, $\underline{w} = \pm 0.22$.)

Figure 5-6: An arrangement of molecules similar to the one in Figure 5-5, except the large α-bungarotoxin cell is assumed to have

$P2_12_12$ symmetry. The other comments in the caption for Figure 5-5 apply to Figure 5-6 as well, except that the noncrystallographic axes are at $\underline{x} = 0.21$, $\underline{y} = 0.17$, plus the three positions related to this one by $P2_12_12$ symmetry.

Figure 5-7: An arrangement of molecules similar to the one in Figure 5-5, except that the large α-bungarotoxin cell is assumed to contain two subcells which have $P2_12_12_1$ symmetry and are stacked one above the other in the $\underline{z}$ direction. This arrangement approximates the true relative arrangement of α-bungarotoxin molecules in the large unit cell. The other comments in the caption for Figure 5-5 apply to Figure 5-7 as well (including the positions of the noncrystallographic axes), except that the molecules shown would produce native Patterson peaks at $\underline{u} = \pm 0.41$, $\underline{v} = \pm 0.34$, $\underline{w} = \pm 0.22$ regardless of whether the molecules located up half a cell in $\underline{z}$ were present or not.

Figure 5-8: 3 Å electron density map of the MAGIC solution for the $Pt(CN)_4^{-2}$ derivative. The map was calculated in space group $P2_12_12$, which, in the [001] projection shown, is equivalent to plane space group pgg. The four possible positions for the noncrystallographic axes are shown, and the correct one (at $\underline{x} = 0.21$, $\underline{y} = 0.33$) is contained in a box. Peaks A, B, and C are discussed in the text. The map was contoured in equal, arbitrary intervals of 75, starting with with the first contour above zero. For reference, the positions of the $Pt(CN)_4^{-2}$ sites which were included in the final three-dimensional refinement are marked with X's. (Note that there is a shift of a/4 in the cell origin when going from space group pgg to the [001] projection of space group $P2_12_12_1$.)

Figure 5-9: 4 Å electron density map of the MAGIC solution for the $Pt(SO_4)_2$ derivative. As with the map in Figure 5-8, this map was computed in space group $P2_12_12$. Peaks A and B are discussed in the text. The origin for this map was shifted by b/2 so that the non-crystallographic axis would coincide with the correct one from the $Pt(CN)_4^{-2}$ map (Figure 5-8). The map was contoured in equal, arbitrary intervals of 75, starting with the first contour above zero.

Figure 5-10: 5 Å $Pt(SO_4)_2$ difference Patterson map in the [001] projection. The peaks labeled A and B are Harker peaks arising from the corresponding sites in Figure 5-9. Site A is at x = 0.26, y = 0.22; site B is at x = 0.15, y = 0.44. Unlabeled peaks marked with an X represent A-B cross vectors. The map was contoured at equal, arbitrary intervals of 20, starting with the first contour above zero. The origin height is shown on the map.

Figure 5-11: 5 Å electron density map of the native α-bungarotoxin structure in [001] projection. As with the map in Figure 5-8, this map was calculated in space group $P2_12_12$. The noncrystallographic axis is at x = 0.21, y = 0.33. The dashed line shows the approximate boundary for the two molecules related by this axis. The map was contoured in equal, arbitrary intervals of 50, starting with the first contour above zero.

Figure 5-12: Two sections from the 5 Å, three-dimensional, large-cell, $Pt(SO_4)_2$ difference Patterson map which are marked with the interatomic vectors generated from the $Pt(SO_4)_2$ solution in large-cell space group $P2_12_12$. These two sections plus the $w = 1/2$ section (which is nearly identical to the $w = 0$ section) contain all the major peaks in the map. The peaks labeled A and B are Harker peaks with each letter representing two sites. The two A sites are at: $x = 0.26$, $y = 0.22$, $z = 0.39$ and $z = 0.89$. The two B sites are at: $x = 0.15$, $y = 0.45$, $z = 0.11$ and $z = 0.61$. Unlabeled peaks marked with an X represent cross vectors among these four sites. The maps were contoured at equal, arbitrary intervals of 15, starting with the first contour above zero. a) $w = 0$ section. The origin height is shown on the map. b) $w = 0.27$ section.

Figure 5-13: Two sections from the $Pt(SO_4)_2$ - $UO_2F_5^{-3}$ cross vector map calculated with 5 Å large-cell data. The $w = 0$ and $w = 1/2$ (not shown) sections are virtually identical, and these three sections ($w = 0, 1/4, 1/2$) contain all the major peaks in the map. The peaks used for superposition are marked with an X. The map was contoured in equal, arbitrary intervals of 50, starting with the second contour above zero. a) $w = 0$ section. b) $w = 1/4$ section.

Figure 5-14: The three-dimensional arrangement of $Pt(SO_4)_2$ sites which was deduced from the $Pt(SO_4)_2$ - $UO_2F_5^{-3}$ cross-vector map. The fractional z coordinates shown in the drawing are relative to the large unit cell. The cell (bold lines) was chosen to be consistent with the two-dimensional maps. For clarity, only the sites in the bottom half of the cell (in z) are shown; the others are related to the ones shown by a translation of approximately half the cell in the z direction. It can be seen that by doubling the indicated z

coordinates, the sites obey $P2_12_12_1$ symmetry (dashed unit cell).

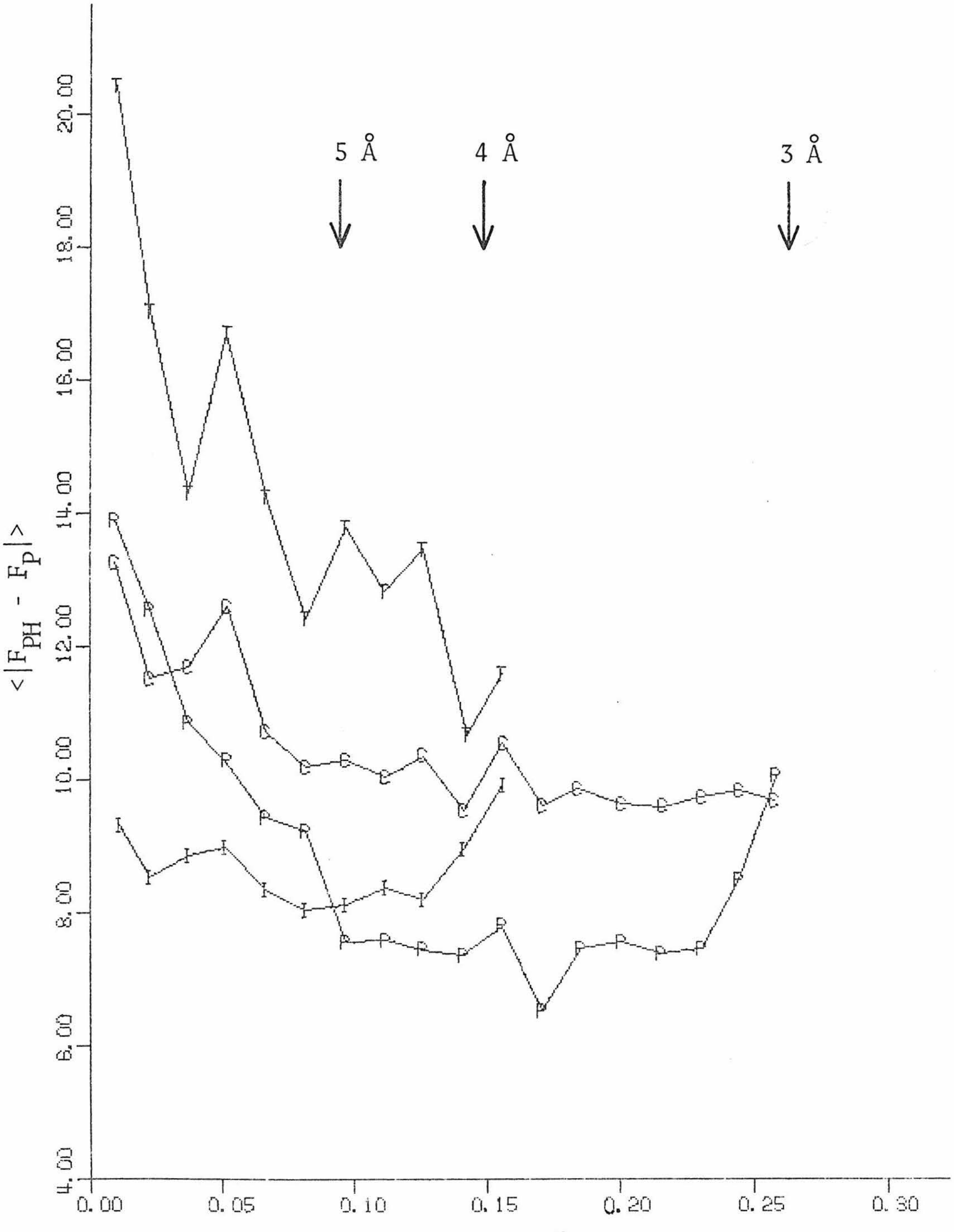
Figure 5-15: Two sections from the 5 Å, three-dimensional, large-cell $Pt(SO_4)_2$ difference Patterson map which are marked with the interatomic vectors generated from the correct $Pt(SO_4)_2$ solution. This map is identical to the one in Figure 5-12, except that the interatomic vectors were generated from four sites with the following coordinates (note that these are relative to the $P2_12_12_1$ cell origin, not the $P2_12_12$ origin used in Figure 5-12): A - x = 0.01, y = 0.22, z = 0 and z = 0.5; B - x = 0.40, y = 0.06, z = 0 and z = 0.5. a) w = 0 section. The origin height is shown on the map. b) w = 0.27 section.

Figure 5-16: Two sections from the 5 Å, three-dimensional, large-cell $UO_2F_5^{-3}$ difference Patterson map which are marked with the interatomic vectors generated from the correct $UO_2F_5^{-3}$ solution. These two sections plus the w = 1/2 section (which is nearly identical to the w = 0 section) contain all the major peaks in the map. The four $UO_2F_5^{-3}$ sites used to generate the interatomic vectors are at: A - x = 0.06, y = 0.27, z = 0 and z = 0.5; B - x = 0.35, y = 0.11, z = 0 and z = 0.5. As in the previous Patterson maps, only the Harker peaks are labeled with these letters. Cross-vector peaks are only labeled with an X. The map was contoured at equal, arbitrary intervals of 15, starting with the first contour above zero. a) w = 0 section. The origin height is shown on the map. b) w = 0.25 section.

Figure 5-17: Three Harker sections from the 5 Å, large-cell, OsO_4^{-2} difference Patterson map which are marked with the interatomic vectors generated from the correct OsO_4^{-2} solution. The four OsO_4^{-2} sites used to generate these vectors are at: A - x = 0.30, y = 0.29, z = 0.11 and z = 0.61; B - x = 0.12, y = 0.45, z = 0.37 and z = 0.87. As in the previous Patterson maps, only the Harker peaks are labeled with these letters; cross-vector peaks are only labeled with an X. The map was contoured in equal, arbitrary intervals of 30, starting with the first contour above zero. The origin height was not recorded. a) u = 0.5 section. b) v = 0.5 section. c) w = 0.25 section.

Figure 5-18: Figure of merit histogram for the 5 Å, small-subcell refinement in space group $P2_12_12_1$. The cross-hatched area is the histogram for only the centric reflections. The mean figure of merit for all reflections was 0.88 and for centric reflections was 0.96.

Figure 5-1a



<|F_PH - F_P|>
20.00
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
0.00
0.05
0.10
0.15
0.20
0.25
0.30
(2 sinθ)²
5 Å
4 Å
3 Å
T
C
P
I

Figure 5-1b

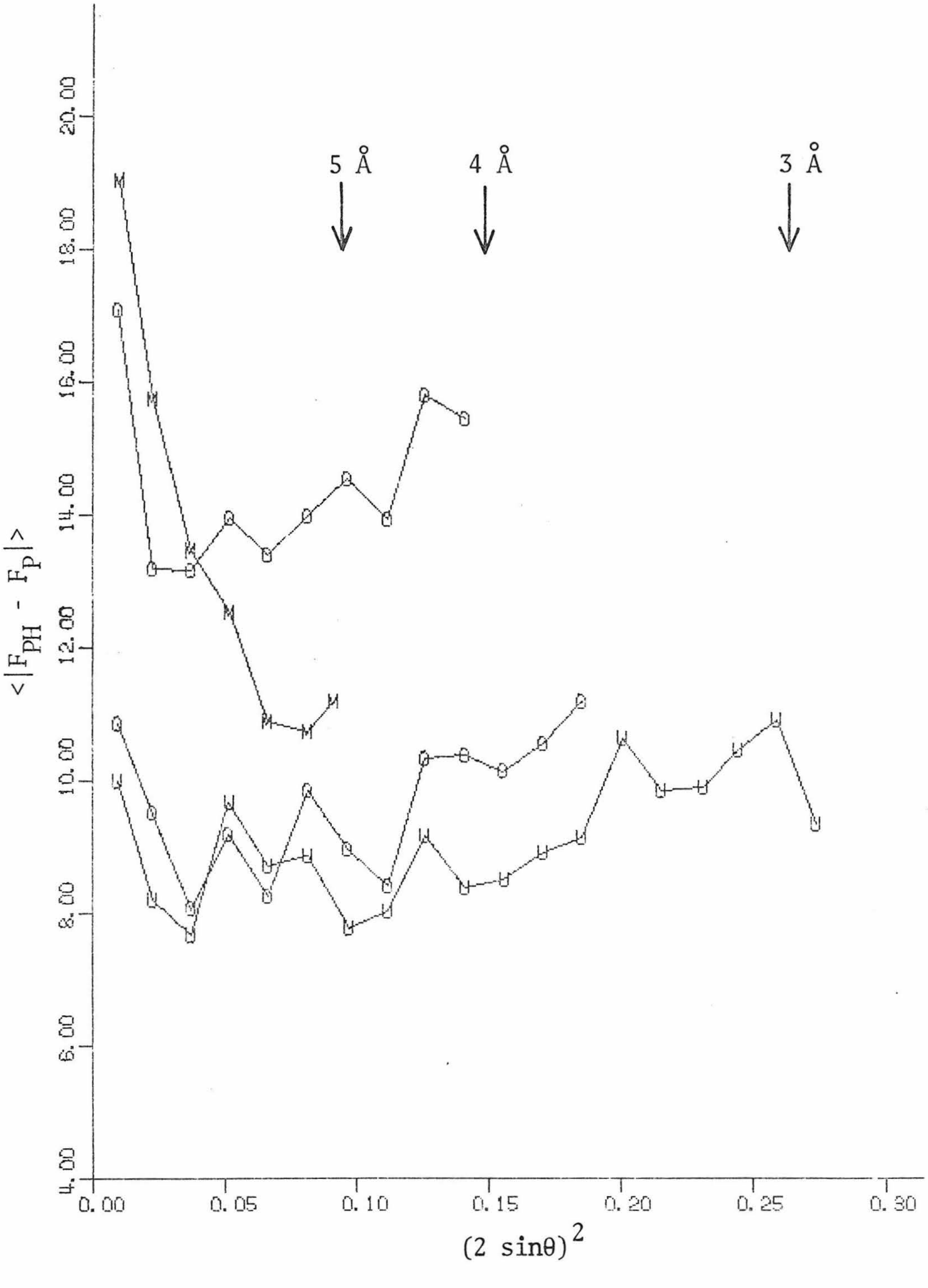

5 Å
4 Å
3 Å
<|F_{PH} - F_P|>
20.00
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
0.00
0.05
0.10
0.15
0.20
0.25
0.30
(2 sinθ)^2
M
O
U

