1989

A reciprocal space approach for locating symmetry elements in Patterson superposition maps

Thomas L. Hendrixson
Iowa State University

Order Number 9014906

A reciprocal space approach for locating symmetry elements in Patterson superposition maps

Hendrixson, Thomas L., Ph.D.

Iowa State University, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A reciprocal space approach for locating symmetry elements in Patterson superposition maps

by

Thomas L. Hendrixson

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Chemistry
Major: Physical Chemistry

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1989

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

In 1934, A. L. Patterson devised a function, now known as the Patterson function, which consists of all of the interatomic vectors in the unit cell and can be obtained by Fourier transforming the observed intensities of the Bragg reflections. Until the early 1950s, almost all crystal structures were solved via analysis of this function. The positions of atoms in a structure were usually obtained by examination of two-dimensional projections of the Patterson function for characteristic interatomic vectors.

A new era dawned in 1952 with the development of a series of equations which showed that, theoretically, complete crystal structures could be obtained from the direct manipulation of the Bragg reflections. Such procedures became known as "direct methods". Much of the initial popularity of direct methods procedures was due to their ability to solve the nearly-equal atom structures, which are difficult to solve using Patterson methods. Today, the direct methods approach is quite often used in the first attempt at solving a crystal structure. While for many structures, direct methods will furnish at least a partial solution, there are still many structures that resist solution by the direct methods approach.

It is these structures that have encouraged a renaissance in the development of Patterson method

techniques by several research groups, notably including this one. Many of the groups have focused their efforts on the development of procedures that make use of known molecular fragments. However, this group has focused its efforts primarily on the development of ab initio methods, where no prior knowledge of the structure is required. Previously, indirect use of the space group symmetry was made by using Harker vectors and planes in these methods. A new procedure which directly makes use of the space group symmetry has been developed and automated in a series of computer programs called HYPAD.

In any structure determination, the primary goal is the determination of the positional and thermal parameters of the atoms present in the structure. These parameters affect the amplitude and phase of the scattered wave. The amplitude of the scattered wave is proportional to the observed intensity of the scattered wave, but the phase is not measurable experimentally. The need to determine the phases of the scattered waves is often referred to as the "phase problem" in X-ray crystallography. A brief description of the phase problem is given in Chapter 2, along with discussions of the approaches used in solving the phase problem; direct methods, Patterson, and Patterson superposition techniques.

Chapter 3 is devoted to the theory of the use of space group symmetry in Patterson superposition maps as utilized

in HYPAD. Reciprocal space approaches for the location of symmetry elements and the location of, and correction for, multiple images in superposition maps are developed. A brief description of the programs available in HYPAD is given in Chapter 4. The programs in HYPAD have been tested using several real structures, some of which were known previously and some which were unknown at the time. These structures are discussed in Chapter 5. Chapter 6 contains a summary of the viability of the HYPAD approach, and a discussion of possible modifications to HYPAD is given in Chapter 7.

Appendix A contains a listing of the keywords used in the various input files for HYPAD. A discussion of crystal structures completed using methods other than HYPAD are also listed in Appendix B.

CHAPTER 2. THE PHASE PROBLEM AND ITS SOLUTION

In every X-ray diffraction structure determination, the objective is the determination of the absolute phase shifts undergone by X-rays when they are scattered by a crystalline solid. This is what is known in crystallography as "the phase problem".

A Physical Description of the Problem

In a single crystal X-ray experiment, the incident X-ray beam interacts with the crystal resulting in a diffracted (or reflected) beam. Being an electromagnetic plane wave, the incident beam can be represented mathematically by

$$\psi(\mathbf{r}) = \psi_o \exp(2\pi i \mathbf{s}_o \cdot \mathbf{r}/\lambda) , \qquad (2.1)$$

where ψ_o is the amplitude of the wave and $\mathbf{s}_o$ is a unit vector in the direction of the propagation of the incident beam. For X-radiation, the wavelength, λ, is on the order of 1Å, which is comparable to the spacing between atoms in a crystalline material. Therefore, a crystal will act as a diffraction grating when it is placed in a beam of X-rays. Because X-rays are scattered most effectively by the electrons in the crystal, those crystals having larger concentrations of electrons will scatter the best. Each electron will scatter X-rays in every direction, and thus,

the j^{th} electron will produce a distribution of scattered waves,

$$\psi_j(r) = f_j \exp(2\pi i s \cdot r/\lambda) \ , \qquad (2.2)$$

where s is a unit vector in the direction of the scattered wave. In general, the direction of s is different from that of s_o. The scattering amplitude for a single electron, f_j, falls off as the scattering angle increases.

In a diffraction experiment, the intensities of the scattered waves are measured. Each scattered wave is the composite of the scattered waves from all of the individual electrons in the crystal, and is formed by the superposition of these waves. The individual waves can either reinforce each other or interfere with each other. If two parallel in-phase incident waves, having a wavelength of λ, are scattered by parallel planes in the crystal, separated by an interplane distance d, then they will scatter in-phase only for specific scattering angles θ. Mathematically, this is given by Bragg's Law[1],

$$\lambda = 2d\sin\theta \ . \qquad (2.3)$$

Similar phase relationships can be obtained when scattering from a distribution of electrons is considered. Suppose there are two scattering centers, C_1 and C_2, as shown in Figure 2.1. These scattering centers can be either individual electrons or small localized distributions of electrons (e.g., atoms). For a center C_j, the path

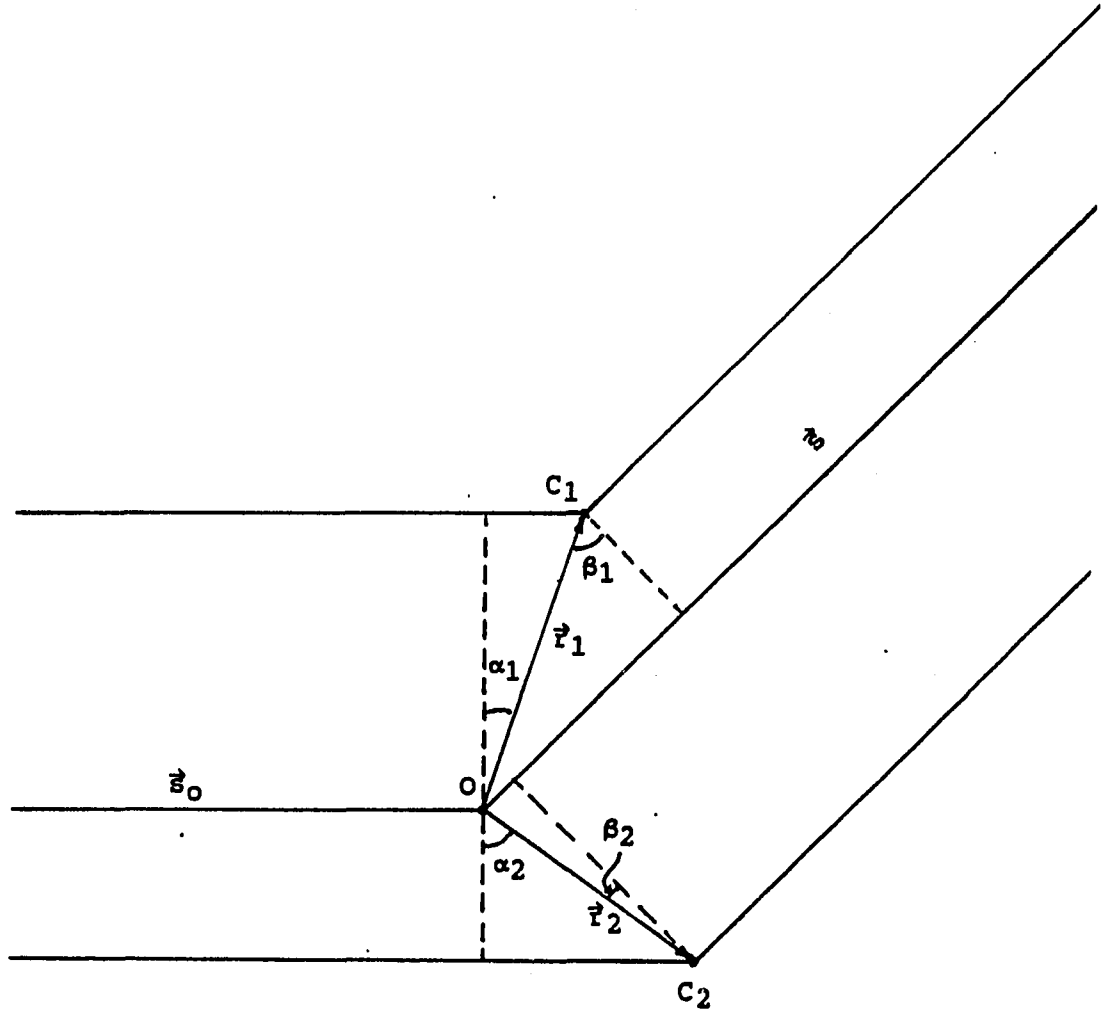


Figure 2.1. Schematic diagram for scattering from two points C_1 and C_2

difference between a wave scattered by the center and a wave scattered by an arbitrary origin O is given by

$$\Delta_j = a_j - a_{jo} = |r_j| \, (\sin\beta_j - \sin\alpha_j) \, . \qquad (2.4)$$

This gives a phase difference of

$$2\pi\Delta_j = 2\pi(s - s_o)\cdot r_j/\lambda = 2\pi h\cdot r_j \qquad (2.5)$$

where **h** is usually referred to as the diffraction vector.