Figure 5-2a

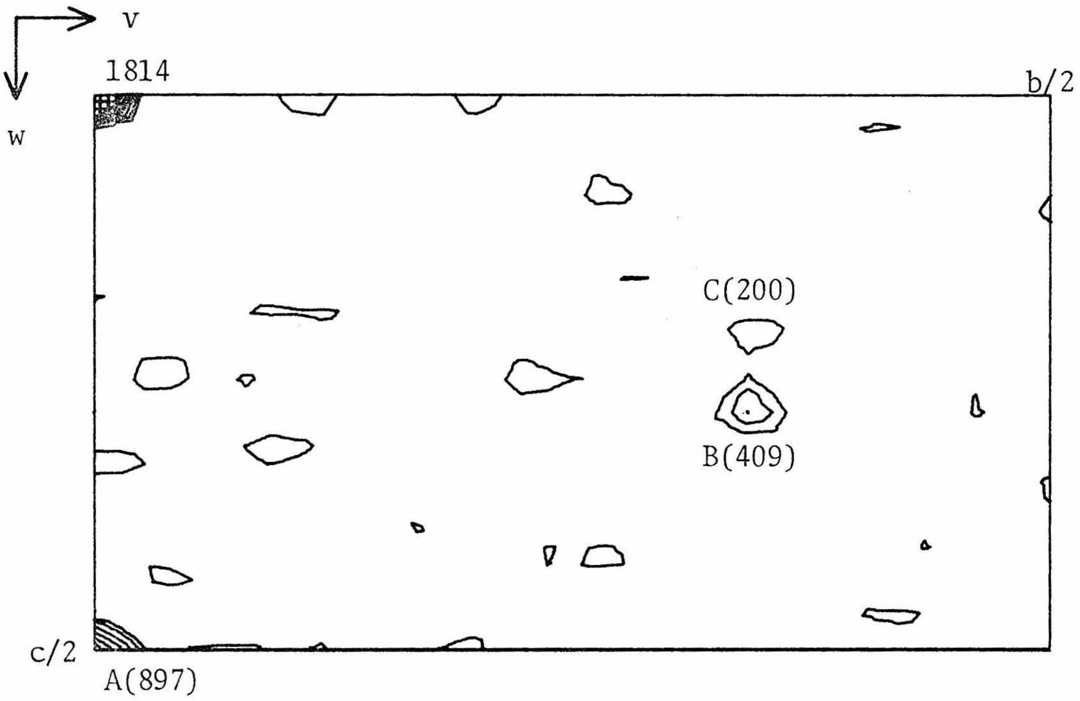


Figure 5-2b

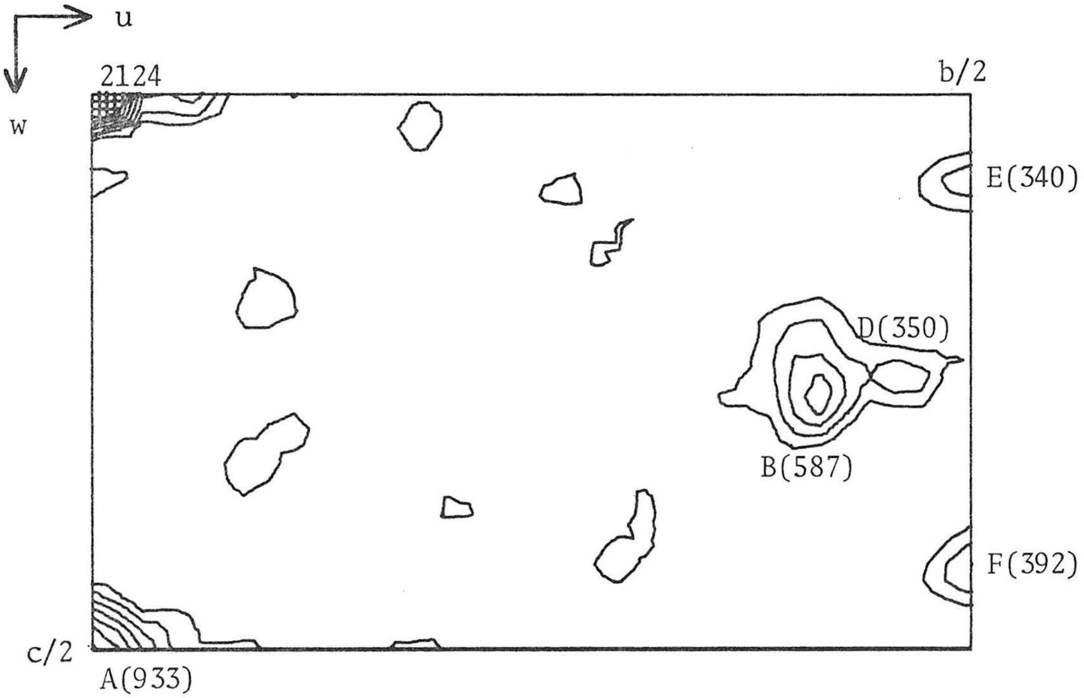


Figure 5-2c

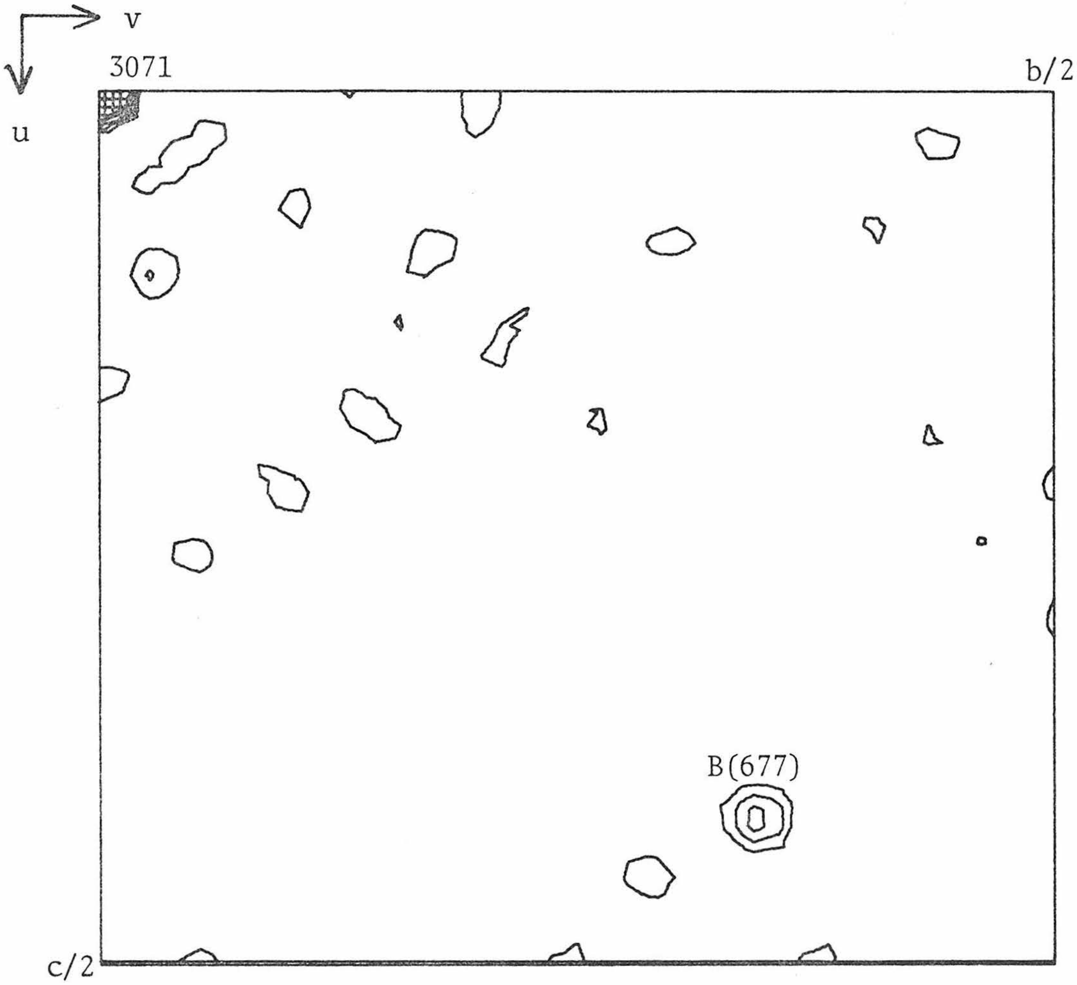


Figure 5-3

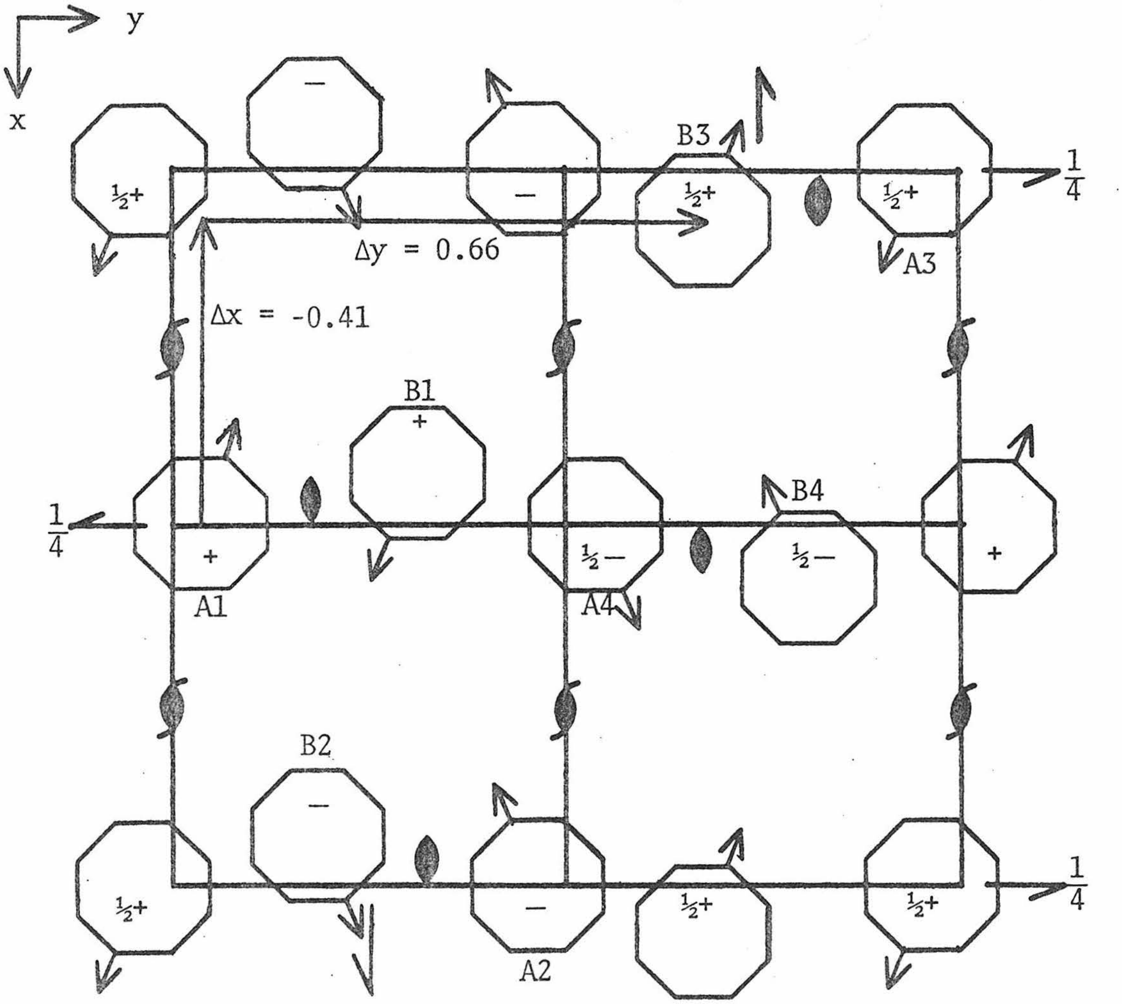


Figure 5-4

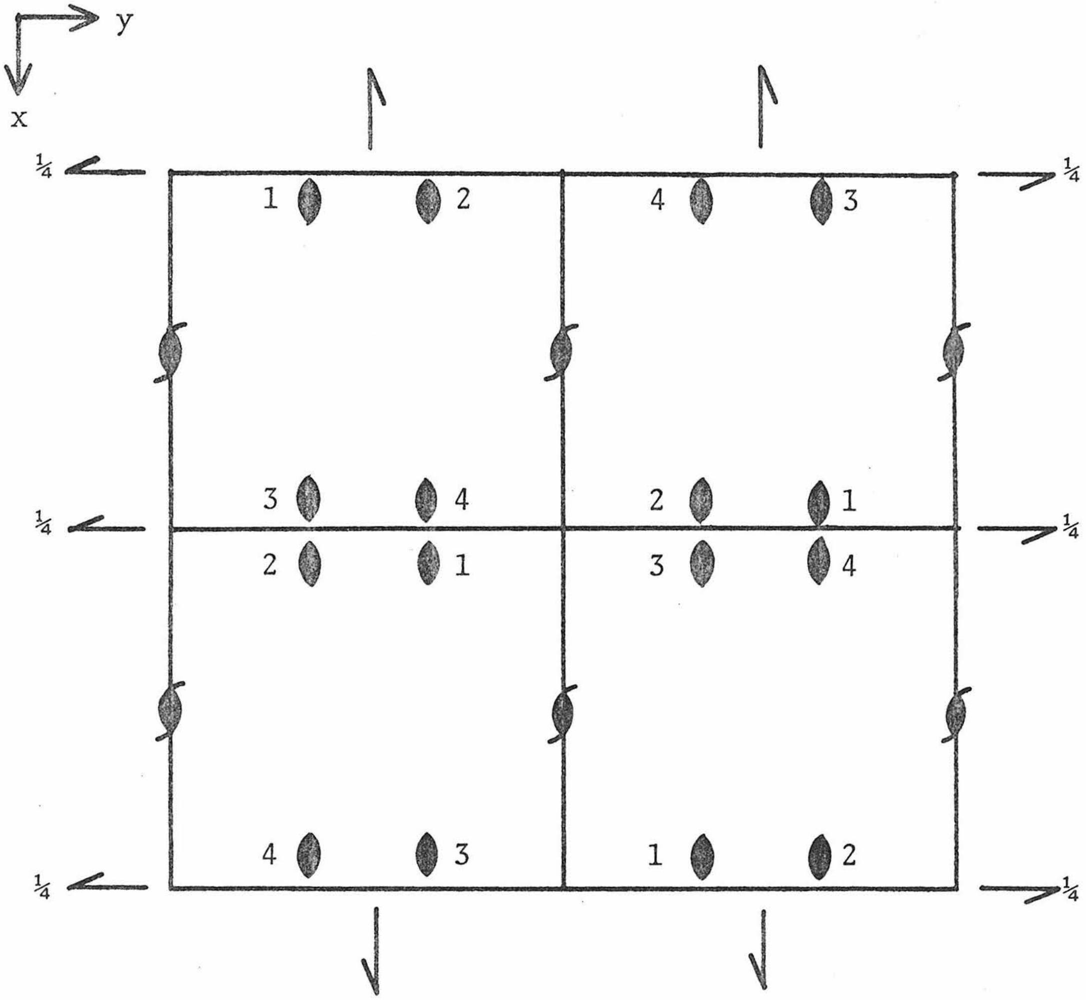


Figure 5-5

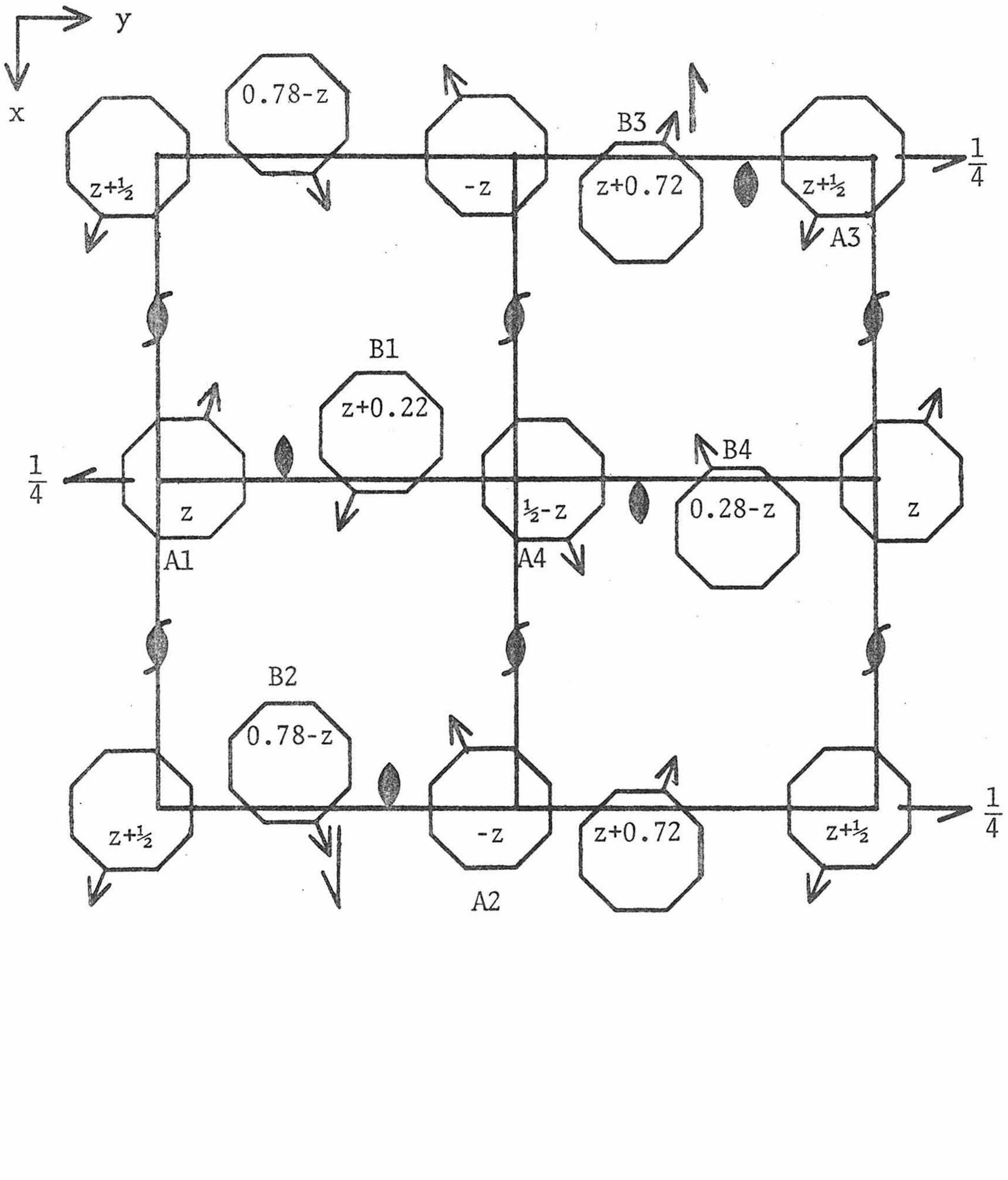

y
x
0.78-z
B3
z+½
-z
z+0.72
z+½
A3
1/4
B1
z+0.22
B4
1/4
z
½-z
0.28-z
z
A1
A4
B2
0.78-z
z+½
-z
z+0.72
z+½
1/4
A2

Figure 5-6

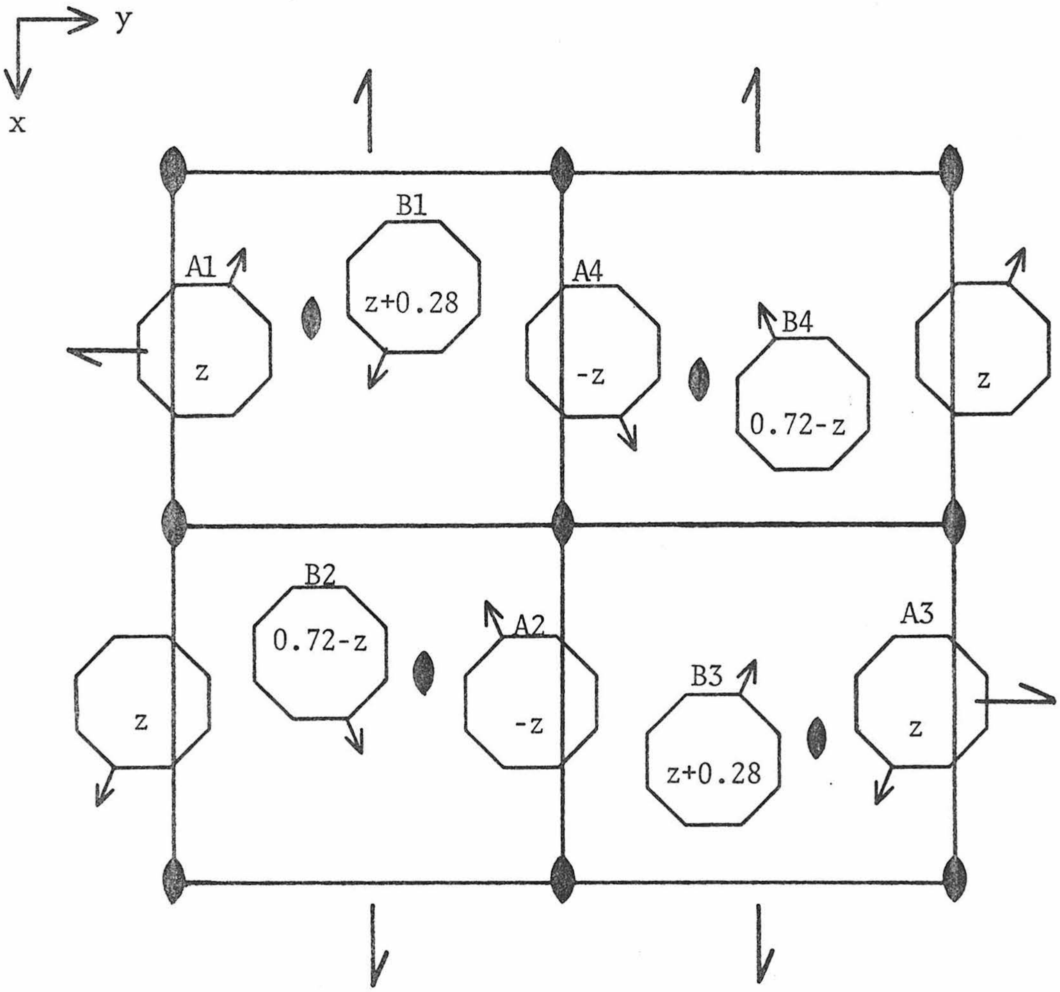


Figure 5-7

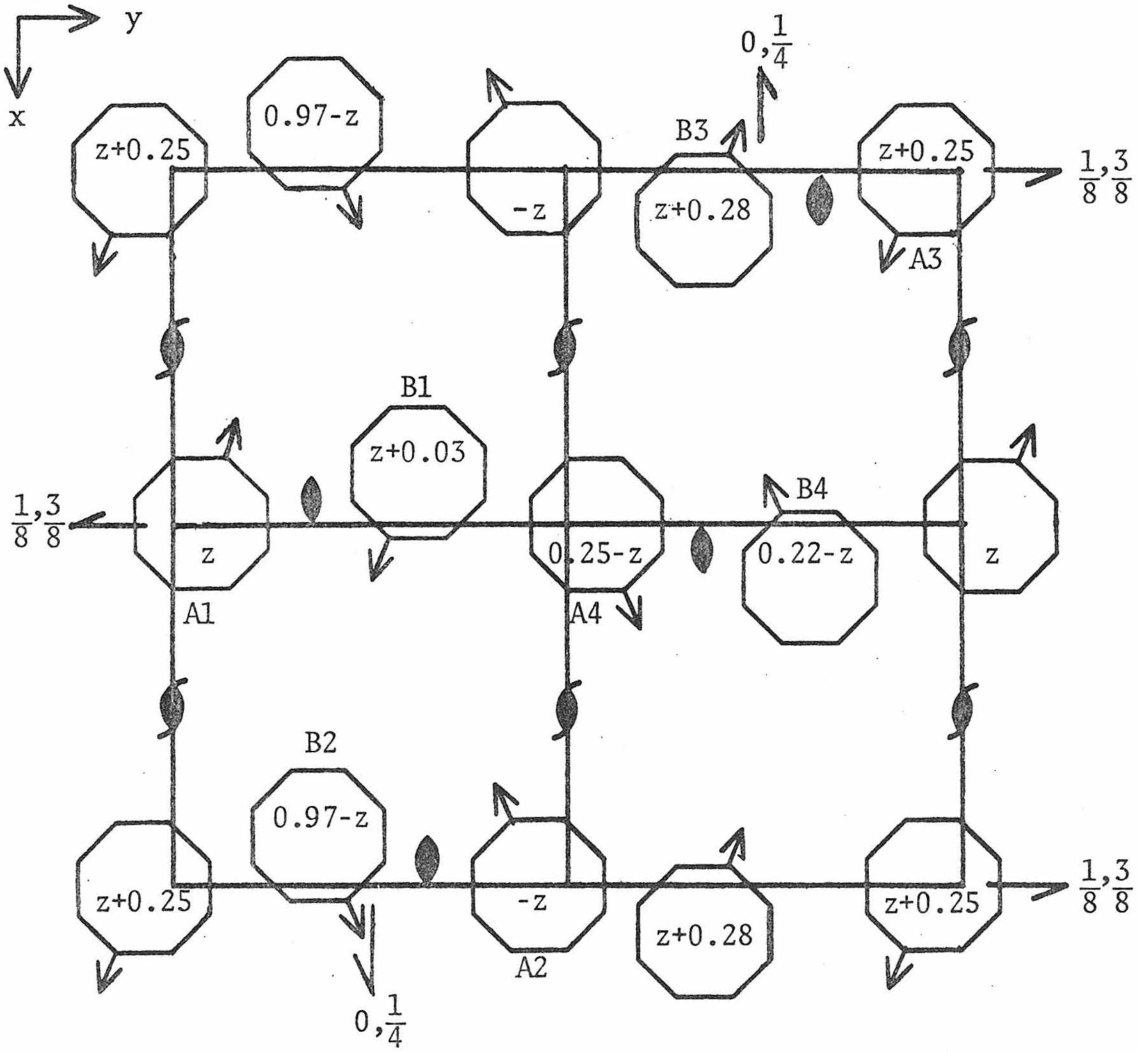

y
x
0.97-z
z+0.25
-z
B3
z+0.28
0,$\frac{1}{4}$
z+0.25
$\frac{1}{8}$,$\frac{3}{8}$
A3
B1
z+0.03
$\frac{1}{8}$,$\frac{3}{8}$
z
A1
0.25-z
A4
B4
0.22-z
z
B2
0.97-z
z+0.25
-z
A2
z+0.28
z+0.25
$\frac{1}{8}$,$\frac{3}{8}$
0,$\frac{1}{4}$