The composite of the scattered waves is given by the structure factor, F_h, and is the sum of the individual scattering distributions

$$F_h = \sum_k^{N_e} \psi_k(h) = \sum_k^{N_e} f_j \exp(2\pi i h\cdot r_k)$$

$$= \sum_j^{N_a} f_j \exp(2\pi i h\cdot r_j) \quad , \qquad (2.6)$$

where N_e is the number of electrons in the unit cell, N_a is the number of atoms in the unit cell and f_j is the atomic scattering factor. The magnitude of F_h, $|F_h|$, is related to the extent to which the atoms in the crystal scatter in-phase for a particular reflection (i.e., the relative phase differences of the diffracted waves), and is independent of the absolute phase of the reflection.

The electron density of a structure, $\rho(r)$, can be written as the Fourier transform of the structure factors

$$\rho(r) = \int F_h \exp(-2\pi i h\cdot r) \, dh \quad , \qquad (2.7)$$

where the integral is over the entire space of wave vectors **h**, and $|F_h|^2$ is proportional to the observed intensity, I_h. Since, in general, the structure factors are complex

numbers, they can be expressed in terms of a phase and a magnitude

$$F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(2\pi i \phi_{\mathbf{h}})$$
$$= A_{\mathbf{h}} + iB_{\mathbf{h}} \quad . \tag{2.8}$$

Since, in practice, the integral in equation (2.7) is evaluated as a summation over a finite number of discrete points having integer values of $\mathbf{h} = (h,k,l)$, the electron density can be rewritten as

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{h}} |F_{\mathbf{h}}| \exp(i\phi_{\mathbf{h}} - 2\pi i \mathbf{h}\cdot\mathbf{r}) \quad . \tag{2.9}$$

Thus, if both the phase and the magnitude of all of the structure factors are known, then complete information about the structure can be obtained. However, the phases of the structure factors, unlike the magnitudes, can not be obtained directly from a diffraction experiment. This loss of phase information is referred to as the "phase problem" in crystallography.

For a structure having a reasonable amount of complexity, there are essentially two basic methods to use to solve the phase problem; a real space method and a reciprocal space method. The remainder of this chapter is devoted to a brief discussion of the background and application of these methods.

Direct Methods

The term "direct methods" is usually taken to mean those methods which attempt to derive the phases of the structure factors using only the intensity information. The birth of direct methods can be traced back to the late 1940s with the development of the Harker-Kasper inequalities[2] and the Karle-Hauptman determinantal inequalities[3], which are based on the fact that the electron density is positive everywhere. While some very simple centrosymmetric structures could be solved using these inequalities, they made very little impact in the practical solution of crystal structures.

In addition to the condition of non-negativity of the electron density, the concept of atomicity (i.e., the electron density can be resolved into separate regions representing atoms) can be imposed. This led to the derivation in 1952 of what is now known as the Σ_2 relation by Sayre[4], and independently by Cochran[5] and Zachariasen[6]. If the electron density function of a structure which consists of equal resolved atoms is squared, then the result looks similar to the original electron density function. The peaks in the squared function will be sharper and higher, but they will be in the same positions as in the original function. In terms of reciprocal space quantities, the S_2 relation is given by,

$$E_{\mathbf{h}} = \kappa \langle E_{\mathbf{h}-\mathbf{k}} E_{\mathbf{k}} \rangle_{\mathbf{k}} \quad , \tag{2.10}$$

where $\langle\rangle_{\mathbf{k}}$ represents the average over a complete set of vectors $\mathbf{k}$, and the $E_{\mathbf{h}}$ are the normalized structure factors. The vectors $\mathbf{h}$, $\mathbf{k}$, and $\mathbf{h}-\mathbf{k}$ form a closed triangle, or vector triplet, in reciprocal space.

Much of the early application of direct methods centered around the solution of centrosymmetric crystals. For centrosymmetric crystals, equation (2.10) becomes

$$s(\mathbf{h}) \simeq s\Big[\sum_{\mathbf{k}} s(\mathbf{k})\, s(\mathbf{h}-\mathbf{k}) \Big] \quad , \tag{2.11}$$

where $\simeq$ means "probably equals", $s(\mathbf{h})$ is the sign of $F_{\mathbf{h}}$, and the terms on the right-hand side of equation (2.11) are pairs of known signs which give an indication for $s(\mathbf{h})$. The probability that $s(\mathbf{h})$ is positive, given many contributor pairs was first derived by Cochran and Woolfson[7] and is given by

$$P_{+}(\mathbf{h}) = \frac{1}{2} + \frac{1}{2} \tanh\left[\frac{\sigma_3}{\sigma_2^3} |E_{\mathbf{h}}| \sum_{\mathbf{k}} E_{\mathbf{k}}\, E_{\mathbf{h}-\mathbf{k}} \right] \quad , \tag{2.12}$$

where σ_n is

$$\sigma_n = \sum_{j} (Z_j)^n \tag{2.13}$$

and Z_j is the atomic number of the j^{th} atom.

It was not until a few years later that a probability distribution for the general phase relationship

$$\phi_{\mathbf{h}} = \phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}} \tag{2.14}$$

was developed by Cochran[8]

$$P(\phi_h) = \frac{1}{2\pi I_0(\kappa_{h,k})} \exp(\kappa_{h,k}\cos(\phi_h - \phi_k - \phi_{h-k})) \quad , \quad (2.15)$$

where

$$\kappa_{h,k} = \frac{\sigma_3}{\sigma_2^{3/2}} |E_h \, E_k \, E_{h-k}| \quad (2.16)$$

and $I_0(\kappa_{h,k})$ is a modified Bessel function,

$$I_0(x) = \sum_{n=1}^{\infty} \frac{x^{2n}}{2^{2n} \, ((2n)!)^2} \quad . \quad (2.17)$$

The probability expressions quantify what would otherwise be a qualitative manipulation of the phases.

In 1956, Karle and Hauptman[9] introduced an equation which provides an estimate for $\phi(h)$ when several pairs of known phases are available. This was given in the form of the tangent formula which is

$$\tan \phi_h \simeq \frac{\sum_k \kappa_{h,k} \sin(\phi_k + \phi_{h-k})}{\sum_k \kappa_{h,k} \cos(\phi_k + \phi_{h-k})} \quad . \quad (2.18)$$

These developments gave crystallographers the tools to solve some of the structures that had previously been difficult to solve, such as structures containing atoms of equal or nearly equal atomic number, and led to the development of several structural solution strategies. One of the earliest such strategies was Symbolic Addition, developed by Karle and Karle[10]. In Symbolic Addition, algebraic symbols are assigned to the phases of a group of reflections, known as the starting set. The starting set is

chosen by taking those structure factors that have large $|E_h|$ and that can be combined with other members of the starting set to give phase estimates of many other reflections. Once a consistent set of symbolic phases has been obtained, the values of the phases can be calculated.

Probably the most widely used direct methods approach is MULTAN, developed by Germain, Main and Woolfson[11]. This approach is similar to that of Symbolic Addition, except that numerical values for the phases are substituted for symbols at the outset. These numeric phases are then refined using equation (2.10) and new estimates of the phases are calculated using equation (2.18).

Since the introduction of MULTAN, there have been many advances in the direct methods strategy. With the addition of various concepts, such as magic integers[12], the neighborhood principle[13], quartets[13-17], random starting phases[18], and the maximum-entropy method[19-21] to name a few, many fairly complex compounds can be solved using direct methods. However, there still remain numerous compounds that resist solution via direct methods. This has prompted a renewal of interest in real space, or Patterson methods.

Patterson Methods

Although the connection between Fourier theory and X-ray diffraction had been well known since 1913, when it was

first noted at a Solvay Conference, it was was not widely used until 1934, when A. L. Patterson introduced a Fourier series expression, now known as the Patterson function[22], which could be calculated directly from the observed intensities

$$P(\mathbf{u}) = \int \rho(\mathbf{r})\, \rho(\mathbf{r} + \mathbf{u})\, d\mathbf{r}$$
$$= \sum_{\mathbf{h}} |F_{\mathbf{h}}|^2 \exp(2\pi i \mathbf{h}\cdot\mathbf{u}) \quad . \tag{2.19}$$

The square of the magnitude of the structure factors can be written as

$$|F_{\mathbf{h}}|^2 = F_{\mathbf{h}}\, F_{-\mathbf{h}}$$
$$= \sum_{j}^{N} \sum_{k}^{N} f_j\, f_k \exp(2\pi i \mathbf{h}\cdot(\mathbf{r}_j - \mathbf{r}_k)) \quad . \tag{2.20}$$

Thus, peaks in the Patterson function occur when two atoms are separated by the vector **u**. Furthermore, the magnitude of a peak is proportional to a contribution of $Z_j Z_k$ for atoms j and k such that $\mathbf{r}_j - \mathbf{r}_k = \mathbf{u}$. The concept that the Patterson function represents all of the interatomic vectors in the unit cell is important to the understanding and interpretation of the function.

In a vector set notation where $\{a_i - a_1\}$, i=1,N, represents the set of interatomic vectors between atom 1 and all the atoms in the unit cell (<u>i.e.</u>, one image of the structure), the Patterson function is given as

$$[P(\mathbf{u})] \equiv \{a_i - a_1\} \cup \{a_i - a_2\} \cup \ldots \cup \{a_i - a_N\}$$
$$\equiv \{a_i - a_j\}, \quad i,j=1,N \quad . \tag{2.21}$$

Thus, for a structure containing N atoms, the Patterson function contains N images of the structure, superimposed upon one another. This can make interpretation of the Patterson function difficult, even for simple structures.

To illustrate, consider the hypothetical structure shown in Figure 2.2, ("Elephantine"[23]). This structure contains 4 heavier atoms, such as iron, per molecule and a long chain of carbons. In the Patterson map generated by this molecule, the dominant images are the iron images (i.e., those images formed by vectors having iron atoms at the tail of the vector). The Patterson map shown in Figure 2.3 is composed of only the iron images and thus contains 16 images of the structure. One can see that even for this relatively simple structure, the Patterson map is very complex.

In 1936, the concept of symmetry was introduced into Patterson analysis by D. Harker by the use of Harker vectors[24]. Harker vectors are interatomic vectors between symmetry-related atoms. The Harker vectors for a given space group have the same general form, regardless of the actual position of the atoms in question. Table 2.1 lists the Harker vectors for the space group $P2_1/c$. For structures containing only a few heavier atoms, which produce the more prominent peaks in the Patterson map, the Harker vectors and Harker planes can be used to infer positional information of some of the atoms.

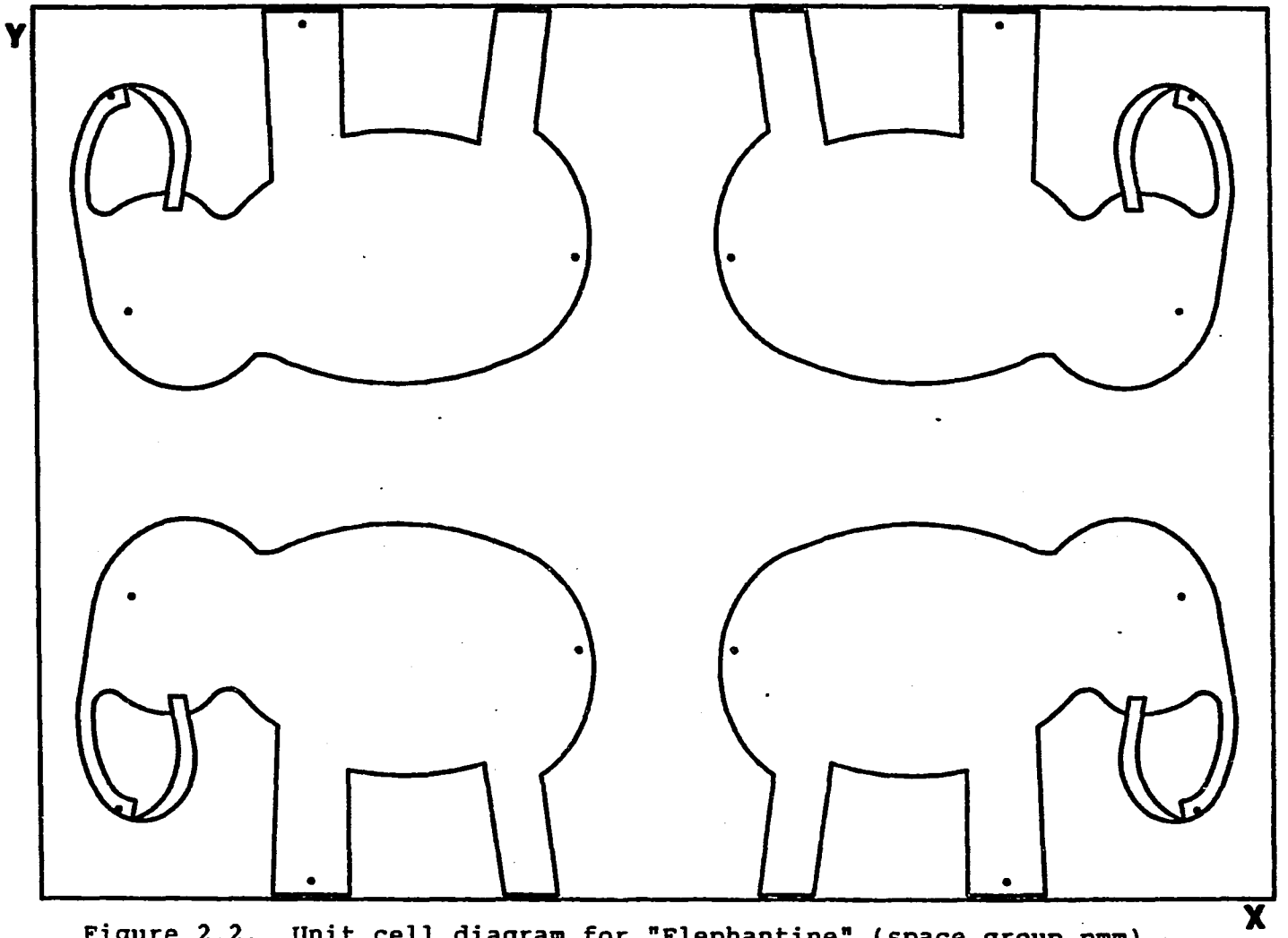


Figure 2.2. Unit cell diagram for "Elephantine" (space group pmm)

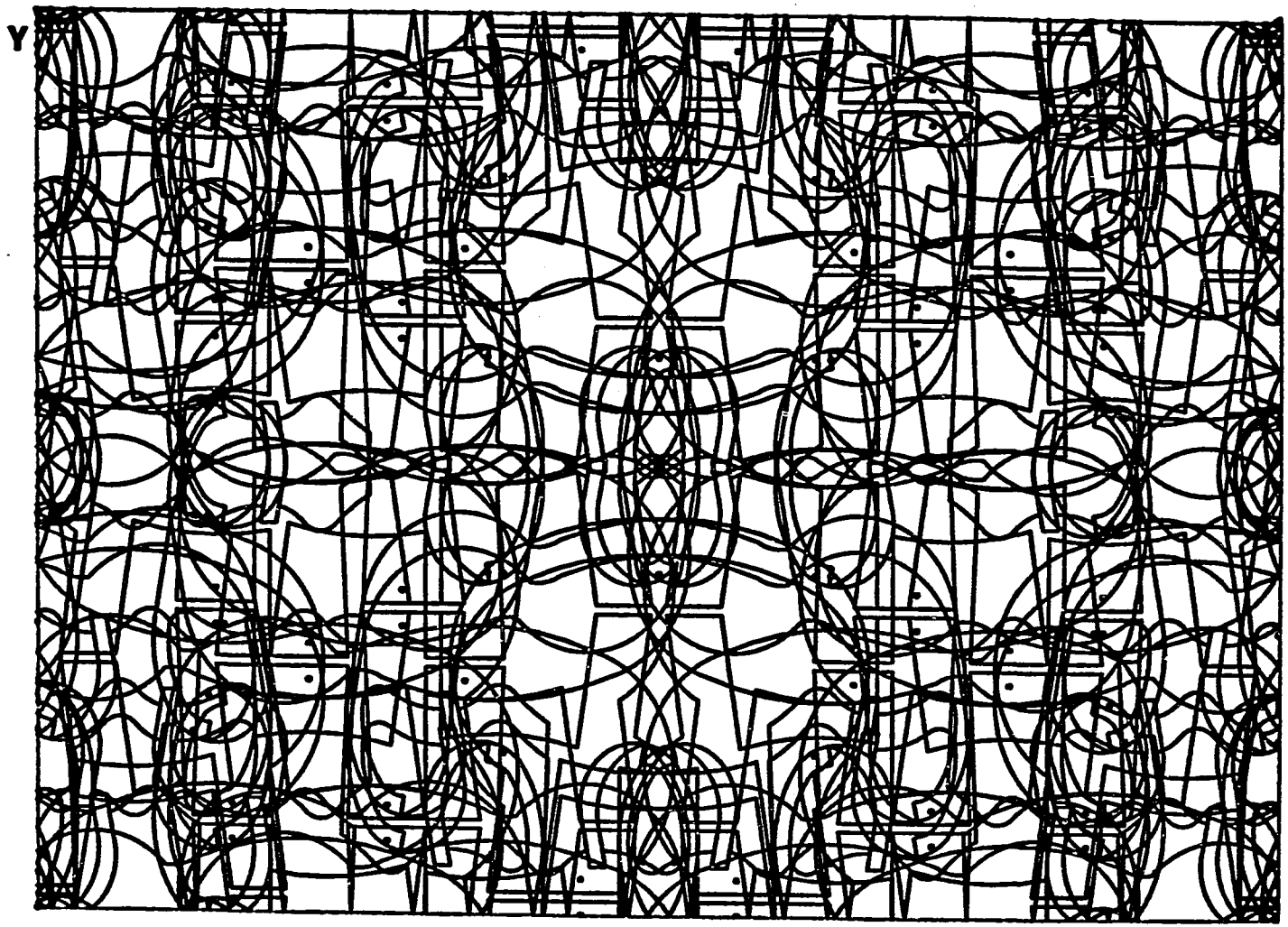


Figure 2.3. Patterson map for "Elephantine". Contains 16 complete iron images

Table 2.1. Harker vectors for the space group $P2_1/c$

	x,y,z	-x,y+1/2,1/2-z	x,1/2-y,1/2+z	-x,-y,-z
x,y,z	0	2x,1/2,1/2+2z	0,1/2+2y,1/2	2x,2y,2z
-x,y+1/2,1/2-z	-2x,1/2,1/2-2z	0	-2x,2y,-2z	0,1/2+2y,1/2
x,1/2-y,1/2+z	0,1/2-2y,1/2	2x,-2y,2z	0	2x,1/2,1/2+2z
-x,-y,-z	-2x,-2y,-2z	0,1/2-2y,1/2	-2x,1/2,1/2-2z	0

However, an ever-present problem in Harker analysis is that the 2_1-screw vector, for example, of one atom may accidentally combine with the glide vector of an unrelated atom to give the position of a peak which is actually present in the Patterson map but is not the inversion vector for either atom. A method, referred to as Vector Verification[25], has been developed to help combat this problem. In Vector Verification, the interatomic vectors based on the atomic positions obtained from the Harker analysis are calculated and then verified by checking to see if they are in the Patterson map.

Quite often, if the structure contains a small number of heavy atoms, the positions of these atoms can be used to estimate the phases of the F_h. Calculation of an electron density map using these phases will then provide possible positions for the rest of the structure. Of course, the assumption inherent in this approach is that the phases from the partial structure are close to the phases obtained from the entire structure. Obviously, this is not always the case.