Figure 5-8

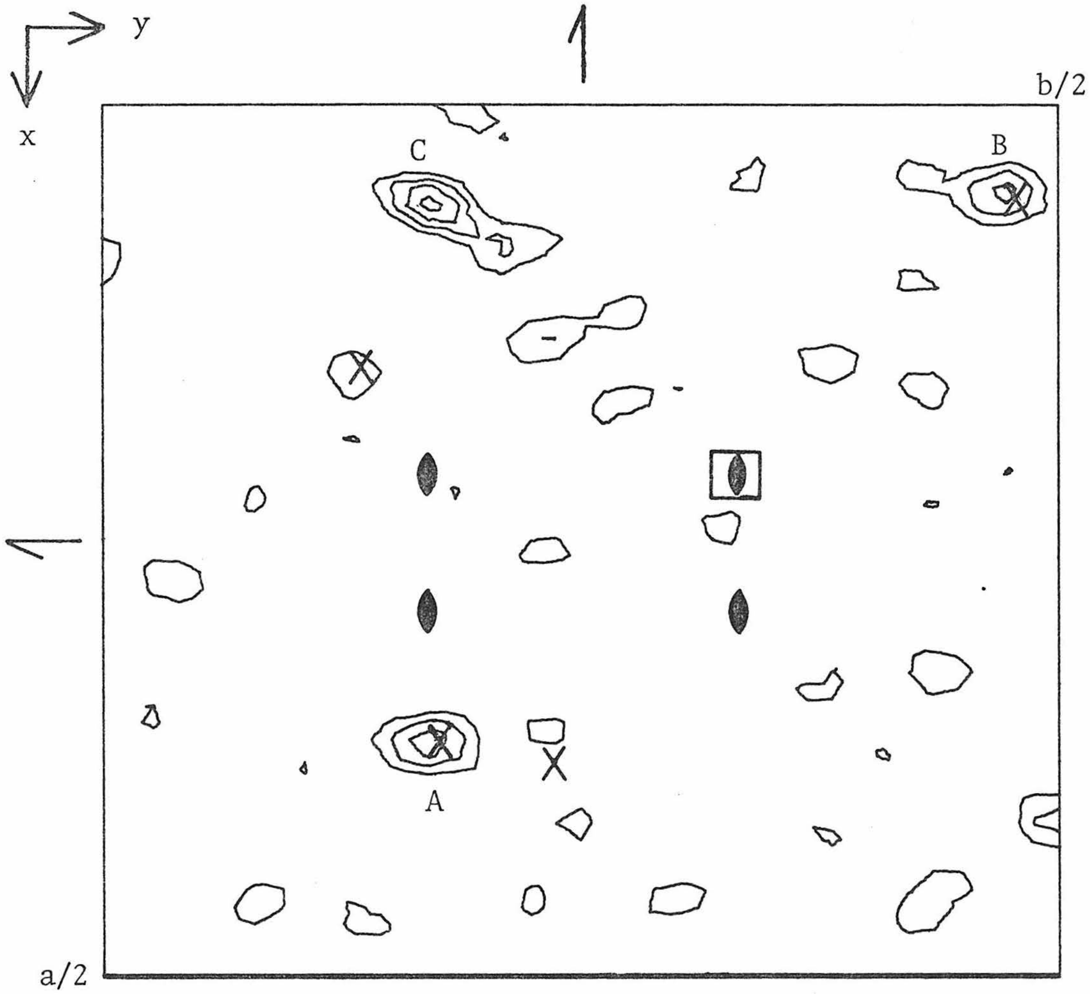


Figure 5-9

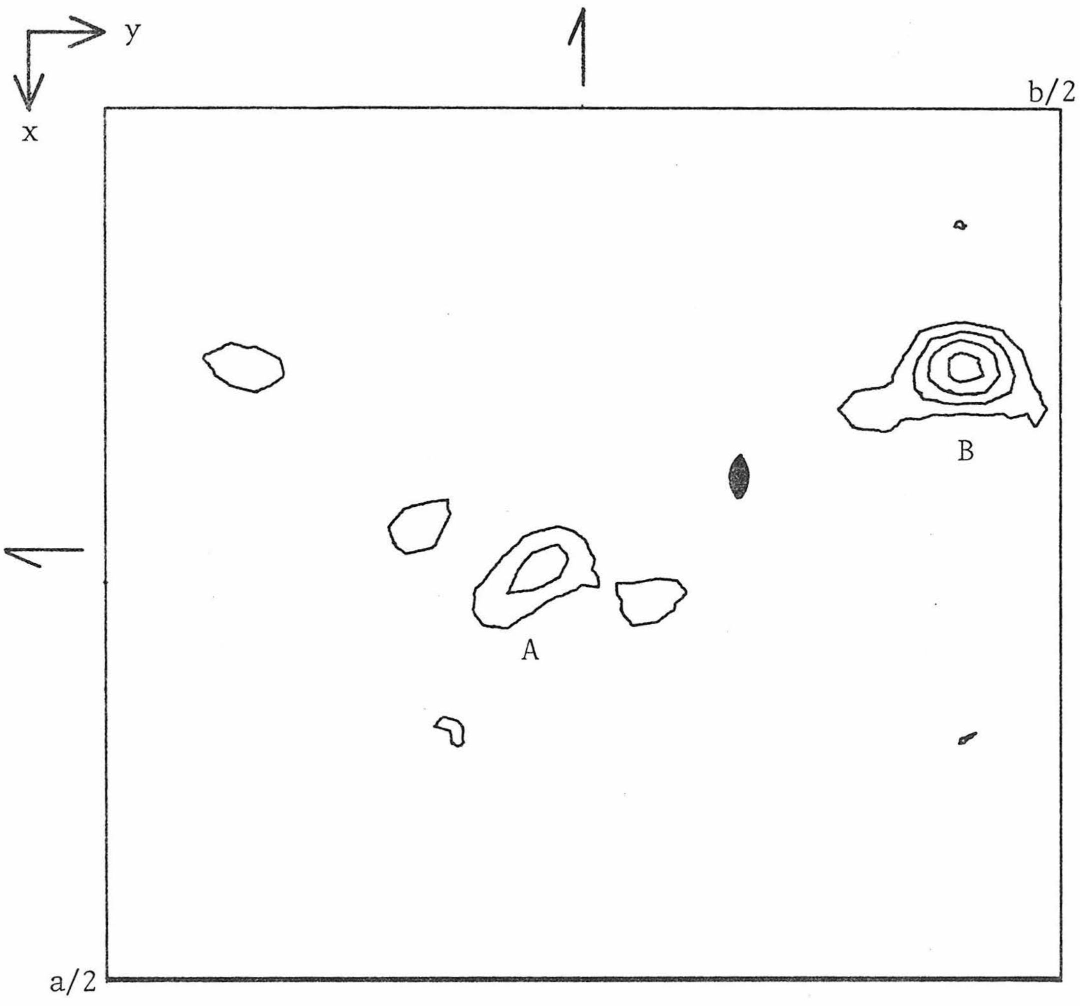


Figure 5-10

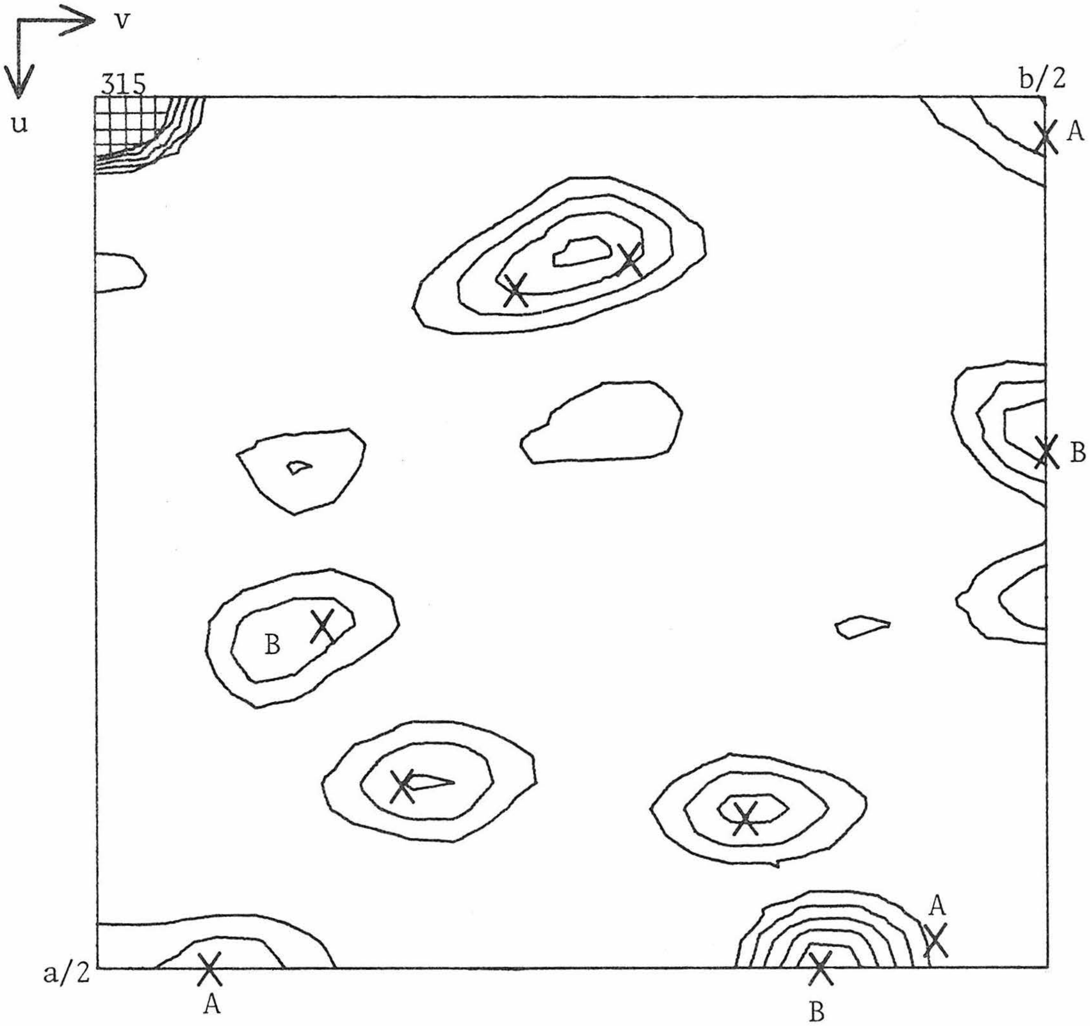


Figure 5-11

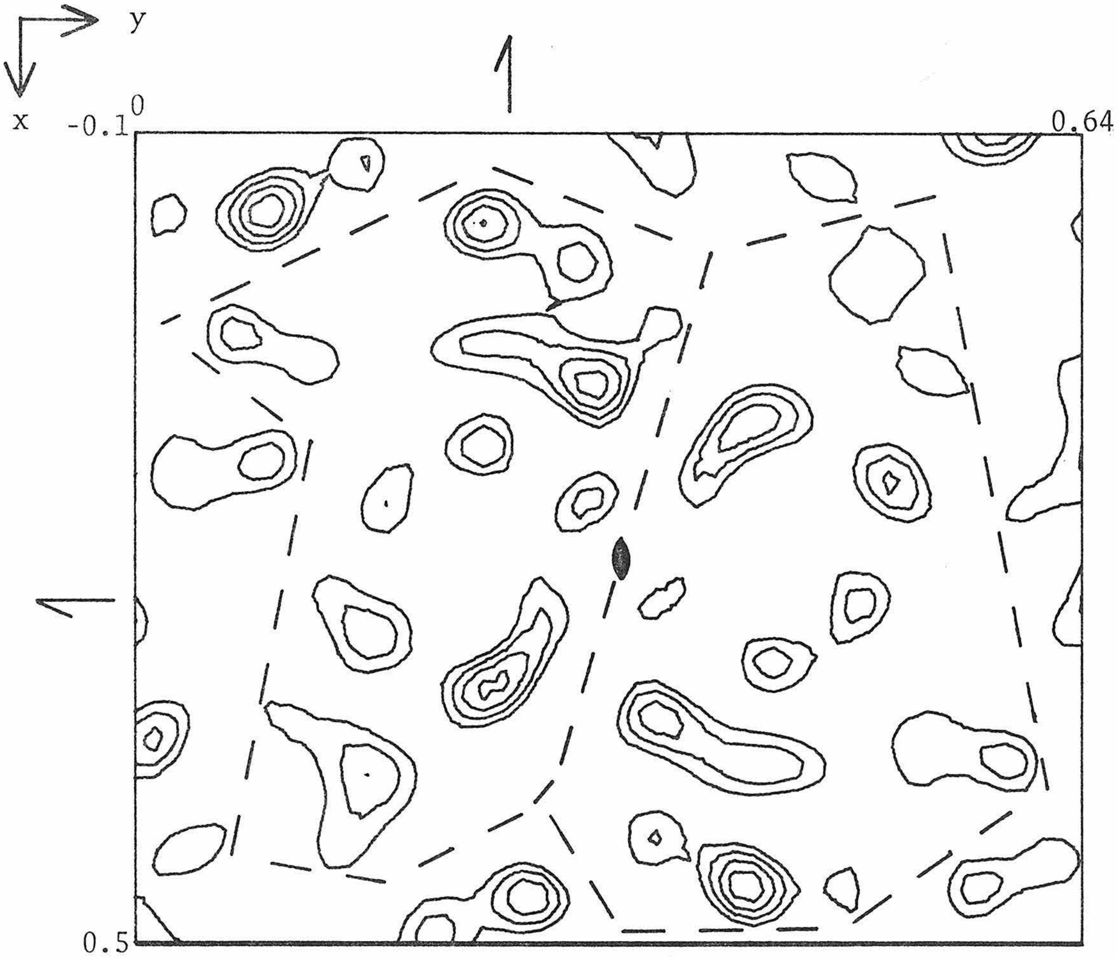


Figure 5-12a

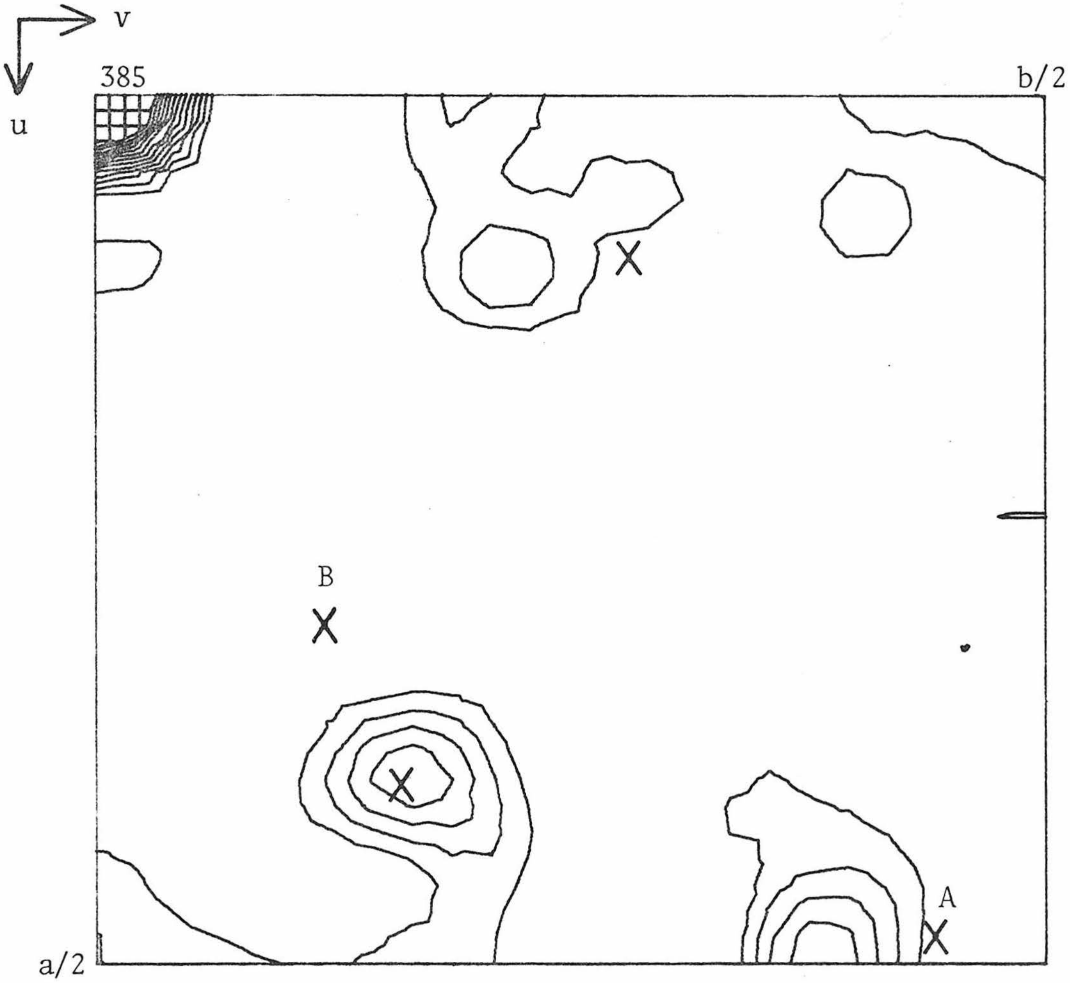


Figure 5-12b

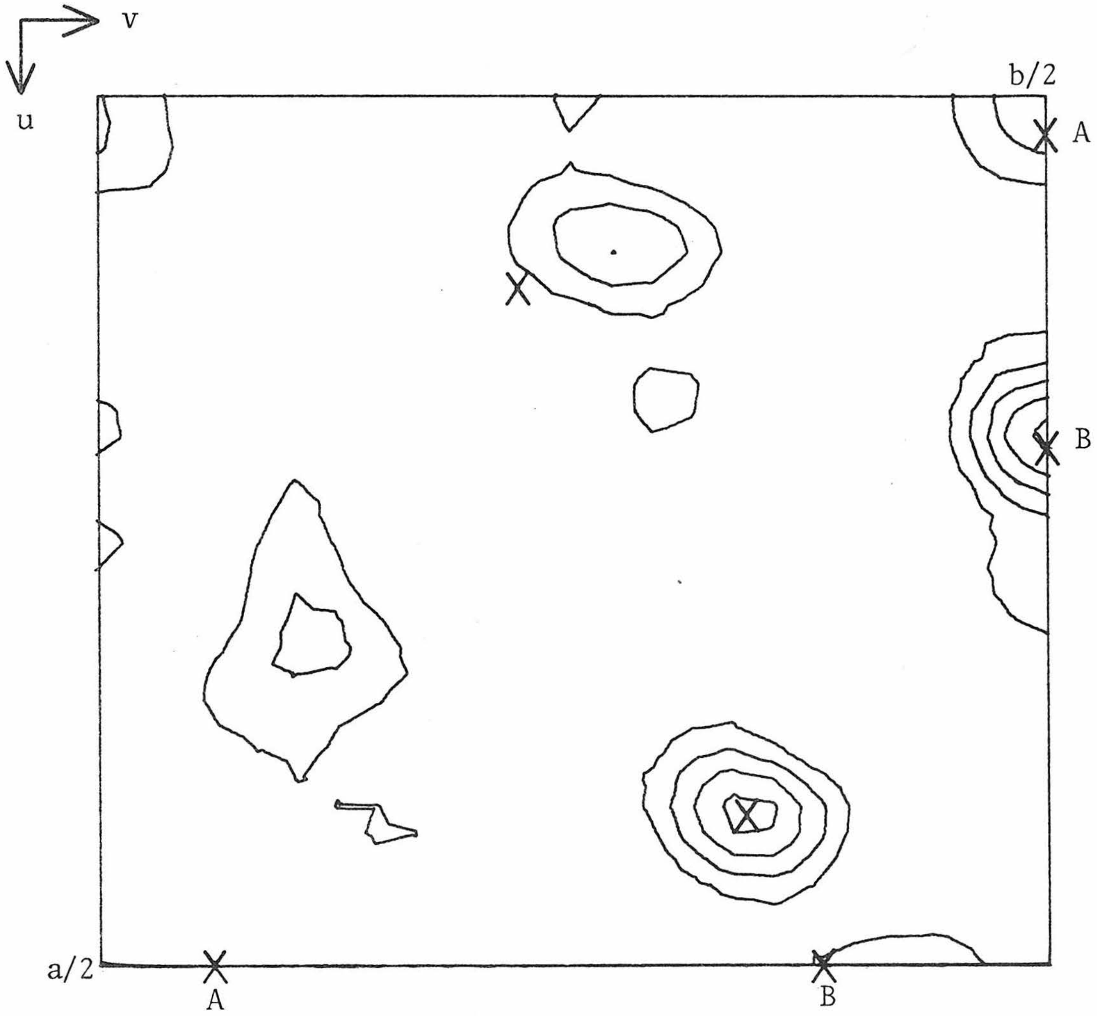


Figure 5-13a

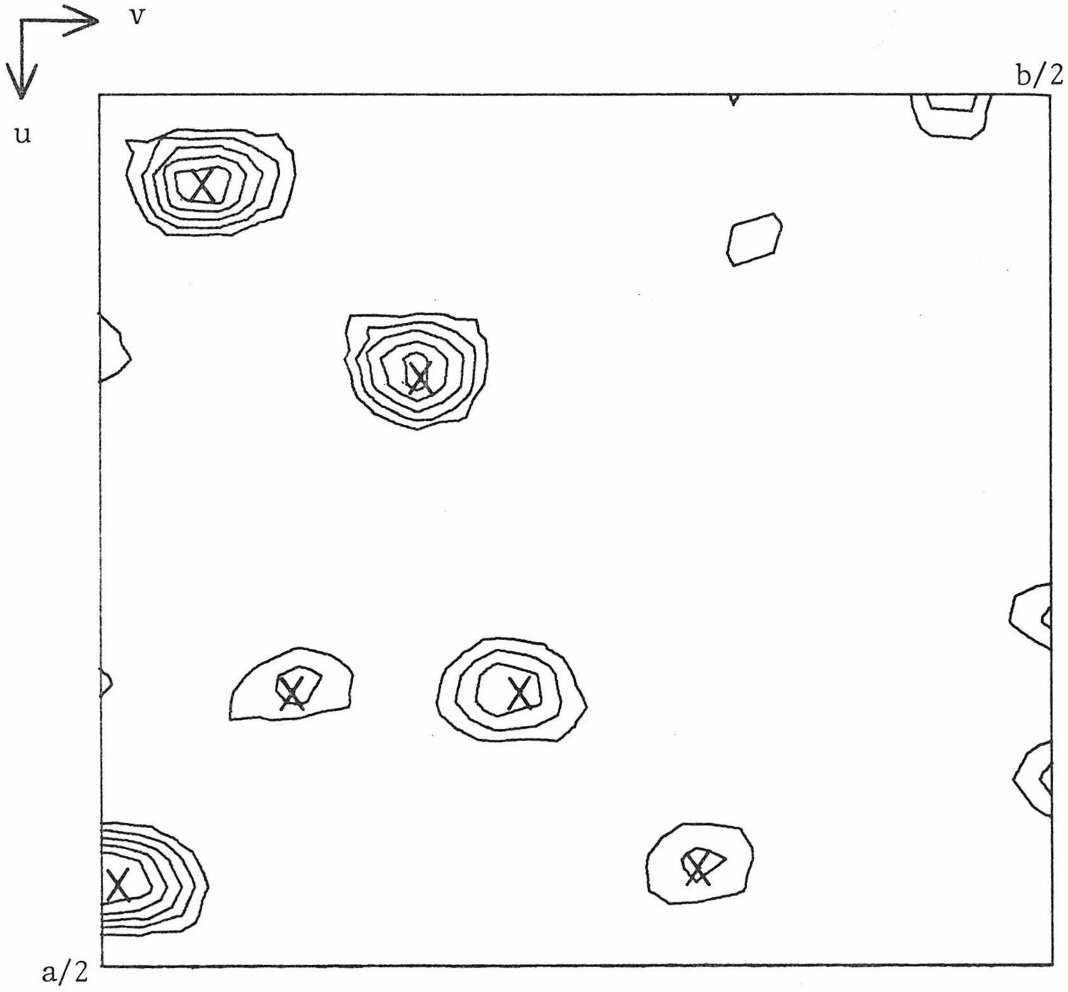


Figure 5-13b

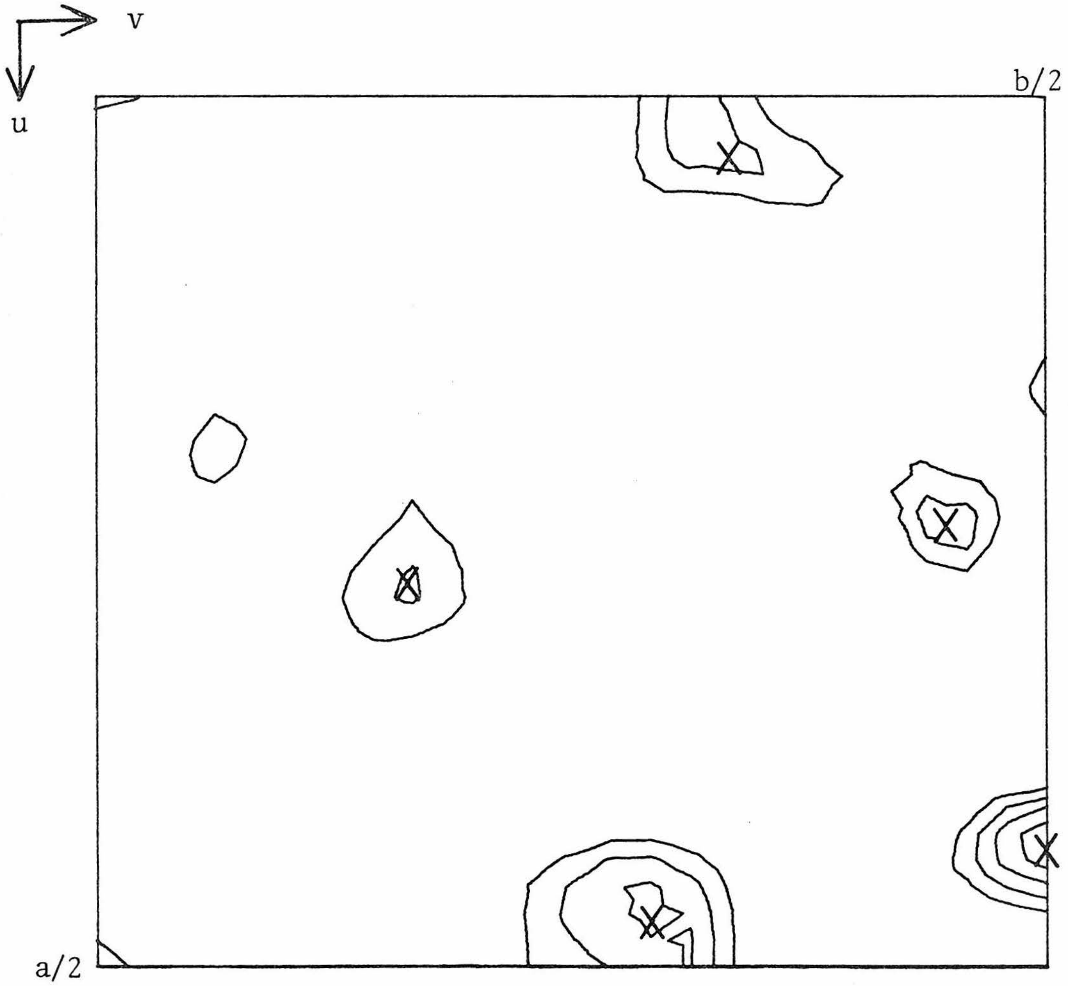


Figure 5-14

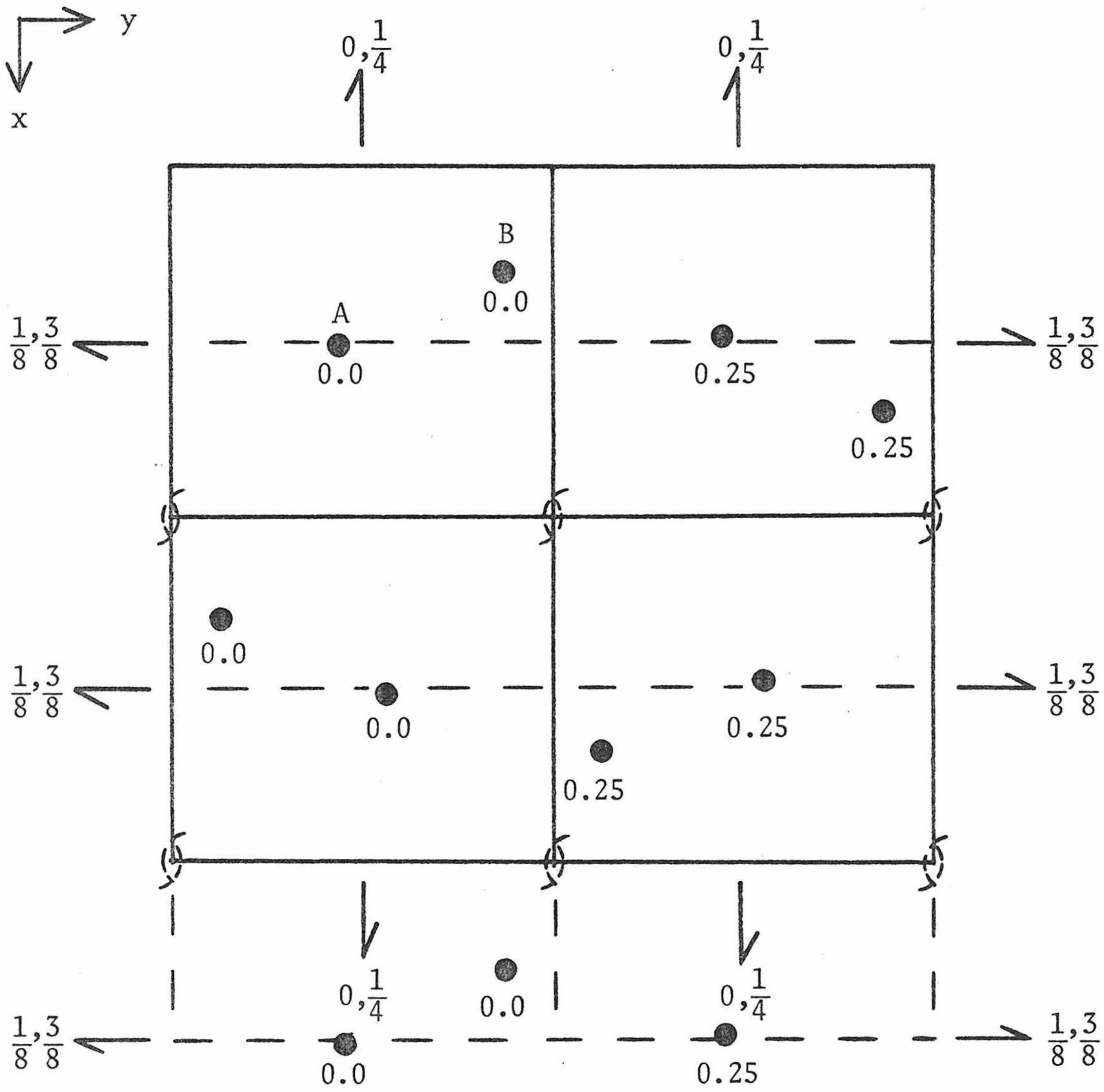

y
x
0,$\frac{1}{4}$
0,$\frac{1}{4}$
B
0.0
A
$\frac{1}{8}$,$\frac{3}{8}$
0.0
0.25
$\frac{1}{8}$,$\frac{3}{8}$
0.25
0.0
$\frac{1}{8}$,$\frac{3}{8}$
0.0
0.25
$\frac{1}{8}$,$\frac{3}{8}$
0.25
0,$\frac{1}{4}$
0.0
0,$\frac{1}{4}$
$\frac{1}{8}$,$\frac{3}{8}$
0.0
0.25
$\frac{1}{8}$,$\frac{3}{8}$