In 1950, it was discovered by Clastre and Gay[26] and Garrido[27] that the superposition of a shifted Patterson map onto an unshifted map resulted in the superposition of a relatively small number of vectors. This set of vectors is referred to as the "reduced vector set". This superposition of Patterson maps results in the partial deconvolution of

the Patterson function. Much of the early systematic analysis of the superposition was done by Buerger[28], who showed that application of this procedure could lead to the deconvolution of the Patterson map into one image.

This can be best illustrated in terms of the vector set notation introduced earlier. The superposition can be thought of as the intersection of the set of unshifted vectors and the set of shifted vectors. If the shift vector s is a single (i.e., unique) interatomic vector, such that s = $(a_2 - a_1)$, then the superposition can be written as

$$\begin{aligned}[M(r)] &\equiv [\{a_i - a_j\} + (a_2 - a_1)] \cap [\{a_i - a_j\}] \\ &\equiv [\{a_i - a_1\} + \{a_2 - a_j\}] \cap [\{a_i - a_j\}] \\ &\equiv \{a_i - a_1\} \cup \{a_2 - a_j\} \ , \ i,j=1,N \ . \end{aligned} \tag{2.22}$$

Thus, two images remain after the superposition. One is an image as "viewed" from the atom at a_1 and the other is the inverse of the image as "viewed" from the atom at a_2. These two images are related by an inversion center at $(a_2 - a_1)/2$. Figure 2.4 shows the result of a superposition using a single vector for "Elephantine". If a second vector, $s_1 = a_3 - a_1$, is chosen to use in an additional superposition, then the result is

$$\begin{aligned}[M(r)] &\equiv [\{a_i - a_j\} + (a_2 - a_1)] \cap [\{a_i - a_j\} + (a_3 - a_1)] \cap [\{a_i - a_j\}] \\ &\equiv [\{a_i - a_1\}] \ , \ i,j=1,N \ . \end{aligned} \tag{2.23}$$

Thus, theoretically, one image of the structure can be obtained after two superpositions using single vectors.

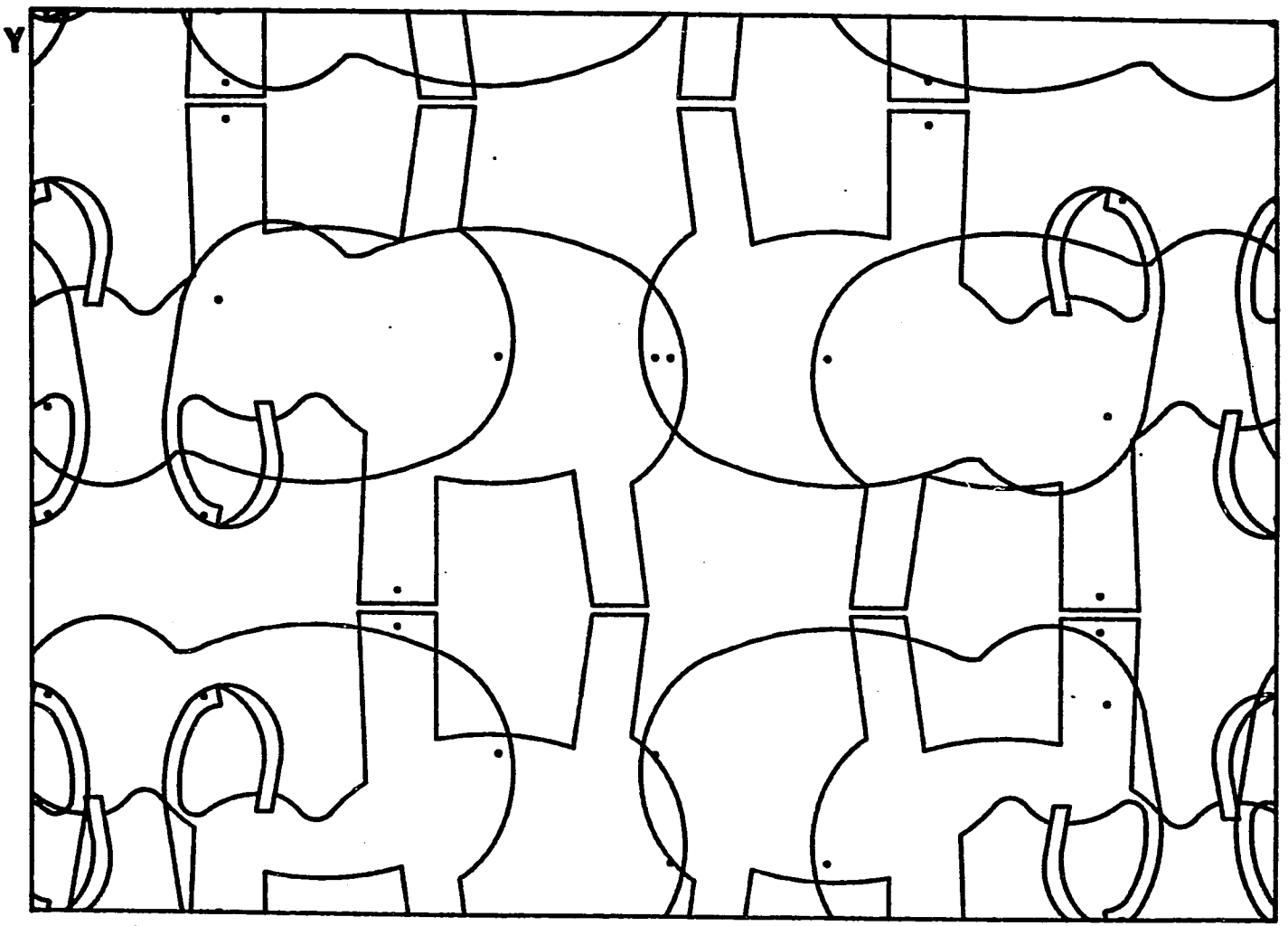


Figure 2.4 Superposition map for "Elephantine". Contains 2 complete iron images

Unfortunately, since multiple vectors are easier to find in a Patterson map, they are more likely to be used in the superposition. This can cause difficulties because the superposition map will then contain multiple images. If the shift vector is a multiple vector, $s = (a_2 - a_1) = (a_4 - a_3)$, then the superposition is

$$[M(r)] = \{a_i - a_1\} \cup \{a_2 - a_j\} \cup \{a_i - a_3\} \cup \{a_4 - a_j\} \ , \ i,j=1,N \qquad (2.24)$$

Two of the images are correctly oriented (from a_1 and a_3) and two of the images are inverse images (from a_2 and a_4).

There are a number of ways to do a Patterson superposition. The preferred way is do a "minimum" convolution of the Patterson function and the shifted Patterson function,

$$M(r) = \min (P(r), P(r + s)) \ . \qquad (2.25)$$

In practice, this is calculated by taking the point-wise minimum of $P(r)$ and $P(r + s)$ at all points r in the unit cell.

If the shift vector chosen for the superposition is an interatomic vector connecting atoms of different atomic number, then the vectors in the reduced vector set will have incorrect peak heights. If the shifted Patterson is weighted by the ratio of the atomic numbers of the atoms at the head and tail of the shift vector, then the reduced vector set will have correct weights[29]. For example, if the

shift vector is given by $s = (a_2 - a_1)$ and $\tau = Z_1/Z_2$, then the weighted superposition function is given by

$$M(r) = \min (P(r), \tau P(r + s)) \quad . \qquad (2.26)$$

If we refer to equation (2.22), a weighted superposition will produce heights for the vectors in the reduced vector set corresponding to the smaller of Z_iZ_1 and τZ_iZ_2, and Z_2Z_j and τZ_1Z_j, respectively. As an example, in order for the vector $(a_4 - a_1)$ to be appear in the superposition map, the vector $(a_4 - a_2)$ must be shifted by $(a_2 - a_1)$ in order to superimpose upon the vector $(a_4 - a_1)$ in the original map. In an unweighted superposition, the height of this vector would be the minimum of Z_4Z_2 and Z_4Z_1, or Z_4Z_2, which is the incorrect height for this vector. In a weighted superposition, the height would be the minimum of $(Z_1/Z_2)*(Z_4Z_2) = Z_1Z_4$ and Z_1Z_4, or Z_1Z_4. Weighting the superposition has the effect of suppressing one image relative to the other image. Thus, weighted superpositions are done whenever the atomic numbers of the atoms of the shift vector can be estimated and the weighting factor would be different from unity.

The study of macromolecules such as proteins and viruses in the 1950s led to the discovery that identical, but not crystallographically symmetry-related, units can often be found throughout the structure. This discovery prompted the development of a technique now known as Molecular Replacement[30], where the locations and orientations of these

units in electron density space could be calculated by locating the positions and orientations of the characteristic Patterson space patterns. This method gave rise to the development of translational and rotational functions[31] which measure the overlap of the Patterson function and a copy of the Patterson which has been transformed about a non-crystallographic symmetry element.

A method which is applicable when partial molecular geometry information is available was developed by Nordman[32]. In this approach, all of the interatomic vectors in the molecular fragment are calculated. This pattern of vectors is then rotated and translated in order to provide the best match with a portion of the Patterson map.

In the past decade, there has been a resurgence of interest in Patterson and Patterson superposition methods. Many computer programs, such as ALCAMPS[23], SHELXS[33], XFPS[34], IMPAS[35], and HASSP[36] have appeared. However, for most crystallographers, Patterson techniques still remain the last alternative, something to be tried only when direct methods techniques have failed completely, because the manipulations tend to be somewhat cumbersome.

The EG Relationship

One of the advantages of Patterson techniques is that they are easy to start. For a superposition, one simply

needs to choose a vector in the map. However successive application of superpositions entails a large amount of risk since all vectors used in the superposition must come from the same image. On the other hand, with a good starting set of phases, direct methods techniques will often result in a solution. However, with a bad set of starting phases, direct methods fail to provide a solution. Quite often, it can be difficult to find a good starting set of phases.

It would be advantageous if a method could be found whereby the easy start of the Patterson methods and the stability of the direct methods - given a good start - are combined. Such an approach has been examined by this group previously[37]. Using the assumption that the product of the Patterson superposition function and the electron density function looks like the square of the electron density function, then the following relation, called the EG relationship, is obtained

$$E_{\mathbf{h}} = K < E_{\mathbf{k}} G_{\mathbf{h}-\mathbf{k}} >_{\mathbf{k}} \quad , \tag{2.27}$$

where the $G_{\mathbf{h}-\mathbf{k}}$ are the Fourier transform coefficients of the Patterson superposition map. Unlike the $E_{\mathbf{h}}$, both the phase and the magnitude of the $G_{\mathbf{h}}$ are known, thus enabling a large number of terms to be included in equation (2.27) at the outset of a phase refinement.

CHAPTER 3. SYMMETRY ELEMENT LOCATION

In the past, procedures that have attempted to find symmetry elements in superposition maps have been real space methods. Typically, the major focus of attention has been an examination of the characteristic Harker vectors in an attempt to locate the symmetry elements. What will be presented here is a reciprocal space approach.

The Superposition Map as a Pseudo-electron Density Map

There exist relationships between the normalized structure factors, E_h, which arise from the symmetry elements present in the unit cell. However, the Fourier transform coefficients of the superposition map, G_h, only have P1 symmetry, in general. Since the structure under study usually has higher symmetry than P1, it is desirable to transform the G_h into some G'_h, where the G'_h have symmetry relationships that arise from the symmetry elements present.

If we examine the ideal case where the structure consists of point atoms and a unique (i.e., single) vector is chosen as the superposition vector, the resulting superposition map would contain two images of the structure and no extraneous peaks. If we further add the condition that the superposition be weighted, then the superposition

map will contain predominantly one image of the structure, with the inverse image substantially reduced by comparison.

Since the origin of the superposition map lies on an atom, and, in general, the origin of the electron density map is defined relative to the symmetry elements, the superposition map can be considered not only to be a first approximation to an electron density map, but one with the origin in the wrong position. The coefficients G_h would then be related to the F_h by

$$G_h = \kappa F_h \exp(-2\pi i \mathbf{h}\cdot\mathbf{t}_{org}) \quad , \tag{3.1}$$

where $\mathbf{t}_{org}$ is the vector from the location (in the superposition map) of the electron density map origin to the origin of the superposition map, and κ is a scale factor. Equation (3.1) can be rewritten as

$$\begin{aligned} F_h &= \kappa G_h \exp(2\pi i \mathbf{h}\cdot\mathbf{t}_{org}) \\ &= \kappa G_h^{\mathbf{t}_{org}} \quad , \end{aligned} \tag{3.2}$$

which produces the structure factor-like quantities, $G_h^{\mathbf{t}_{org}}$. These new coefficients should have the desired symmetry relationships.

Therefore, once the location of the electron density map origin in the superposition map has been determined, the G_h can be transformed using equation (3.2). The phases of the transformed coefficients can be used as a first approximation to the phases of the F_h and these phases can

in turn be further refined by use of the Σ_2 relation or the EG relation.

The Location of Symmetry Elements in Superposition Maps

Previous attempts at locating the position of symmetry elements have focused their efforts on examining the characteristic Harker vectors for solutions which occur frequently. One of the difficulties with this type of analysis is that the peak positions from the superposition map are approximate. This means that only rarely will exact matches be found; the user must decide on a tolerance within which matches will be considered "exact". However, if one could use the Fourier transform coefficients of the superposition map to determine the location of the symmetry element, then the problem of "how close is close" can be eliminated.

The unit cell of an alternate form of the elephant compound is represented in Figure 3.1. This form of the compound contains a glide line perpendicular to the y axis at y = 1/2. (The glide line reflects the image in the y direction and translates 1/2 of the unit cell in the x direction.) Placement of the glide line at some arbitrary position t, as shown in Figure 3.2, will generate the dashed image. If the glide line is incorrectly placed, then the dashed image will only partially overlap the the original,

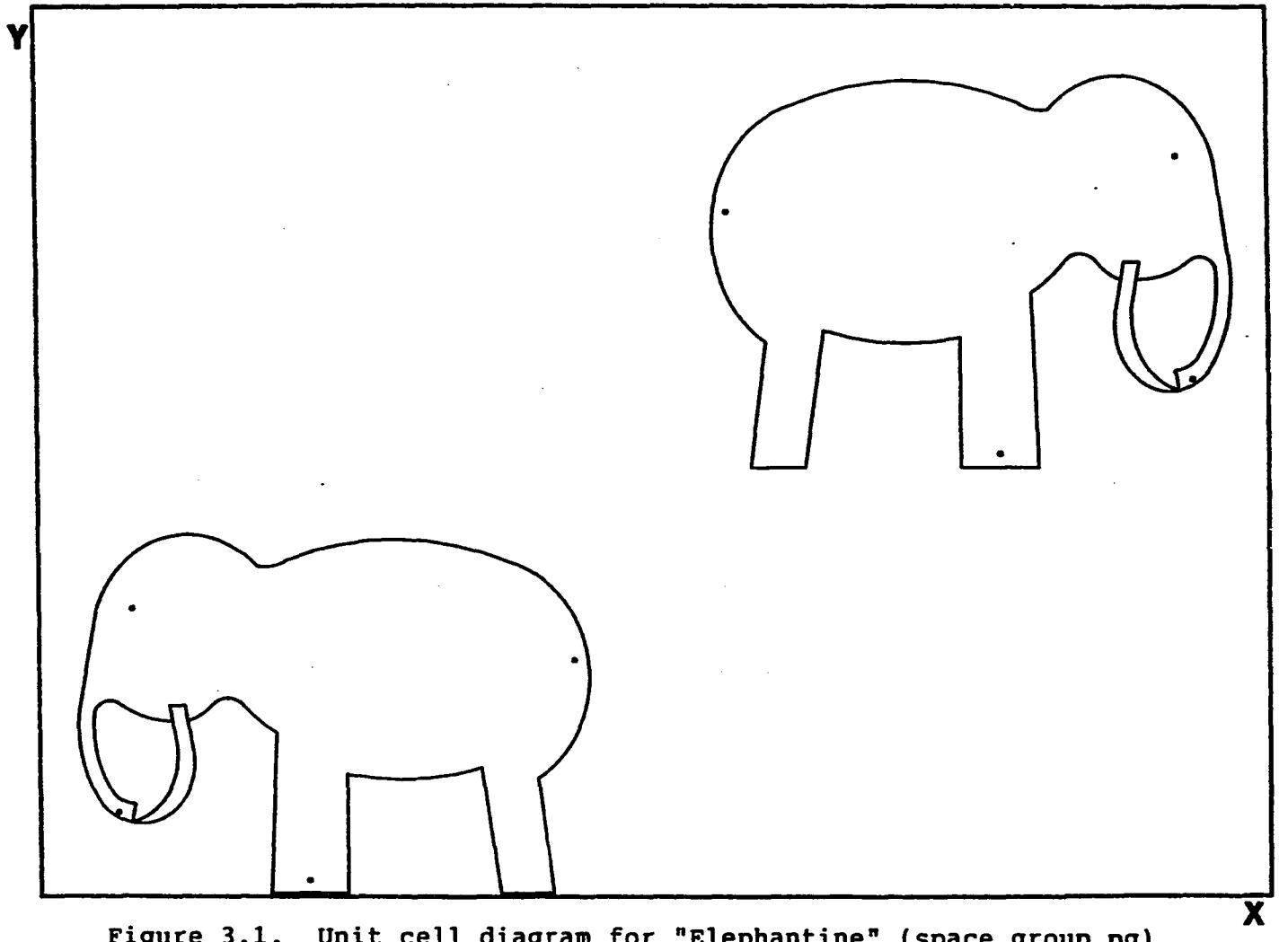


Figure 3.1. Unit cell diagram for "Elephantine" (space group pg)