Figure 5-15a

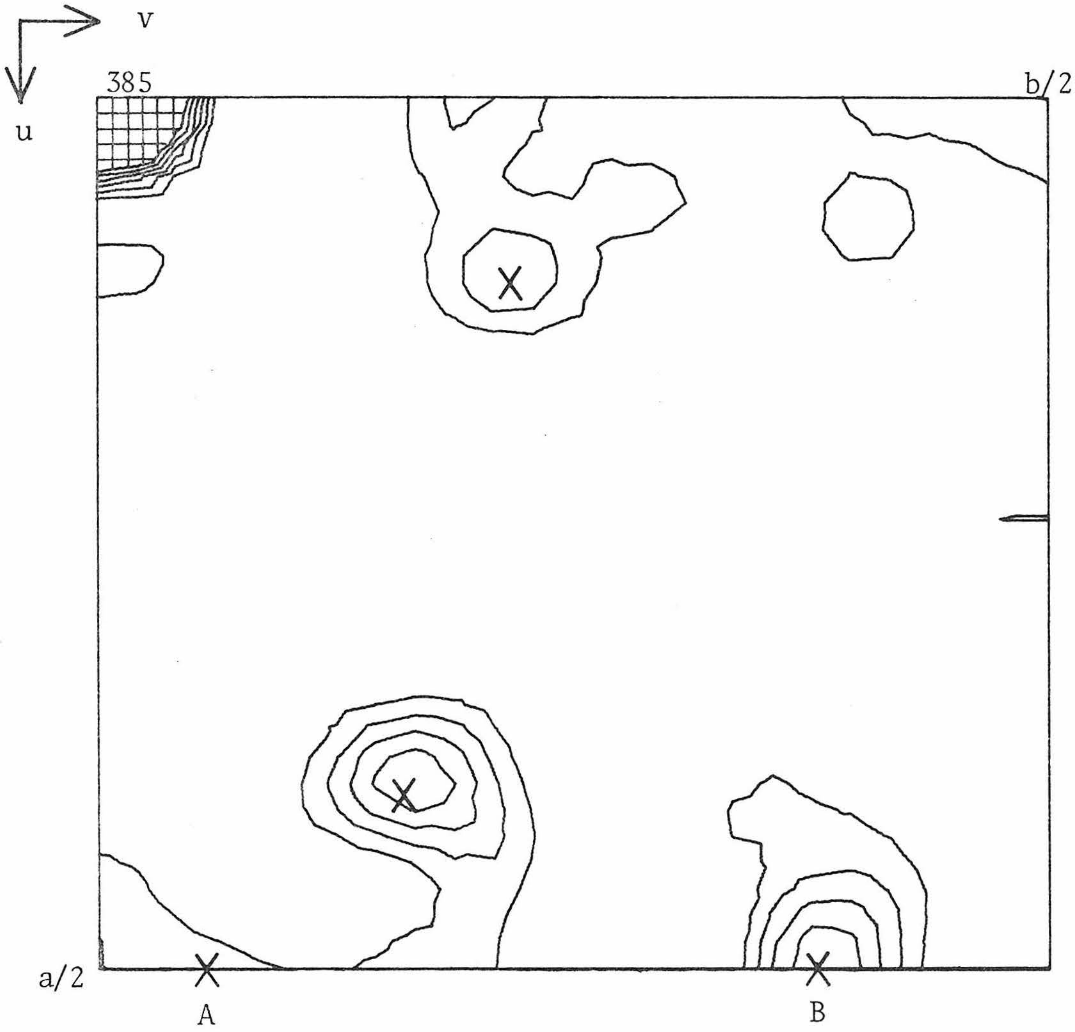


Figure 5-15b

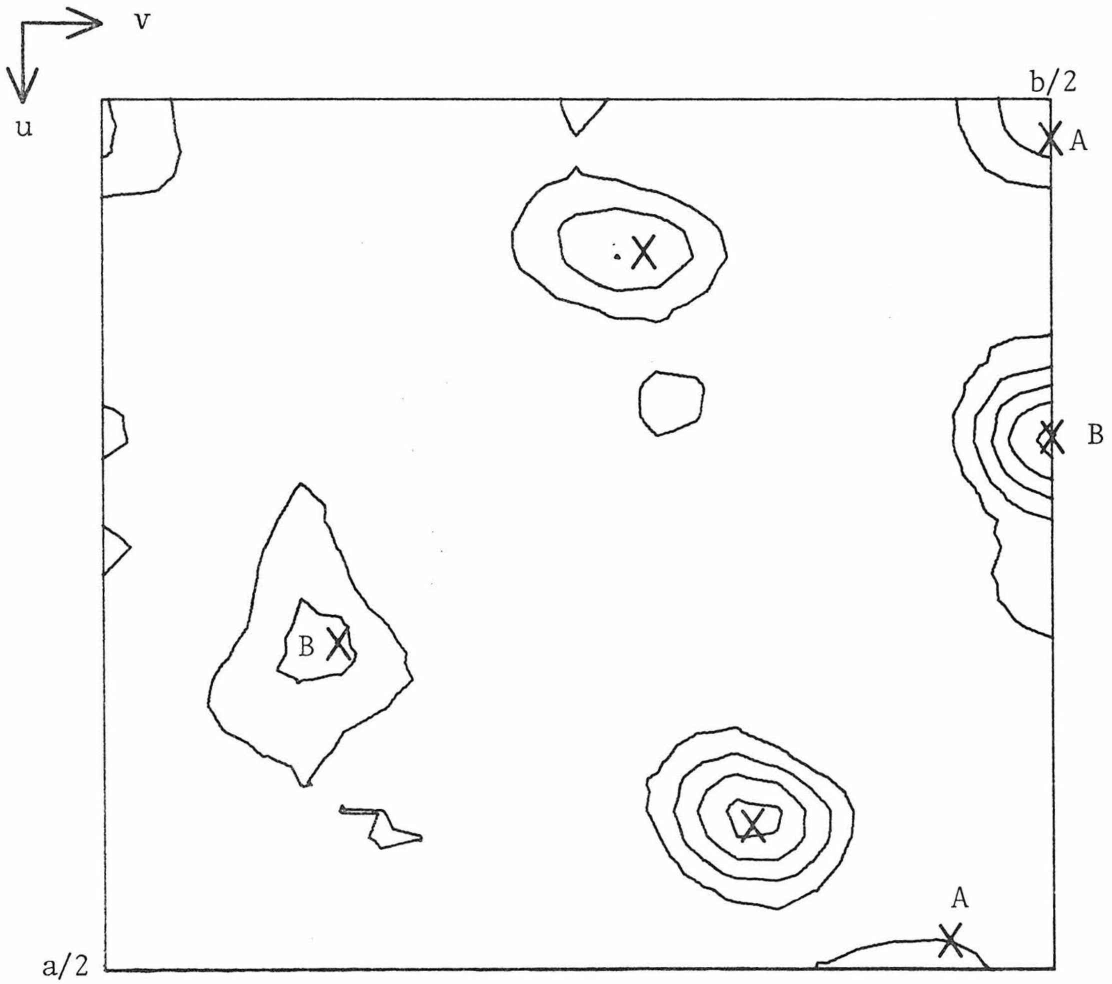


Figure 5-16a

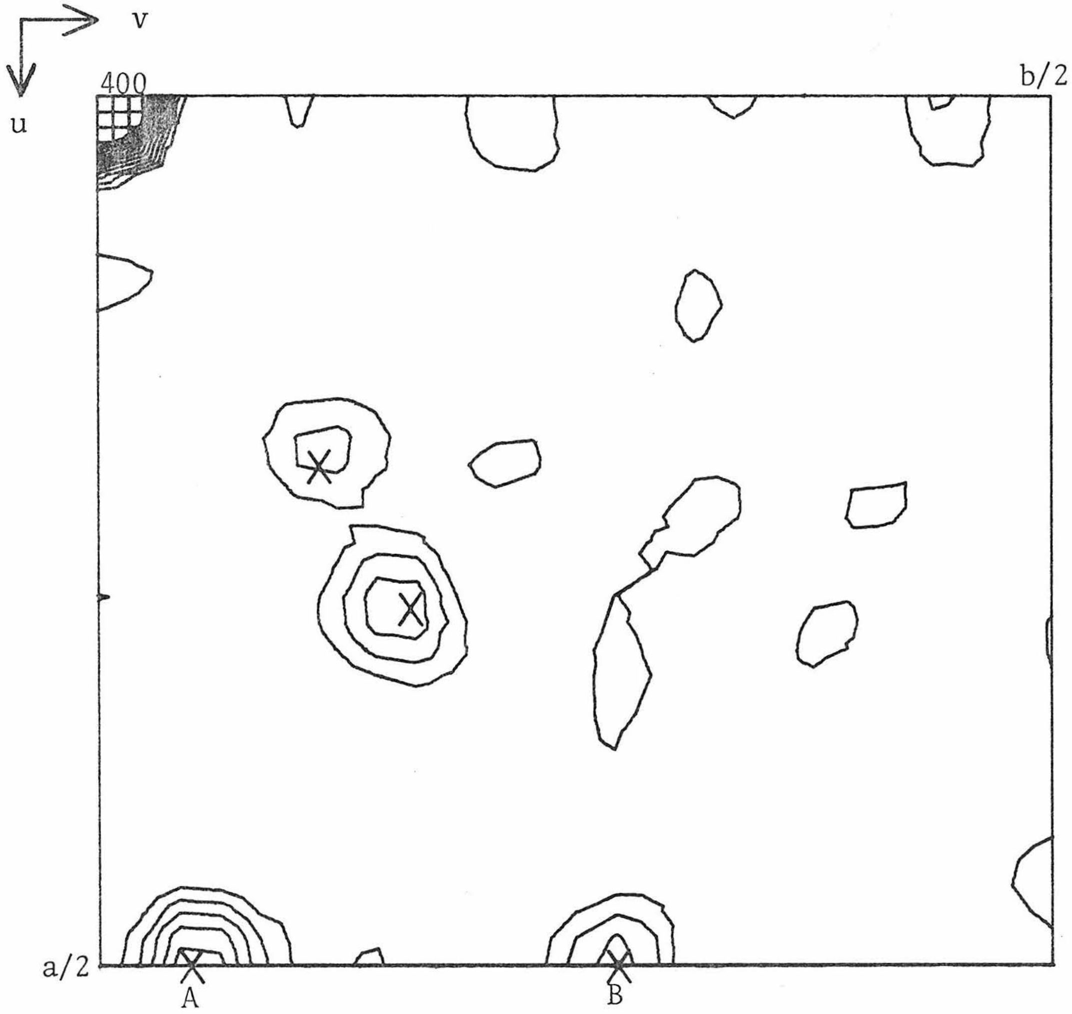


Figure 5-16b

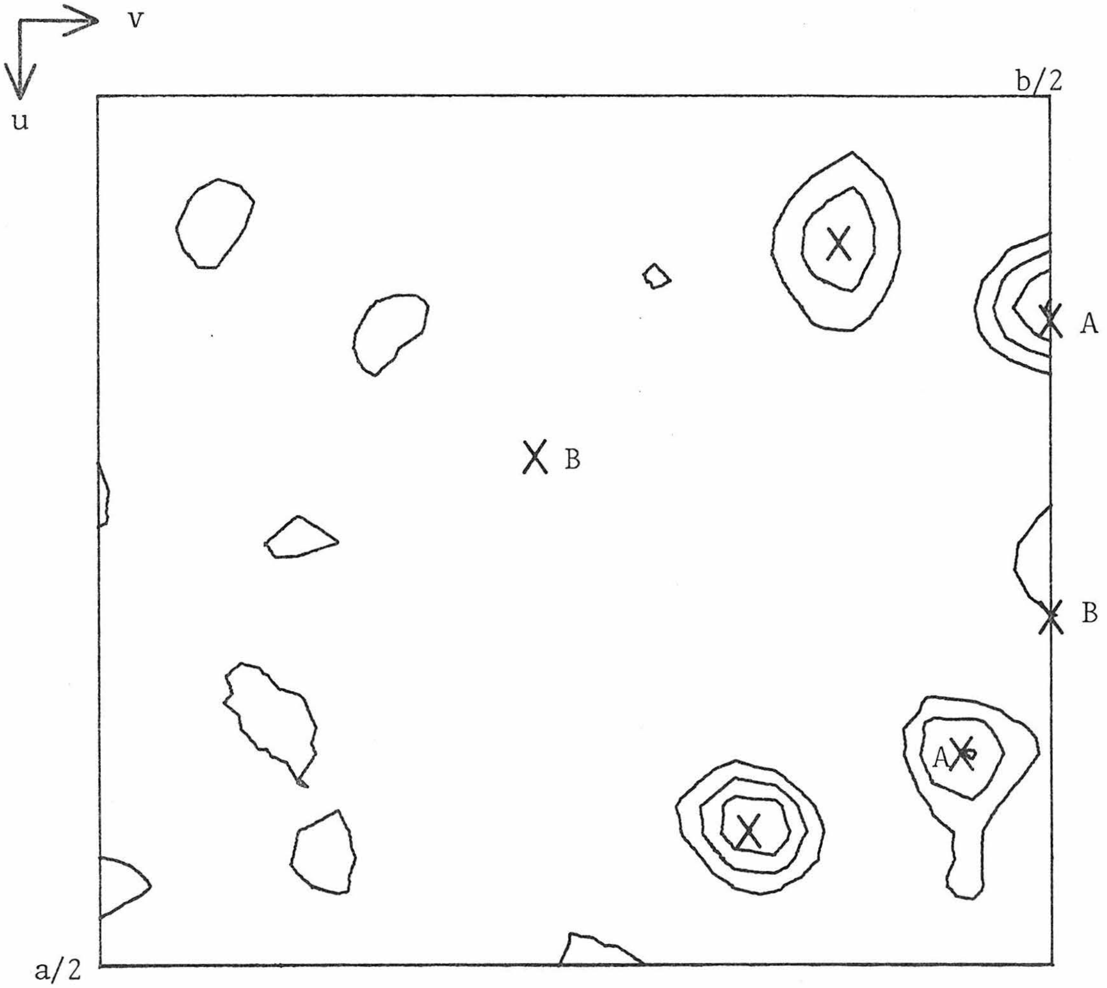


Figure 5-17a

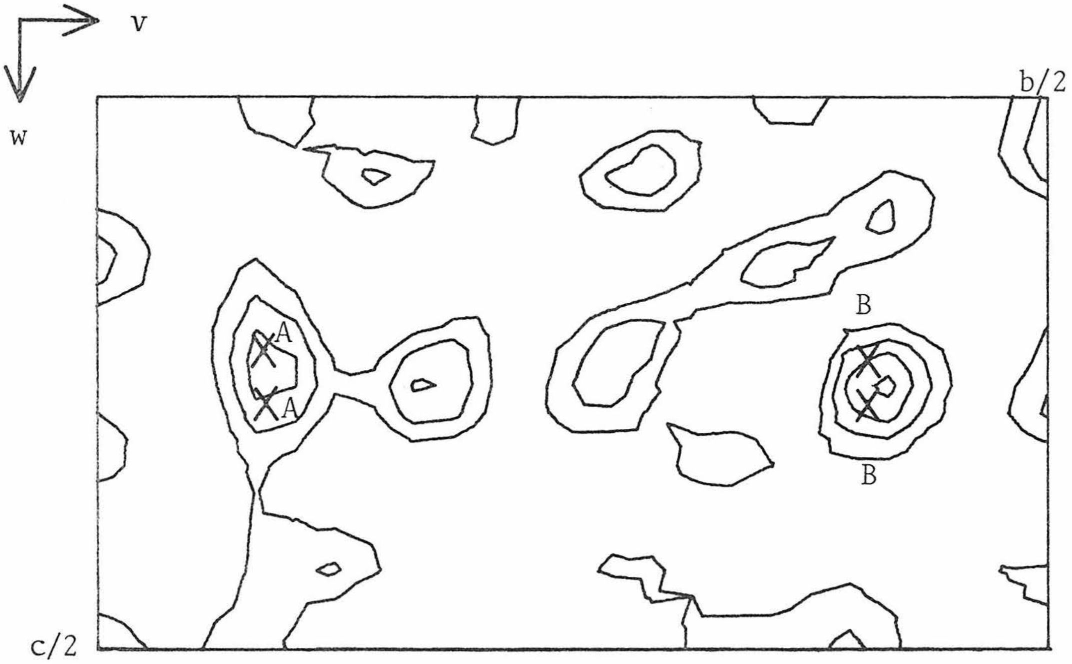


Figure 5-17b

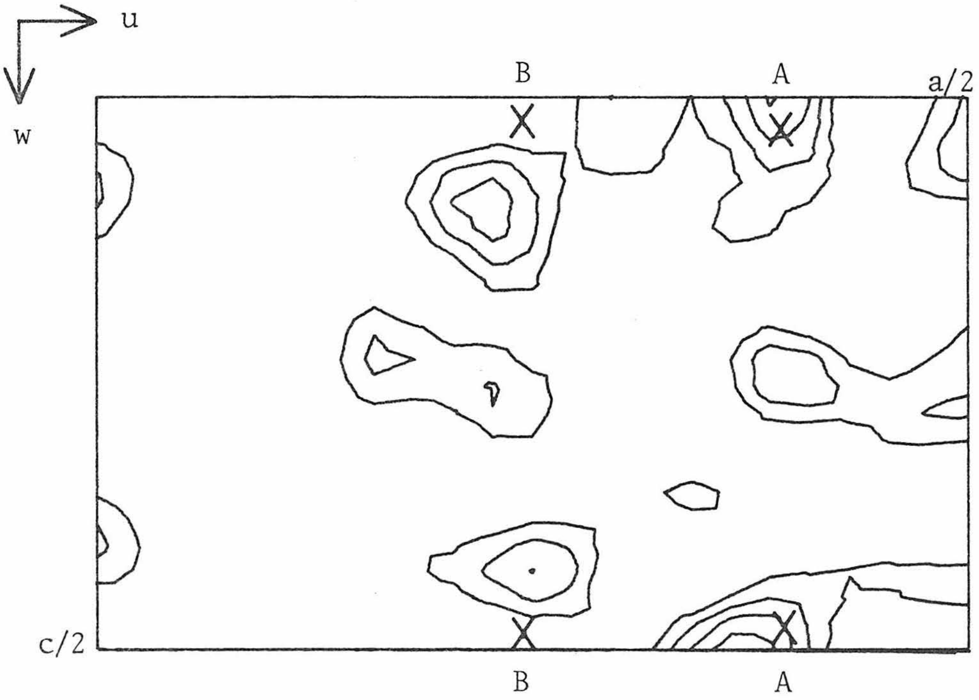


Figure 5-17c

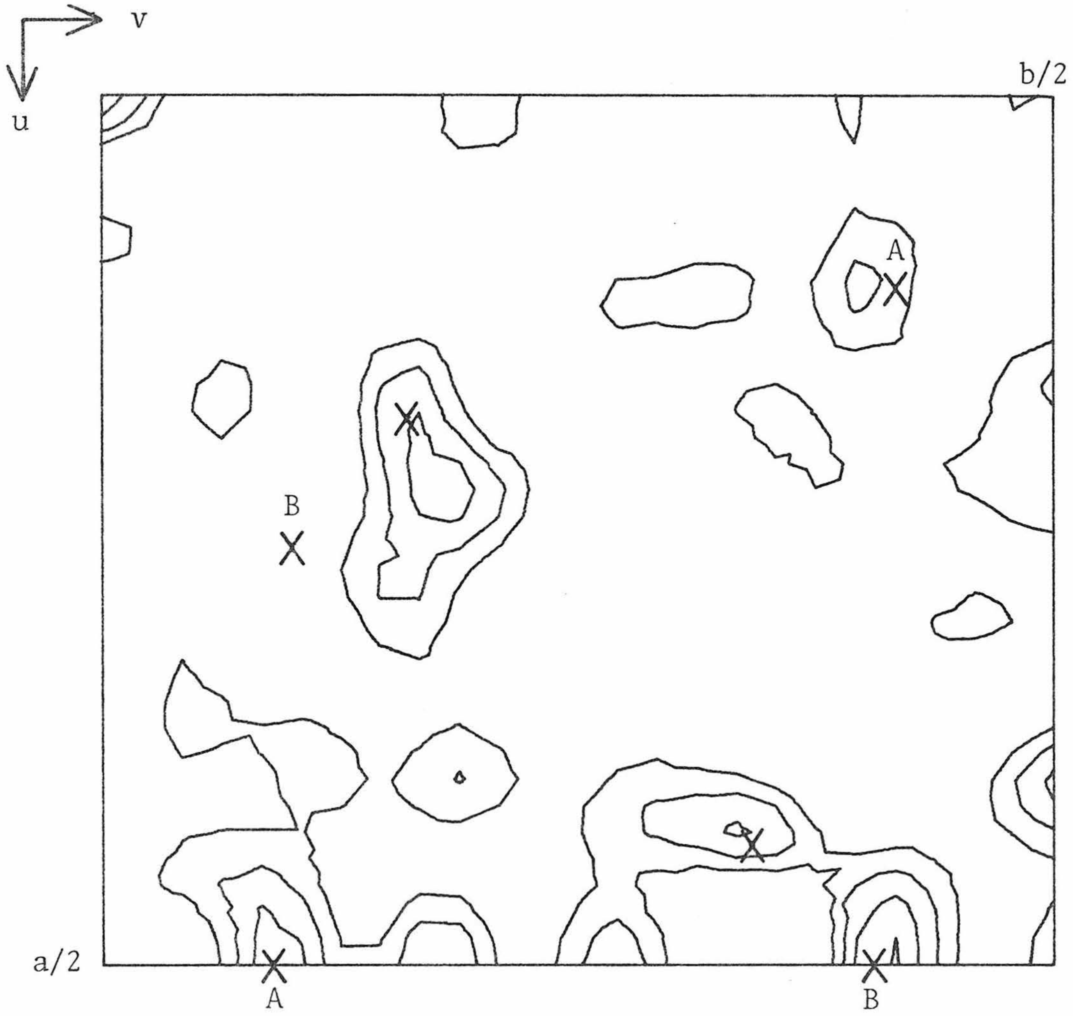
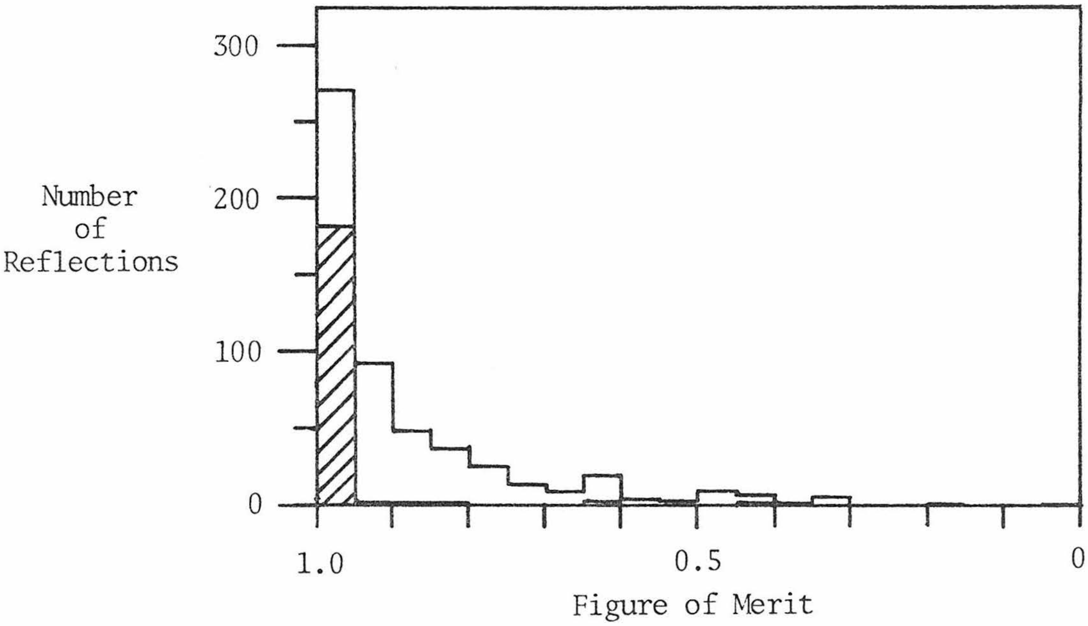


Figure 5-18



CHAPTER VI

The 5 Å Structure of α-Bungarotoxin

"Plop, plop,
Fizz, fizz,
Oh, what a relief it is!"

--Alka-Seltzer commercial

"The truth shall make you free."

--Motto of the California Institute of Technology

Caveat

Interpretation of a 5 Å map of a protein structure is, at best, a tricky business. It is not possible at this resolution to identify most amino acid side chains, and it is often difficult even to trace the route of the polypeptide chain. Therefore, the structural results presented in this chapter will undoubtedly be subject to revision, perhaps major revision, once the high-resolution structure is determined. For this reason, the reader is warned that the interpretations of the α-BuTX map presented here are preliminary in nature, and these results should therefore not be accepted as established fact.

Erabutoxin b structure

Before discussing the structure of α-BuTX, it is valuable to mention briefly the main structural features of erabutoxin b (EbTX). This toxin is the only one whose high-resolution structure has been solved (Low et al., 1976; Tsernoglou and Petsko, 1976).[1] It is one of the short neurotoxins, and contains 62 amino acids linked together by four disulfide bridges.

The structure consists of three broad loops of β-sheet linked together at one end by a knot of four disulfide bonds. The amino-

---

[1] See Chapter I, footnote 5.

terminal and carboxyl-terminal loops are shorter than the central loop, and so, in projection, the structure somewhat resembles a fig leaf. The long, central loop (residues 23-44 in Table 1-2) contains many of the nonstructural conserved residues and is presumably the "active site" of the molecule.

5 Å α-Bungarotoxin Map

A 5 Å electron density map was calculated for native α-BuTX by using the small-subcell data with MIR phases and figures of merit determined from refinement of the eight derivatives (see Chapter V for details). This map was contoured in equal intervals of 1.5 times the estimated error level of the map (Dickerson, Kendrew and Strandberg, 1961), which translated into intervals of about 0.12 $e^-/Å^3$. The zero contour was omitted. A photograph of a portion of this map, viewed down the <u>z</u> axis and centered on the noncrystallographic axis, is shown in Figure 6-1. The molecule on the left in the photograph is referred to below as molecule 1 and the one on the right as molecule 2. Figure 6-2 shows views of a wooden model of the density for molecule 1.

The α-BuTX map was relatively clean and clearly showed molecular boundaries and solvent regions. The molecules in the dimer pair (i.e. those molecules related by the noncrystallographic axis) were in close contact with each other, as well as with small regions of symmetry-related molecules. Despite this, it was possible to assign a boundary both to the dimer pair and to each molecule in this dimer with a high degree of confidence. With the exception of a long "tail," described below, one molecule could be confined within an ellipsoid approximately 40 x 28 x 21 Å, and was therefore comparable in size to EbTX (Low et al., 1976).

The most distinctive feature in the map was a "tail" about 31 Å long extending away from the main body of the toxin. This tail contained several distinct kinks, and was definitely attached to the remainder of the molecule at only one point. These characteristics made it virtually certain that the tail was the carboxyl terminus of the toxin.

The fact that the C-terminal segment of α-BuTX extended well away from the remainder of the toxin instead of interacting with other residues was surprising, and led to speculation about its function. It had been suggested (Lee, 1972) that the long toxins evolved from the short ones. If this were indeed the case, it seemed probable that the extended tail performed some useful function. Since it obviously did not contribute to the maintenance of the three-dimensional structure of the toxin, it appeared likely that it conferred some advantage to the toxin in its encounter with the acetylcholine receptor. It could have been designed to form secondary interactions (ion-pair bonds, hydrogen bonds) which would cause the toxin to bind more tightly to the receptor. Such an arrangement could explain the apparent irreversibility of α-BuTX binding, as compared to cobrotoxin (Lee, 1972), and could also explain the difference between the long and short toxins in the kinetics of their interaction with the receptor (Chicheportiche et al., 1975).

Despite this knowledge of the position of the α-BuTX carboxyl terminus, it was possible to trace the polypeptide chain backwards for only a short distance. Once the chain rejoined the main region of positive density, the large number of disulfide bonds (five) made it difficult to follow the main chain. At this resolution, disulfide bonds did not appear to have a significantly higher electron density than the polypeptide chain and so it was not possible to determine on this basis which one was which. Therefore, despite the fact that the polypeptide chain was readily traceable in several regions of the map, it was not possible to trace unambiguously the complete chain.

Structure Prediction

Since the structure of EbTX consists largely of β sheet with no α helix, the α-BuTX map was examined for these features. As with EbTX, no α helix was found, although some regions which possibly contained β sheet were identified. To further investigate the possibility of an α helix in α-BuTX, the structure prediction method of Chou and Fasman (1974) was tried. Only one small potentially α-helical region, between residues 37 and 42, was predicted, and even this was marginal. It is unlikely that such a small section of α helix would be identifiable at 5 Å in any case.

The structure prediction technique was also used to predict the positions of β sheet and β turns in the α-BuTX sequence. Three region regions of β sheet (residues 1-8, 22-32 and 57-62) and five β turns (residues 9-12, 17-20, 33-36, 48-51 and 63-66) were predicted.[2] Although it was not possible to identify these residues in the α-BuTX map, the predicted positions for the β turns were close to where they occur in the EbTX sequence[3] (Low _et_ _al_., 1976). The β sheet predictions did not agree quite so well with the EbTX structure, but they were in the right general areas. A determination of the accuracy of these predictions will have to await the high-resolution α-BuTX structure.

Determination of the Noncrystallographic-Axis Position

To determine the position in the 5 Å map of the noncrystallographic axis, as well as to determine the precise nature of its symmetry, a program was written for an IBM 370 computer to rotate around this axis a portion of the map containing a dimer pair (molecules 1 and 2 in Figure 6-1), and then to translate this piece of the

---

[2] A recent paper by Hseu _et_ _al_. (1977) also predicts the α-BuTX structure by the Chou and Fasman method. Their predictions of β turns are consistent with those given above, but some of their predicted regions of β sheet are different.

[3] With the exception of the final β turn, which would be at the end of the EbTX sequence.

map with respect to the unrotated map. In order to locate the axis in the $\underline{x}$-$\underline{y}$ plane, the map was first rotated 180°, then translated in a grid-search pattern over the unrotated map. For each new position of the rotated map, the sum of the products $\rho_1\rho_2$, where ρ_1 is the electron density at each grid point in the rotated map and ρ_2 is the density at the corresponding point in the unrotated map, was calculated. When necessary, eight point linear interpolation was used to calculate densities between grid points in the unrotated map.

This sum was quite sensitive to shifts as small as 0.001 of a cell in the noncrystallographic axis position, and appeared to be monotonic around the maximum value. The maximum occurred with the noncrystallographic axis at $\underline{x} = 0.540$, $\underline{y} = 0.670$.[4] This position is identical to the one found when just the [001] projection of the native structure was rotated in this same manner.

Once the position of the noncrystallographic axis was fixed, rotation angles other than 180° were tried, as were translations of one molecule in the dimer pair with respect to each other. All of these changes resulted in a poorer match between the two molecules. No attempt was made to tilt the rotation axis, but any significant tilt should have been apparent in the electron density map.

---

[4] This position for the noncrystallographic axis was a consequence of the portion of the map that was chosen for the rotation. Operation by the cell symmetry $(-\underline{x}, 1/2 + \underline{y})$ returns this axis to the familiar location of $\underline{x} = 0.46$, $\underline{y} = 0.17$.

Once the position of the noncrystallographic axis was determined, the rotated and unrotated maps were averaged together. A view of this map down the noncrystallographic axis is shown in Figure 6-3. This map shows the same portion of the unit cell as is shown in Figure 6-1, although the contour levels are somewhat different.

The averaging process largely cancelled out the density for the other molecules in the cell which were related to the dimer pair by crystallographic symmetry, and it also reduced the number of peaks in the solvent region. Both of these changes helped to clarify the boundary of the dimer pair in those regions of close contact with other molecules.

The averaged molecules looked basically the same as they had originally, except that in certain regions the main chain was easier to follow and in other regions it was less clear. In general, those regions which were well connected and clear in the original map were also clear in the averaged map, while those regions which were poorly connected or weak became worse after averaging. Any real differences between the two molecules in the dimer pair were, of course, obliterated.

Comparison of Erabutoxin to α-Bungarotoxin

In order to compare the EbTX and α-BuTX structures, as well as to help interpret the low-resolution α-BuTX map, a preliminary set of

EbTX coordinates--graciously provided by Dr. Demetrius Tsernoglou and Dr. Gregory Petsko of Wayne State University--was rotated through all possible orientations until a best fit of the EbTX structure to the α-BuTX electron density was found. The rotation was performed by first calculating a "center-of-mass" for the EbTX main chain and sulfur atoms and a "center-of-density" for molecule 1 in the α-BuTX map. These centers were then made coincident, and the EbTX structure was rotated about this center through all unique values of the three Eulerian angles.[5] Initial rotations were in steps of 20°. The "best" fit was defined as that orientation which maximized the sum of the electron density measured at the positions of the EbTX main-chain atoms and disulfide bonds. Since the absolute configuration for the α-BuTX structure was not known, both enantiomorphs were tried.

The result of this search procedure was one orientation which was clearly better than the others. A fine angular search was performed about this orientation, and then the EbTX center-of-mass was shifted parallel to the three orthogonal axes until an optimum fit was found. A final, fine angular search was then performed and resulted in the EbTX orientation shown in Figure 6-4. It was from this orientation of EbTX that a right-handed coordinate system was chosen as the one giving the correct α-BuTX enantiomorph.

An encouraging result of this fitting process was that the

---

[5] See Rossmann and Blow (1962) for a description of the Eulerian system.

The residues at the tip of this loop (Thr 38 to Ser 29) were completely out of density. The map did contain positive density near these residues, however, and it is possible that this density contained the tip of the α-BuTX active loop. Indeed, this section of positive density had an offshoot which may have been the ring of Phe 32.

The difference in this region between the two structures was not surprising, since the α-BuTX sequence has four amino acid residues and an extra disulfide bond inserted at this point, relative to the EbTX sequence. What was unusual, however, was the poor quality of the map in this area. In EbTX, this loop consists of a single layer of β sheet, unsupported by other parts of the molecule. It is likely that such a structure is fairly flexible, so it is possible that the poor quality of the map in this region was due to the presence in the crystals of several conformations of this loop. It has been reported (D. Tsernoglou, personal communication) that this loop was also the worst part of the EbTX map.

Between Tyr 24 and the amino terminus, EbTX forms the amino-terminal loop. In the α-BuTX map, the EbTX sequence rejoined the positive α-BuTX density around Ser 22 and fit it fairly well--although not as well as the fit to the C-terminal loop--until Gln 12. At this point, the EbTX sequence looped down close to the body of the toxin, whereas the α-BuTX density extended out into the solvent before turning back to form a loop. Once the α-BuTX density rejoined the main body of the toxin, it was not possible to follow the main chain any farther. A region of positive density between the C-terminal loop and the active loop may have been the amino terminus, although it did not agree with the position of the EbTX amino terminus.

Thus, at low resolution the two toxin structures appeared to have significant differences between them. The differences between the active loops were expected because of the major sequence differences at this point, but it was somewhat surprising that the amino-terminal loops were different. It is possible that flexibility in the loops could account for some of these differences. In any case, it will be necessary to get a high-resolution α-BuTX structure before these differences can accurately be determined.