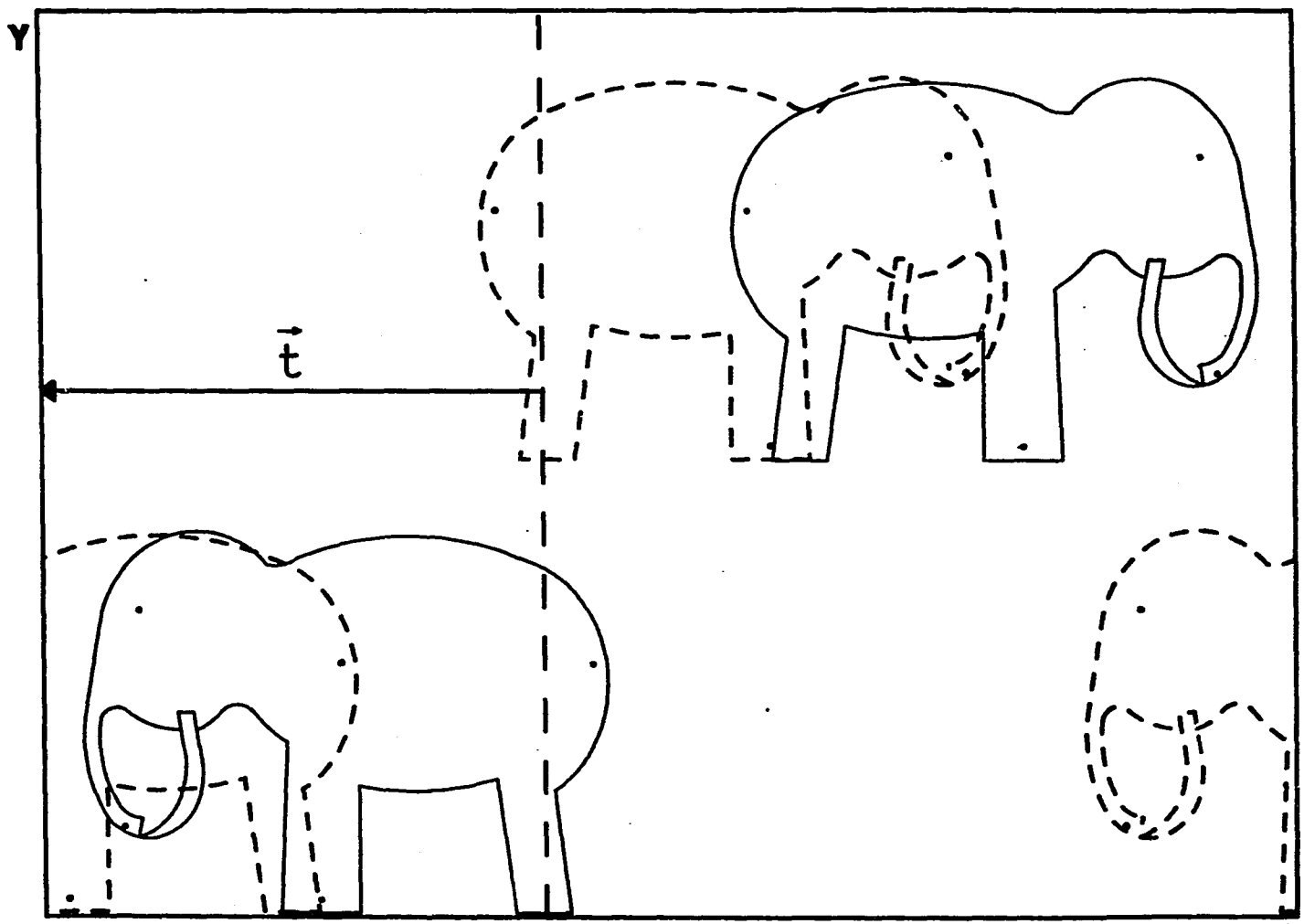


Figure 3.2. Q(S̈,t) diagram for "Elephantine" (space group pg)

solid image, if at all. However, when the glide line is placed in the correct location, the two images will overlap completely, maximizing the amount of overlap.

This overlap can be expressed in terms of the product of the original map and the map produced by the symmetry operator, and is given mathematically by

$$Q(\hat{S},t) = \int M(r|\hat{E},0)\, M(r|\hat{S},t)\, dr \quad , \tag{3.3}$$

where $M(r|\hat{S},t)$ is the map produced by placing the symmetry operator $\hat{S}$ at the position t, $\hat{E}$ is the identity operator, and the integral is over all space.

In general, a symmetry operator $\hat{S}$, located at a position t (with t defined as shown in Figure 3.2), can be written in terms of its rotational component, R, and its translational component, T. The result of a symmetry operator $\hat{S}$ operating on a point r is given by

$$\hat{S}r = R\cdot r + (R - I)\cdot t + T \quad , \tag{3.4}$$

where I is the identity matrix. Using this definition, the Fourier transform coefficients of $M(r|\hat{S},t)$, H_h, are given by

$$\begin{aligned} H_h &= \sum_j f_j \exp(2\pi i h\cdot \hat{S}r_j) \\ &= \sum_j f_j \exp(2\pi i h\cdot [R\cdot r_j + (R - I)\cdot t + T]) \\ &= G_{hR} \exp(2\pi i h\cdot [(R - I)\cdot t + T]) \quad , \end{aligned} \tag{3.5}$$

where G_{hR} is a Fourier transform coefficient of the original superposition map. Thus, the map produced by the symmetry operator can be expressed in terms of the original map.

Using the Fourier transform of the maps $M(\mathbf{r}|\hat{E},0)$ and $M(\mathbf{r}|\hat{S},t)$, the expression for the overlap, $Q(\hat{S},t)$ is

$$Q(\hat{S},t) = \sum_{\mathbf{h}} \sum_{\mathbf{k}} G_{\mathbf{h}} H_{\mathbf{k}} \int \exp(-2\pi i(\mathbf{h} + \mathbf{k})\cdot\mathbf{r})\, d\mathbf{r}$$

$$= \sum_{\mathbf{h}} G_{\mathbf{h}} H_{-\mathbf{h}} \quad . \tag{3.6}$$

Substituting the definition of the coefficients $H_{\mathbf{h}}$ into equation (3.6), the result

$$Q(\hat{S},t) = \sum_{\mathbf{h}} |G_{\mathbf{h}}||G_{\mathbf{h}'}| \exp(i(\alpha_{\mathbf{h}} - \alpha_{\mathbf{h}'} + 2\pi\mathbf{h}\cdot[(\mathbf{I} - \mathbf{R})\cdot\mathbf{t} - \mathbf{T}])) \tag{3.7}$$

is obtained, where $\mathbf{h}' = \mathbf{h}\cdot\mathbf{R}$.

Ideally, if the $G_{\mathbf{h}}$'s had exact symmetry, then the term in the exponential would be zero for the correct translation and $Q(\hat{S},t)$ would be a maximum,

$$Q^{max}(\hat{S},t) = \sum_{\mathbf{h}} |G_{\mathbf{h}}||G_{\mathbf{h}'}| \quad . \tag{3.8}$$

However, since the $G_{\mathbf{h}}$'s only have approximate symmetry, $Q(\hat{S},t)$ will be less than $Q^{max}(\hat{S},t)$, but will still likely be a maximum for the correct translation. For the purposes of comparison, all $Q(\hat{S},t)$ values will be referenced to $Q^{max}(\hat{S},t)$

Since $Q(\hat{S},t)$ is a maximum when the term in the exponential is small, one can alternatively try to minimize the phase difference between $G_{\mathbf{h}}$ and $G_{\mathbf{h}'}$. If we define a function

$$S(\hat{S},t) = \frac{1}{N} \Sigma \; |G_h||G_{h'}| \; ((\alpha_h - \alpha_{h'})/2\pi + h\cdot(I - R)\cdot t - h\cdot T)^2 \quad , \tag{3.9}$$

where $|G_h||G_{h'}|$ is included as a weighting function and N is the number of terms in the summation, it should be a minimum for the correct translation.

In practice, both the Q and the S functions are evaluated. Q is computationally faster to evaluate, but the S function is more sensitive to the presence or absence of a symmetry operation. To determine the "best" location from a group of choices, a figure-of-merit (FOM) is used. The figure-of-merit is given by

$$FOM(\hat{S},t) = \frac{1}{\varepsilon}\left[\frac{Q(\hat{S},t)}{Q^{max}_{group}} + \frac{S^{min}_{group}}{S(\hat{S},t)}\right] \quad , \tag{3.10}$$

where Q^{max}_{group} is the maximum $Q(\hat{S},t)$ value in the group, S^{max}_{group} is the minimum $S(\hat{S},t)$ value in the group, and ε is a parameter which takes into account the inability of the superposition to completely remove the Patterson symmetry. The form of ε was established by trial-and-error: for t = {(0,0,0), (0,0,½), (0,½,0), (0,½,½), (½,0,0), (½,0,½), (½,½,0), (½,½,½)},

$$\varepsilon = 1 + \frac{\min(5-j,4)}{4} \quad , \tag{3.11}$$

where j is the number of superposition vectors used to create the map, otherwise $\varepsilon = 1$.

In the preceding derivations, it has been assumed that the symmetry elements of a space exist independently of one another. However, for a given space group, the presence of a symmetry element $\hat{S}_1$, located at a position t_1, may restrict the positions (t_j) of other symmetry elements, $\hat{S}_j$.

One of the properties of a group is that the product of two elements of the group must be itself an element of the group, which means

$$\hat{S}_1\hat{S}_2\mathbf{r} = \hat{S}_3\mathbf{r} \quad . \tag{3.12}$$

Using the definition of a symmetry operator given in equation (3.4), the following relation between the locations of the symmetry operators and the translational components of the symmetry operators is obtained:

$$\mathbf{R}_2\cdot(\mathbf{R}_1 - \mathbf{I})\cdot\mathbf{t}_1 + (\mathbf{R}_2 - \mathbf{I})\cdot\mathbf{t}_2 - (\mathbf{R}_3 - \mathbf{I})\cdot\mathbf{t}_3 = \mathbf{T}_3 - \mathbf{R}_2\mathbf{T}_1 - \mathbf{T}_2 \quad . \tag{3.13}$$

For the product of 3 symmetry operators, the analogous expression

$$\mathbf{R}_3\cdot\mathbf{R}_2\cdot(\mathbf{R}_1 - \mathbf{I})\cdot\mathbf{t}_1 + \mathbf{R}_3\cdot(\mathbf{R}_2 - \mathbf{I})\cdot\mathbf{t}_2 + (\mathbf{R}_3 - \mathbf{I})\cdot\mathbf{t}_3 - (\mathbf{R}_4 - \mathbf{I})\cdot\mathbf{t}_4 = \mathbf{T}_4 - \mathbf{R}_3\cdot\mathbf{R}_2\cdot\mathbf{T}_1 - \mathbf{R}_3\cdot\mathbf{T}_2 - \mathbf{T}_3 \quad , \tag{3.14}$$

is obtained.

Thus, evaluating the functions $Q(\hat{S},t)$ and $S(\hat{S},t)$ while constraining the locations of different symmetry operators to satisfy equation (3.13) will enable the locations of the symmetry operators to be correctly determined.