Heavy-Atom Binding Sites

With one exception, all the heavy-atom binding sites were on the surfaces of the toxin molecules or in the solvent regions between them. The exception, iridium site 7 (Ir7),[7] was in the region of the molecule where the three loops were linked together by the disulfide knot. This iridium site was present in only one of the two molecules in the dimer pair; and this fact, coupled with its low refined occupancy (about 2.3 electrons), suggested that it was probably an artifact.

Most of the remaining sites were arranged in several clusters distributed around the surfaces of the two molecules. Two of these

---

[7] Numbering of the sites is the same as in Appendix B.

clusters, one containing sites Ir3, Tl1 and uranyl 2 and the other containing sites PtS 1 and Hg 1 (as well as their noncrystallographically related mates shown in Table 6-1), were close to one another and occupied a gap approximately 6 Å wide between the C-terminal loop of one toxin molecule and the C-terminal loop of the molecule one unit cell below it in _z_. Based on the fit of the EbTX structure to the α-BuTX density, it appeared that the first cluster (Ir3, Tl1, uranyl 2) was bound to Lys 51 or Lys 52, while the second group (PtS 1, Hg 1) was bound to Glu 56.

It is interesting that crystals left soaking in the three heavy-atom compounds in group one were stable for long periods of time, while crystals soaked in the heavy-atom compounds in group two became quite unstable and cracked after relatively short soaking times. Since both groups bound in equally open regions of the crystals, steric factors could not explain this difference.

It was mentioned in Chapter III that the large cell to small cell conversion appeared to be controlled by the ionization of one or more carboxyl groups. If the heavy-atom compounds in the second group were bound to Glu 56 or close to it, it is possible that they disturbed the pK_a of this group sufficiently to initiate conversion to the small cell, thereby cracking the crystals. Since Glu 56 of one molecule is just above Lys 51 or Lys 52 of the molecule one unit cell below it in _z_, it is not unreasonable to expect that ionization of Glu 56 would tend to cause the molecules to pull towards each other at this point.

TABLE 6-1

Derivative sites related by noncrystallographic symmetry

Derivative	Site[1] on Molecule 1	Site on Molecule 2
$Pt(SO_4)_2$	1	2
$Pt(CN)_4^{-2}$	1	2
	-	3
	-	4
OsO_4^{-2}	2	1
	-	3
$UO_2F_5^{-3}$	2	1
	3	-
$TlCl_6^{-3}$	1	4
	3	2
	-	5
	-	6
	7	-
	-	8
	-	9
	10	-
$Ir(CN)_6^{-3}$	1	5
	3	2
	-	4
	7	-
	8	6[2]
$HgBr_2$	1	2
	4	3
	-	5

Table 6-1 (continued)

[1] Site numbers are the same as in Appendix B.

[2] Sites Ir8 and Ir6 were in interstitial regions partially composed of molecules not related by the noncrystallographic symmetry. Their approximate adherence to the noncrystallographic symmetry may have been coincidental.

The only other derivative for which a reasonable guess could be made as to the identity of its major binding sites was Os. The two major Os sites, Os1 and Os2, appeared to be close to the imidazole ring of His 67 on toxin molecules 2 and 1, respectively.

As was mentioned in Chapter V (Chapter V, footnote 3), the binding of osmium to these sites appeared to alter the protein structure, as indicated by the presence of two negative peaks in the Os difference Fourier map. The first negative peak was centered at approximately $\underline{x} = 0.52$, $\underline{y} = 0.54$, $\underline{z} = 0.35$, which is about 10 Å away (mainly in the $\underline{x}$ direction) from site Os2. The second negative peak was at approximately $\underline{x} = 0.73$, $\underline{y} = 0.84$, $\underline{z} = 0.70$, which is about 11 Å away (mainly in the $\underline{z}$ direction) from site Os1. In the α-BuTX map, the first peak fell in the center of the density for the C-terminal tail of molecule 1, at a position approximately one residue removed from His 67. The second negative peak fell near part of the amino-terminal loop of molecule 2.

It was not clear why the effects of osmium binding should be different for the two molecules, nor was it clear why these relatively sharp negative peaks occurred instead of the broad negative regions which would be expected if the toxin structure was altered. Despite these unanswered questions, it seemed unlikely that these peaks were simply artifacts, since they appeared in both the difference map and the double difference (or lack-of-closure) map[8] for this derivative. While this result is not conclusive, it suggests that the negative peaks resulted from real features in the Os crystals.

The remaining heavy-atom binding sites were not identifiable as specific residues, and many of these heavy atoms did not obey the noncrystallographic symmetry (Table 6-1). Of the thirteen sites not obeying the noncrystallographic symmetry, one (Ir7) has been discussed, three (PtC3, PtC4, Ir4) were in interstitial sites which were partially composed of molecules not related by the noncrystallographic twofold axis and therefore did not have to obey the noncrystallographic symmetry, and three others (T15, T18, Ir4) may have been in such sites. T16 was also in an interstitial site--this one composed of the C-terminal tail of one molecule and the body of the molecule one unit cell below it in z. Slight differences in the positions of the tails in the two molecules in the dimer pair could explain why this site did not obey the noncrystallographic symmetry.

It is not clear why the other five sites (Os3, uranyl3, T17, T19, Hg5) did not obey the noncrystallographic symmetry. All were relatively low-occupancy sites, so it is possible that the matching sites were not above the noise level in the difference map. Slight differences between the two molecules in the dimer may also have made

---

[8] The negative peak near site Os1 was only -71 (arbitrary scale) in the double difference map, as opposed to -132 in the difference map, but the one near site Os2 was only slightly reduced in the double difference map (-109 versus -120). The average negative noise level in the double difference map was around -40.

a site on one molecule less favorable on the other one. The high-resolution structure of α-BuTX will have to be determined before such questions can be answered.

Further Work

The obvious next step is to take the α-BuTX structure to higher resolution. Unfortunately, the best way to go about this is not nearly so obvious. The question of the true space group of the large-cell crystals has still not been solved, and even if it had been, it would still be necessary to determine the exact arrangement of the four (or more, depending on the final space group) molecules in the asymmetric unit. In addition, most of the derivatives are not usable beyond 3 to 4 Å resolution, and so the high-resolution structure could not be determined by the multiple isomorphous replacement method. It would instead have to be obtained from the low-resolution structure by Fourier refinement techniques or by phase extension based on the non-crystallographic symmetry (Jack, 1973; Bricogne, 1974).

Once the structure was solved at some reasonable resolution, other problems would arise when refinement was attempted. Depending on the final space group of the crystals, at least four molecules would have to be refined independently. Also, the large number of weak reflections, especially the $\underline{\ell}$ = odd reflections, would probably make refinement more difficult because of the considerable statistical error in the data. Finally, both the diffraction pattern and the low-resolution map indicate that regions of the structure are disordered in these crystals. Refining the structure in such areas would be quite difficult.

The final result of this analysis is that the crystals used for the low-resolution α-BuTX structure reported here are far from ideal for doing the high-resolution structure. It might therefore be better to use a different crystal form for the high-resolution work. As shown in Table 3-1, a second crystal form was produced at "high pH' (around pH 7) which contained only two molecules in the asymmetric unit and did not have the symmetry problems, or stability problems, of the pH 4 crystals. Unfortunately, these pH 7 crystals never grew large enough for data collection, but other conditions might make them grow larger. Their diffraction pattern was quite similar to the small-cell form of the pH 4 crystals, so the molecular packing must be nearly the same in these two crystal forms.

If this alternate crystal form were grown to a large enough size to use, either at high pH or at room temperature, the problem of solving the high-resolution structure would still remain. A low-resolution structure could readily be obtained by rotation of the current 5 Å structure into the new crystal form (Rossmann and Blow, 1962), but the high-resolution structure would still require the use of one or more of the techniques mentioned earlier. One hopeful note, however, is that many of the heavy-atom compounds tested originally were rejected because they forced the crystals into the small-cell form, and even some which were used encouraged this tendency. Thus, many of these compounds might form usable derivatives with the small-cell crystals, and it is quite possible that the quality of the derivatives used in the work reported here would improve significantly in small-cell crystals.

It appears that the most efficient route to a high-resolution structure of α-BuTX is the simultaneous pursuit of both the large-cell and small-cell crystal structures. Solving the large-cell structure appears to be possible, although with considerable difficulty; while solution of the small-cell structure must await crystals large enough for data collection, but should be easier to solve than the large-cell structure once the crystals are grown. By following both routes, all contingencies should be covered, and one or both of these paths should lead to the high-resolution structure.

References

Bricogne, G. (1974), Acta Crystallogr. A30, 395-405.

Chicheportiche, R., Vincent, J.-P., Kopeyan, C., Schweitz, H., and Lazdunski, M. (1975), Biochemistry 14, 2081-2091.

Chou, P.Y. and Fasman, G.D. (1974), Biochemistry 13, 222-245.

Dickerson, R.E., Kendrew, J.C. and Strandberg, B.E. (1961), Acta Crystallogr. 14, 1188-1195.

Hseu, T.-H., Liu, Y.-C., Wang, C., Chang, H., Hwang, D.-M., and Yang, C.-C. (1977), Biochemistry 16, 2999-3006.

Jack, A. (1973), Acta Crystallogr. A29, 545-554.

Lee, C.Y. (1972), Ann. Rev. Pharmacol. 12, 265-286.

Low, B.W., Preston, H.S., Sato, A., Rosen, L.S., Searl, J.E., Rudko, A.D., and Richardson, J.S. (1976), Proc. Nat. Acad. Sci. USA 73, 2991-2994.

Rossmann, M.G. and Blow, D.M. (1962), Acta Crystallogr. 15, 24-31.

Tsernoglou, D. and Petsko, G.A. (1976), FEBS Letters 68, 1-4.

Figure Captions

Figure 6-1: Part of the 5 Å, small-subcell, α-BuTX electron density map showing two molecules related to each other by the noncrystallographic symmetry. The photograph is centered on the noncrystallographic twofold axis at $x = 0.54$, $y = 0.67$. It shows approximately the portion of the unit cell from $x = 0.25$ to $x = 0.85$, $y = 0.35$ to $y = 0.95$ and $z = 1.30$ to $z = 0.35$. The view is in the $-z$ direction, and the $z = 1.30$ section is closest to the viewer. The molecule on the left is referred to as "molecule 1" in the text, and the one on the right is referred to as "molecule 2." Solvent regions are indicated by an S. The map was contoured in equal intervals of 1.5 σ (approximately 0.12 $e^-/\text{Å}^3$), starting with the first contour above zero.

Figure 6-2: Three views of a wooden model of the electron density for molecule 1 in Figure 6-1. The first contour of the 5 Å map (approximately 0.12 $e^-/\text{Å}^3$) was used in making the model. The labeled features are discussed in the text. a) A view towards the $-x$ direction. b) A view towards the $+y$ direction. c) A view towards the $-z$ direction. In this photograph, the model is in approximately the same orientation on the page as molecule 1 in Figure 6-1.

Figure 6-3: The 5 Å α-BuTX electron density map after averaging the density from the two noncrystallographically related molecules. This photograph contains approximately the same region of the unit cell as Figure 6-1. The map was contoured in equal intervals of approximately 0.21 $e^-/\text{Å}^3$, starting with the first contour above zero. In some areas, contours were also drawn at the 0.10 $e^-/\text{Å}^3$ level (dotted lines) in order to increase the continuity of the map.

Figure 6-4: The α-carbon positions for two erabutoxin b molecules rotated to give the best fit to the 5 Å α-BuTX electron density map in Figure 6-1. The erabutoxin coordinates were actually rotated to fit the density for molecule 1, not both molecules. This drawing is on approximately the same scale as Figure 6-1. Disulfide bonds are indicated by a double line. α-Bungarotoxin sequence numbering is used.

Figure 6-1

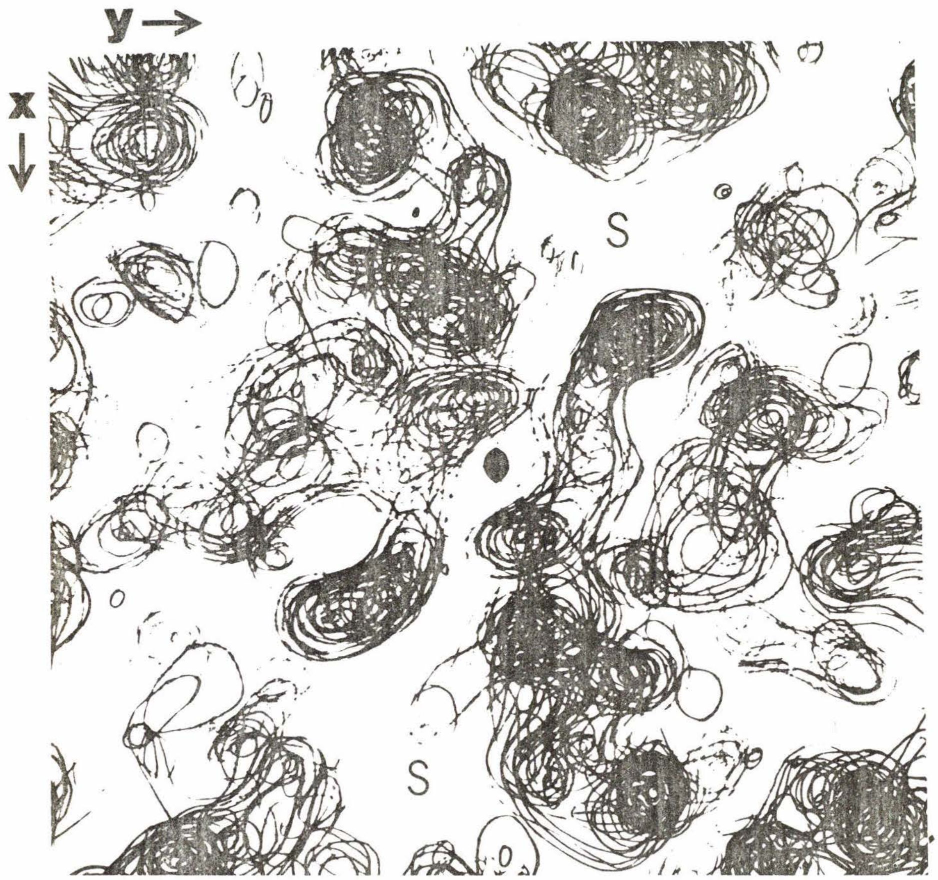


Figure 6-2a

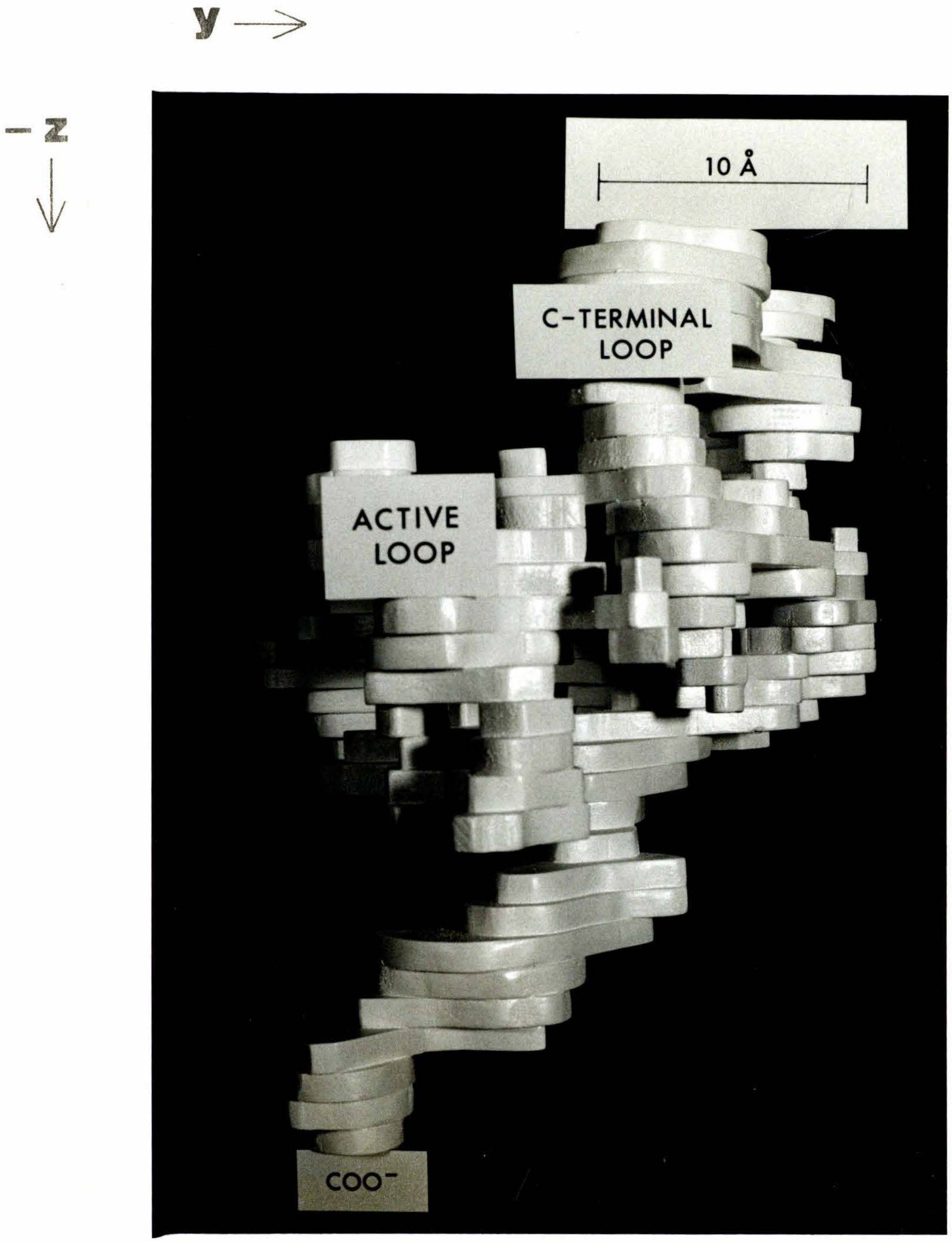


Figure 6-2b

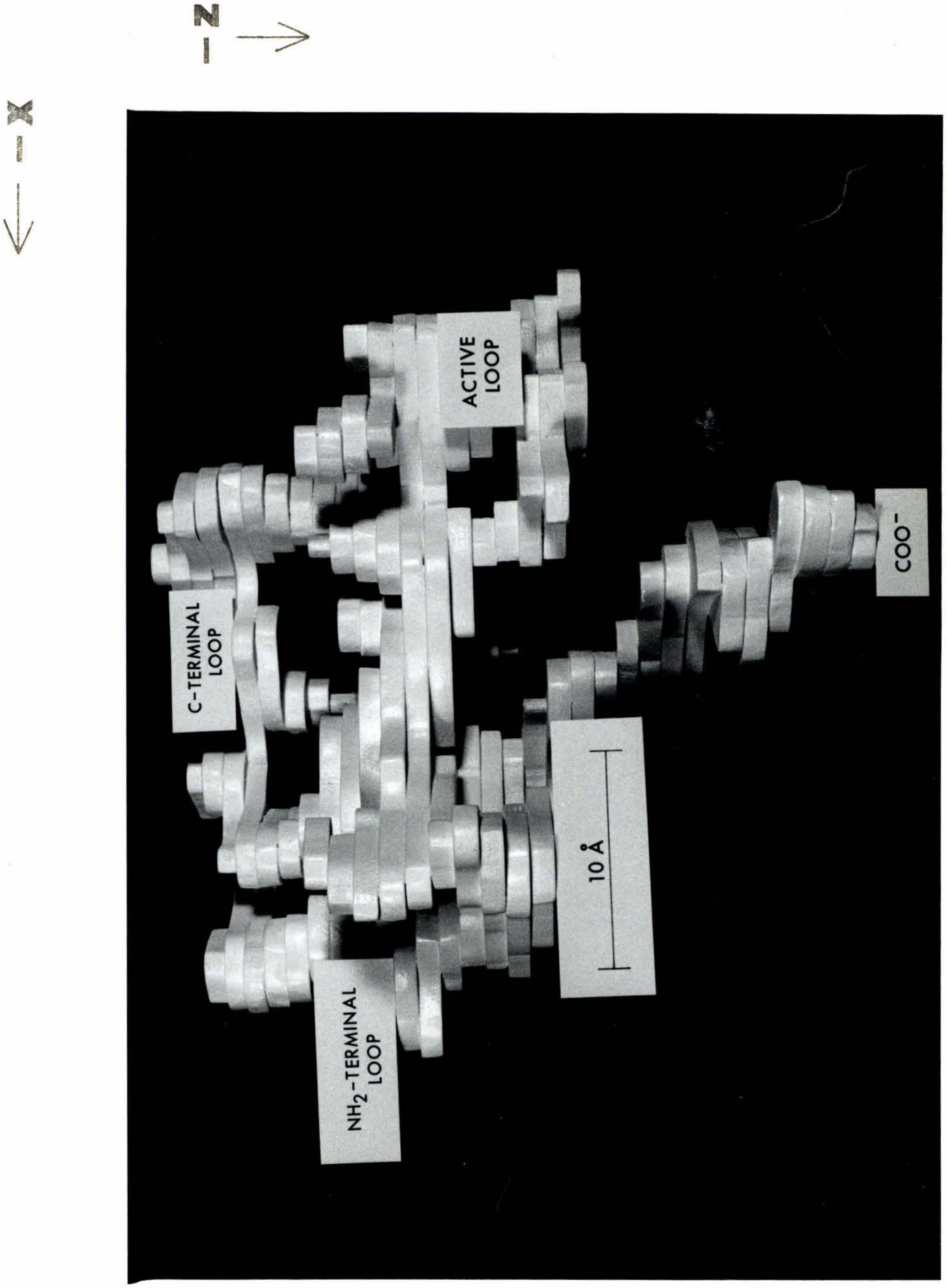


Figure 6-2c

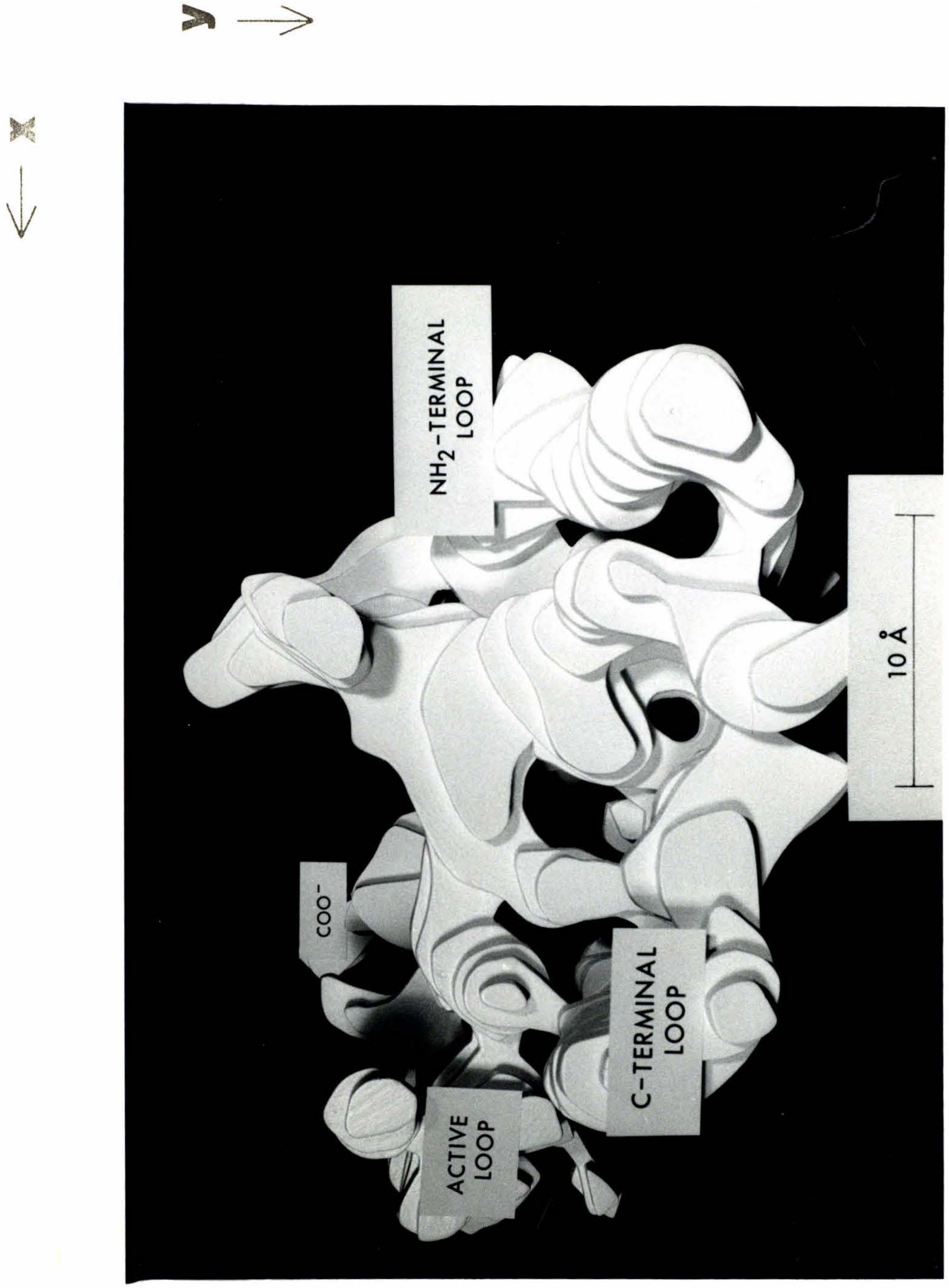


Figure 6-3

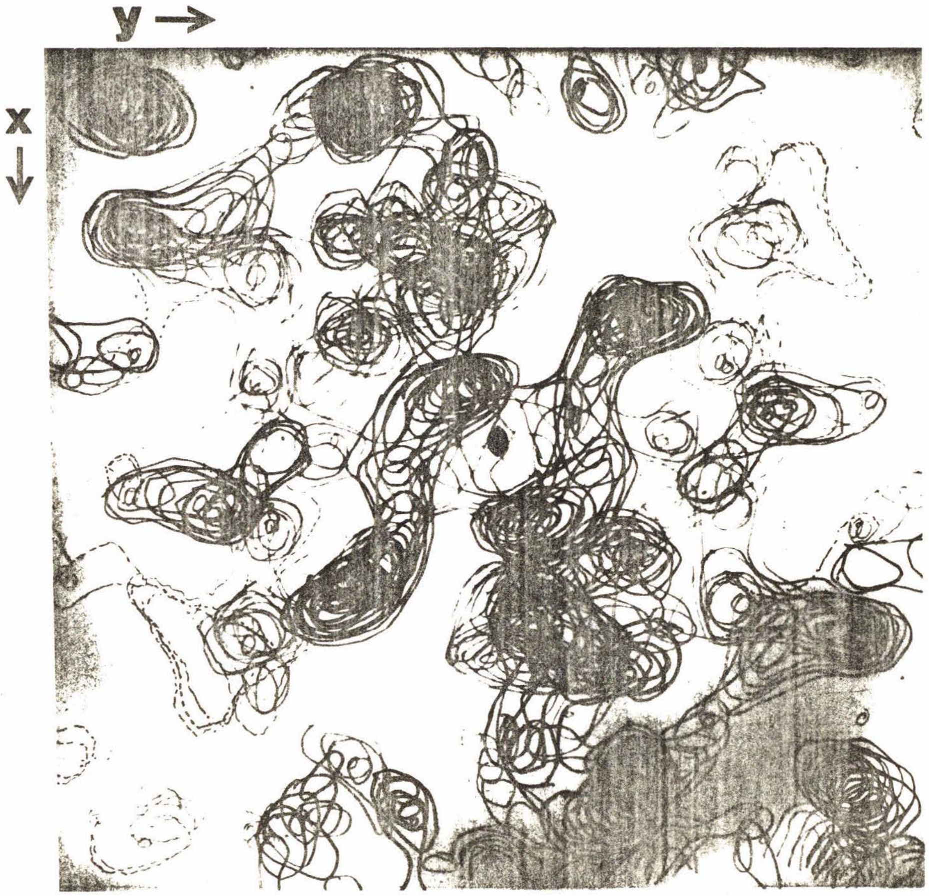
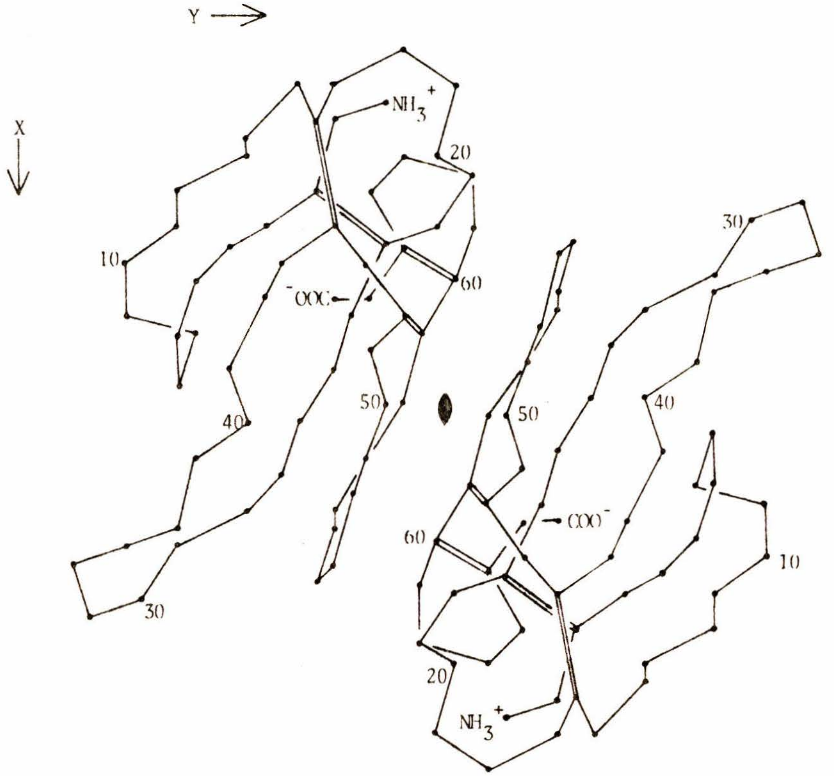


Figure 6-4



Appendix A

Abbreviations

α-BuTX: α-bungarotoxin (*B. multicinctus*)

BgV: bungarovenom (*B. multicinctus* venom)

BIS: N,N'-methylenebisacrylamide

CM-52: Whatman carboxymethyl cellulose ion exchange resin

DTT: dithiothreitol (Cleland's reagent)

EbTX: erabutoxin (*L. semifasciata*); toxin *b* unless otherwise noted

EDTA: ethylenediamine tetraacetic acid, disodium salt

Hg: $HgBr_2$ heavy-atom derivative

Ir: $K_3Ir(CN)_6$ heavy-atom derivative

MIR: multiple isomorphous replacement

NH_4OAc: ammonium acetate

Os: Na_2OsO_4 heavy-atom derivative

PtC: $(NH_4)_2Pt(CN)_4$ heavy-atom derivative

PtS: $Pt(SO_4)_2 \cdot 4H_2O$ heavy-atom derivative

SDS: sodium dodecyl sulfate

SIR: single isomorphous replacement

TEMED: N,N,N',N'-tetramethylethylenediamine

Tl: K_3TlCl_6 heavy-atom derivative

TNBS: trinitrobenzene sulfonate

Tris: tris (hydroxymethyl) aminomethane

Uranyl: $K_3UO_2F_5$ heavy-atom derivative

Appendix B

Final Heavy-Atom Parameters from the 5 Å, Small-Subcell Refinement in Space Group $P2_12_12_1$

Derivative	Site	x	y	z	A[1]	B[2]
$Pt(SO_4)_2$	1	0.404	0.062	0.006	10.65	6.80
	2	0.010	0.222	(-)0.007	5.72	"
$Pt(CN)_4^{-2}$	1	0.303	0.026	0.577	5.44	4.51
	2	0.113	0.174	0.403	4.36	"
	3	0.404	0.366	0.252	4.93	"
	4	0.126	0.233	0.794	2.14	"
OsO_4^{-2} [3]	1	0.305	0.293	0.225	6.50	0.54
	2	0.121	0.453	0.729	4.83	"
	3	0.466	0.271	0.067	1.80	"
OsO_4^{-2} [3]	1	0.305	0.293	0.231	3.38	0.54
	2	0.120	0.453	0.738	3.26	"
	3	0.467	0.265	0.068	0.96	"
$UO_2F_5^{-3}$	1	0.063	0.274	0.005	4.39	0.82
	2	0.353	0.113	(-)0.004	3.43	"
	3	0.304	0.038	0.626	1.80	"
$TlCl_6^{-3}$	1	0.363	0.120	0.069	4.57	4.40
	2	0.191	0.176	0.913	5.57	"
	3	0.223	0.318	0.165	9.34	"
	4	0.043	0.282	0.917	5.80	"
	5	0.387	0.354	0.213	2.59	"
	6	0.113	0.179	0.393	3.86	"
	7	0.303	0.031	0.613	3.61	"
	8	0.278	0.104	0.638	2.16	"
	9	0.311	0.298	0.215	2.20	"
	10	0.111	(-)0.007	0.656	3.41	"
$Ir(CN)_6^{-3}$	1	0.307	(-)0.014	0.502	1.16	0.75
	2	0.052	0.280	0.936	1.55	"
	3	0.362	0.121	0.096	1.65	"
	4	0.388	0.362	0.252	1.62	"

Derivative	Site	x	y	z	A	B
$Ir(CN)_6^{-3}$ (contd)	5	0.111	0.162	0.425	2.58	0.75
	6	0.270	0.087	0.382	1.83	"
	7	0.141	0.412	0.289	0.50	"
	8	0.122	0.225	0.782	0.69	"
$HgBr_2$	1	0.401	0.046	0.062	6.63	4.76
	2	0.016	0.183	(-)0.129	3.97	"
	3	0.112	0.200	0.364	4.79	"
	4	0.300	0.043	0.631	6.73	"
	5	0.377	0.326	0.179	2.20	"

---

[1] Occupancy of an arbitrary scale (approximately 0.22 x absolute value).

[2] Isotropic temperature factor as a function of $(2 \sin\theta)^2$.

[3] Two OsO_4^{-2} data sets from crystals with different degrees of osmium substitution were included in the refinement.

Appendix C

The Mean Values of |ΔF| and Lack-of-Closure as a Function of Resolution for Each Heavy-Atom Derivative. Values are from the 5 Å, small-subcell, $P2_12_12_1$ refinement.

		$2\ \sin\theta^2$					
		0.010	0.025	0.040	0.055	0.071	0.087
$Pt(SO_4)_2$	$<\|\Delta F\|>$[1]	19.6	16.3	13.3	11.8	9.6	8.5
	E_{av}[2]	7.2	5.7	5.5	5.3	5.1	6.1
$Pt(CN)_4^{-2}$	$<\|\Delta F\|>$	19.3	14.6	13.5	10.4	10.7	9.1
	E_{av}	10.6	8.6	7.8	8.9	8.0	7.4
OsO_4^{-2}[3]	$<\|\Delta F\|>$	21.6	16.0	13.7	15.2	13.5	12.8
	E_{av}	13.0	8.2	9.3	10.1	8.8	11.4
OsO_4^{-2}[3]	$<\|\Delta F\|>$	13.5	9.1	7.1	8.5	7.7	7.8
	E_{av}	8.8	5.0	4.4	6.0	5.3	5.9
$UO_2F_5^{-3}$	$<\|\Delta F\|>$	14.0	8.9	10.3	10.2	7.7	7.8
	E_{av}	7.4	4.2	5.6	4.8	4.2	4.7
$TlCl_6^{-3}$	$<\|\Delta F\|>$	22.2	21.7	17.1	18.1	15.8	14.0
	E_{av}	15.1	10.0	8.0	8.3	10.8	9.1
$Ir(CN)_6^{-3}$	$<\|\Delta F\|>$	8.8	8.3	8.9	8.0	7.1	7.0
	E_{av}	4.9	5.1	5.1	5.7	4.7	5.1
$HgBr_2$	$<\|\Delta F\|>$	23.6	15.5	12.3	12.2	11.4	10.7
	E_{av}	11.3	8.5	8.2	9.1	8.4	8.1

[1] $|\Delta F| = |k \cdot F_{PH} - F_P|$, where k is the Kraut scale factor.

[2] Average lack-of-closure for all reflections.

[3] Two OsO_4^{-2} data sets from crystals with different degrees of osmium substitution were included in the refinement.