Location of Multiple Images

The Patterson superposition map is a partial deconvolution of the Patterson map. Only rarely does it approach a full deconvolution, which is an electron density map, or a single image of the structure. As such, it would be expected that the superposition map would contain more than one image. As shown in Chapter 2, an unweighted superposition or a superposition using a multiple vector will contain more than one image. These extra images, which could be entire images or fragments of images, will have the largest effect on the phases of the G_h. The phases as obtained from the Fourier transform of this multiple image map would give erroneous values for the phases of the F_h if a transformation such as equation (3.2) were used. These errors, when introduced into the EG relation, or a Σ_2 relation, would likely cause the refinement to diverge. Thus it would be advantageous to be able to locate the additional image(s) in the superposition map and correct the G_h for this(these) images.

Let us again consider the glide form of the compound "Elephantine". If we have a superposition map containing two images of the structure, as shown in Figure 3.3, then shifting a copy of the map by the appropriate translation t should place the dashed image on the solid image. Thus the overlap of the two maps would be a maximum. Similar to the

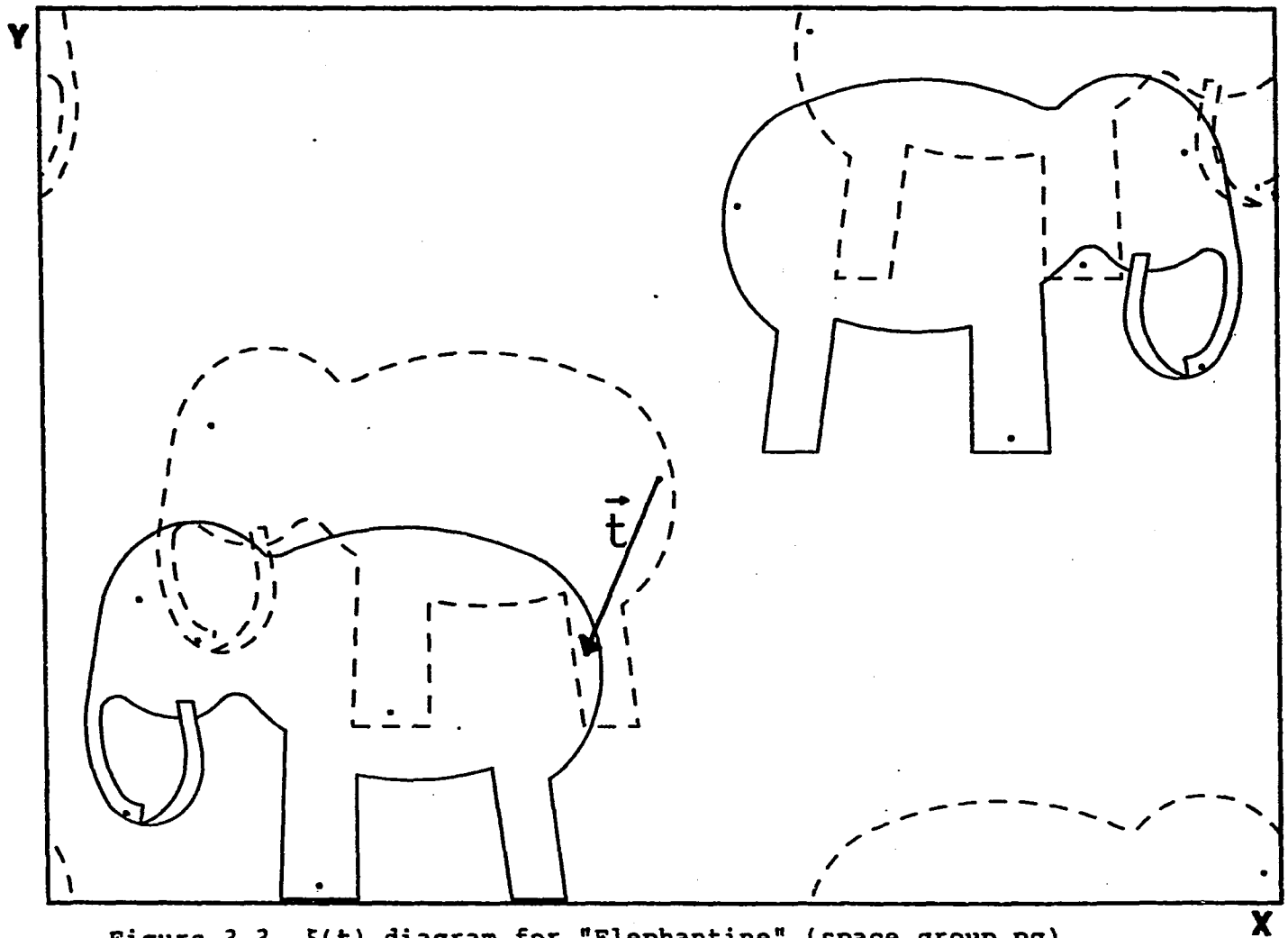


Figure 3.3 $\xi(t)$ diagram for "Elephantine" (space group pg)

case of the $Q(\hat{S},t)$ function, this overlap function, $\xi(t)$, is given by

$$\begin{aligned}\xi(t) &= \int M(r|\hat{E},0)\ M(r+t|\hat{E},0)\ dr \\ &= \sum_h \sum_k G_h\ G_k \exp(-2\pi i k\cdot t) \int \exp(-2\pi i(h+k)\cdot r)\ dr \\ &= \sum_h G_h\ G_{-h} \exp(2\pi i h\cdot t) \\ &= \sum_h |G_h|^2 \exp(2\pi i h\cdot t) \quad . \end{aligned} \tag{3.15}$$

One also has to be concerned that there may be an inverse image in the map. This situation is analogous to the presence of a non-inverted image and thus, the overlap is given by

$$\xi'(t) = \sum_h G_h\ G_h \exp(-2\pi h\cdot t) \quad . \tag{3.16}$$

Once the translation for the location of an additional image has been located, the G_h of the original map should be modified so as to remove the effect of the extra image. There are two reasonable methods of doing this. The first method would be to say that the transformed, or new, map is the sum of the original map and the map backshifted by the translation t. This is given by

$$M^{new}(r) = M(r) + M(r-t) \quad . \tag{3.17}$$

The Fourier transform coefficient of the new map, in terms of the Fourier transform coefficients of the original map, are

$$G_h^{new} = G_h + G_h \exp(2\pi i h\cdot t) \quad . \tag{3.18}$$

This definition has the disadvantage that if there is a peak at r but not at r - t in the original map, there will still be a peak in the new map at r.

The second method would be to say that the new map is the product of the original map and the backshifted map,

$$M^{new}(r) = M(r)\, M(r - t) \quad . \tag{3.19}$$

Using this definition, the Fourier transform coefficients of the new map are

$$G_h^{new} = \sum_k G_{h-k}\, G_k \exp(2\pi i k \cdot t) \quad . \tag{3.20}$$

This definition removes the disadvantage found in the additive map, but it has the disadvantage of distorting the original map if the structure is not an equal-atom case.

In the case of an unweighted superposition, the presence of the inverse image causes a small, local maximum where the global minimum should be (i.e., the location of the symmetry element) in the $S(\hat{S},t)$ function. A modified version of $S(\hat{S},t)$ has been developed which "mirrors" those phase differences, Δ $(-\pi/2 \leq \Delta \leq \pi/2)$, that are greater than $\pi/4$ across $\pi/4$,

$$S(S,t) = \frac{1}{4\pi^2} \sum_h |G_h||G_{h'}|\, \Phi^2 \tag{3.21}$$

where

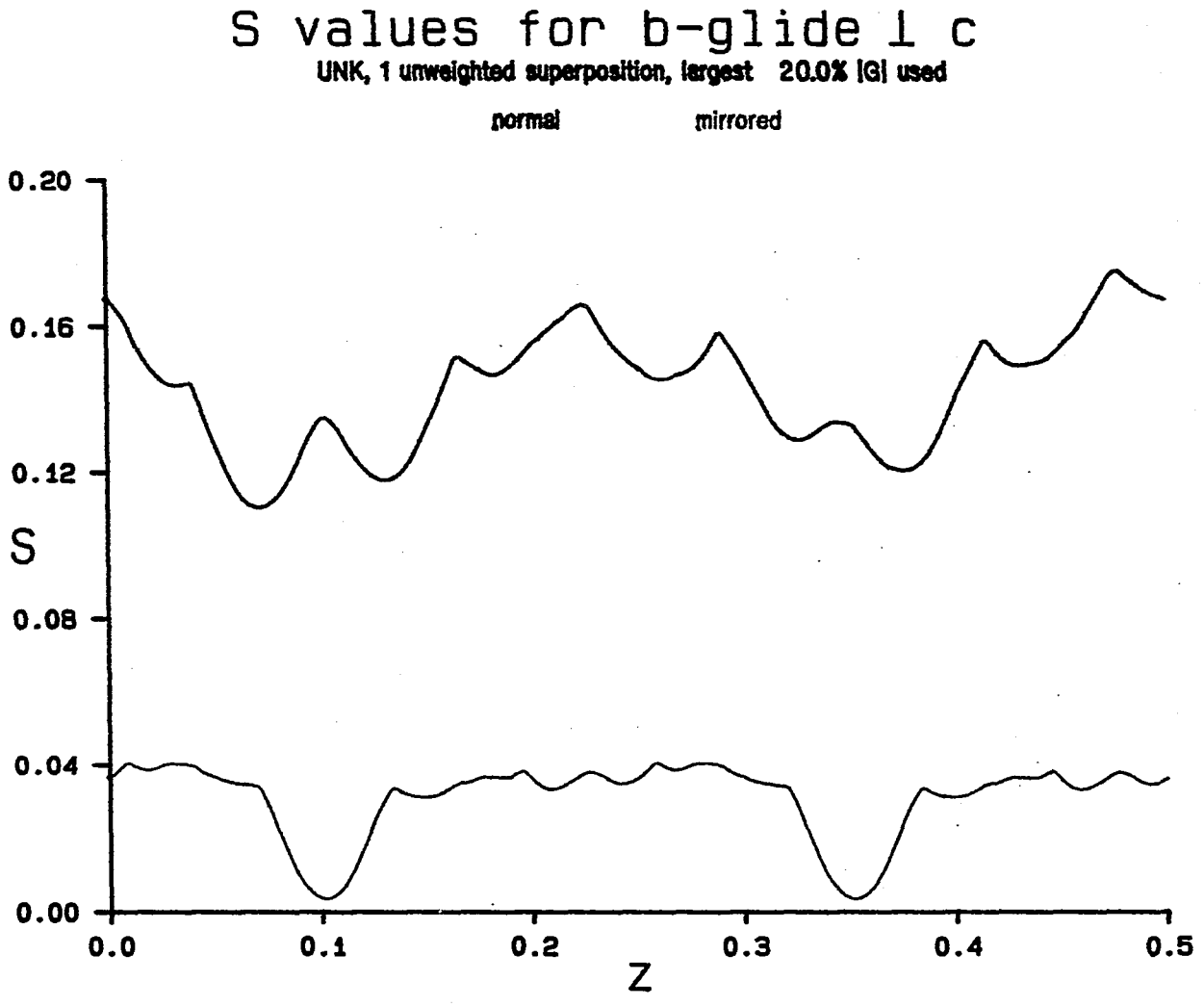


Figure 3.4. "Mirrored" and normal $S(\hat{S},t)$ functions

$$\Phi = \begin{cases} \frac{\pi}{2} - \Delta, & \Delta > \frac{\pi}{4} \\ \Delta, & -\frac{\pi}{4} \leq \Delta \leq \frac{\pi}{4} \\ -\frac{\pi}{2} - \Delta, & \Delta < \frac{\pi}{4} \end{cases} , \tag{3.22}$$

and

$$\Delta = \alpha_{\mathbf{h}} - \alpha_{\mathbf{h}'} + 2\pi \mathbf{h}\cdot(\mathbf{I} - \mathbf{R})\cdot\mathbf{t} - \mathbf{h}\cdot\mathbf{T} \quad . \tag{3.23}$$

A plot of the "mirrored" $S(\hat{S},t)$ function and the normal $S(\hat{S},t)$ function is shown in Figure 3.4. The "mirrored" $S(\hat{S},t)$ function removes the local minimum at the symmetry element location, but it also produces a solution 1/4 of the unit cell away from the actual solution. This is simply a result of the "mirroring" and such pseudo-solutions can be eliminated after examination of the $Q(\hat{S},t)$ function.

CHAPTER 4. DETAILS OF THE HYPAD PROCEDURE

Briefly, the HYPAD (HYbrid PAtterson superposition - Direct methods) procedure can be broken down into the following steps:

(1) calculation of a superposition map, using the program SUPR in the CHES.CAT routine or other suitable program.

(2) GCALC - calculation of the Fourier coefficients of the superposition map.

(3) LOCSYM - location of the symmetry elements in the superposition map using a reciprocal space approach

(3a) MULTIM - location of multiple images, both forward and inverse.

(3b) TRIM - removal of peaks in the superposition map using a symmetry operator and its location in the superposition map.

(4) ORSHIFT - translation of the electron density map origin to the origin of the superposition map.

(5) PHASS - direct transfer of the phases of some of the G_h to the E_h.

(6) EGOR - refinement of the some of the other phases, via the EG relationship.

(7) FOUR - calculation of an electron density map and peak interpretation.

The procedure, except step (1), will be discussed in more detail. In addition, there are several graphics programs which are available and will also be discussed. The input files that must be created by the user and that are not created by the CHES.CAT procedure use keyword-style input. These files and their keyword options are listed in Appendix A. All of the plots in Chapter 5 were generated using these programs.

There are several companion programs which set up the input files for the individual steps. These programs, as well as ORSHIFT and PHASS, make extensive use of the screen managment facility of VAX/VMS. This allows parameters to be entered using a menu-format.

HYPAD

GCALC calculates the Fourier transform coefficients of the superposition map. The program calculates the coefficients by reading in the positions and the heights of the peaks in the superposition map. The "atomic" number of the peaks, or "atom-peaks" is estimated from the height of the origin peak, and the G_h are calculated using an expression similar to equation (2.6). The F_h are transformed into E_h and a Wilson plot is calculated, providing both an approximate least-squares scale factor and an average isotropic thermal parameter value. A listing of

the number and percentage of the total of both G_h and E_h as a function of magnitude are output. This can provide early information as to how "clean" the superposition was. The closer the G_h and the E_h distributions agree, the likelier it is that the superposition map closely resembles the electron density map.

LOCSYM uses the theory presented in Chapter 3 to determine the location and/or existence of symmetry elements in Patterson superposition maps. For symmetry planes, a preliminary test can be made to determine the plausibility of the existence of a symmetry plane. This is most useful when the extinction conditions do not clearly indicate the presence of a symmetry plane.

The translation choices are determined by first calculating $Q(\hat{S},t)$. Then for those translations that appear most promising, $S(\hat{S},t)$ is calculated and the figure-of-merit is used to determine the "best translation choice. The translations for the symmetry elements are then compared for consistency based on the requirements for forming a group (i.e., in class mmm, a 2-fold in x operating on a 2-fold in y must give a 2-fold in z), and those symmetry elements that show consistency are indicated.

For one- and two-dimensional translations, it takes a small amount of computer time to calculate the entire spectrum of translations (within 0.0001 for one dimension, 0.01 for two dimensions). However, it requires a larger

amount of computer time to calculate the three dimensional translation spectrum. One-dimensional operators take about a minute of CPU time, two-dimensional operators take 1-4 minutes, and three-dimensional operators take 5-10 minutes. (Of course, these times depend upon how many data are being using to do the calculations; larger data sets require longer times than smaller ones.) Experience has shown that inclusion of G_h having magnitudes less than unity do not make a qualitative difference in the appearance of the $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions.

In practice, the $S(\hat{S},t)$ function is usually calculated according to equation (3.9). However, the user is given the option of using an alternate weighting function, $|E_h||G_h||G_{h'}|$.

MULTIM uses the theory presented in Chapter 3 to locate translations between multiple images. For centered cells, one of the solutions will be the centering operation. For example, a C-centered cell will produce values of $\xi(t)$ that are approximately unity for the translations (0,0,0) and (½,½,0).

TRIM reduces the list of peaks in the superposition map based on the presence of symmetry related peaks. This is an optional procedure, which I have found is very useful for eliminating extraneous peaks in the map. The symmetry operator to be used and its translation (as given by LOCSYM) are input and the program then searches the peaks list and

eliminates all peaks for which a match cannot be found. The range of peaks to be considered and the tolerance, in Å, for finding a match are input by the user.

After execution of TRIM, GCALC and LOCSYM should be rerun, using the new peaks list. This will provide new phases, which should be better, for use in the rest of the package.

ORSHIFT modifies the phases of the G_h so that they are referenced to the electron density map origin, according to equation (3.2). Once the phases have been modified, the symmetry-related G_h are averaged. For centrosymmetric structures, the phases of the G_h are constrained to be 0 or π.

PHASS sets the phases of some of the larger E_h by directly transferring the phase of the corresponding G_h. The criterion for phase transferability is given by Wang[37]. At this step, the user is given the option of creating a set of Fourier coefficients suitable for use with FOUR (step 7). The coefficients are constructed by combining the magnitude of the E_h with the phase of the G_h. Either the complete set of reflections or the set containing only those reflections for which phases can be transferred can be created. For simple centrosymmetric structures, direct transfer of all of the phases and then calculation of an electron density map can often result in a solution.

EGOR refines the phases of the E_h using the EG relationship. The weighted tangent equation, as given by Wang[37], is used for the refinement. In centrosymmetric space groups, the phases of the E_h are constrained to be 0 or π. Only those E_h with a magnitude greater than or equal to a minimum magnitude, supplied by the user, are refined. The user can set the maximum number of cycles of refinement, but the refinement will automatically terminate when no new phases are added or modified. A phase is added (or modified) when the weight for the new phase is greater than 1. - n/100 (or the current weight), where n is the cycle of iteration.

FOUR is actually a series of programs that calculate and interpret an electron density map. These routines were taken from the CHES.CAT procedure and modified to accept keyword input.

Graphics Programs

There are six graphics programs available in HYPAD. These are called QPLANES, QSPLOT, QROTLINE, QDATA, 2DDRAW, and 3DDRAW. Two of the programs, 2DDRAW and 3DDRAW, are interactive plotting routines that make use of the DI3000 graphics routines.

2DDRAW is a two-dimensional contouring program. The user is given control over both the region to be plotted and

the range of contours. Individual contour lines are drawn every ten percent of the contouring range. Depending upon the device used, the contours are drawn either in color or with different linestyles.

3DDRAW is a contouring program that draws three-dimensional "fishnet" projections. The user has more limited control over the appearance of the plot, and with because of the addition of a third dimension, all contour lines are drawn with the same color and linestyle.

QPLANES generates a "plot" containing the $Q(\hat{S},t)$ values for all space group symmetry planes perpendicular to the axis indicated. All five symmetry plane types (mirror, glides in the other two directions, n-glide, and d-glide) are plotted on the same graph. Plots for more than one axis can be generated in the same run. The "plot" is written out to a file as a series of HP graphics language commands and can be plotted on an HP plotter.

QROTLINE generates a plot of the $Q(\hat{S},t)$ values for a rotation symmetry operator, where one of the translation directions is fixed at some value. As in QPLANES, the plot is written out to a file as a series of HP graphics language commands.

QSPLOT generates a plot containing the $Q(\hat{S},t)$ and $S(\hat{S},t)$ values for a symmetry plane. The $Q(\hat{S},t)$ values are plotted in red and the $S(\hat{S},t)$ values are plotted in green. The calculation of the $S(\hat{S},t)$ values is the most time consuming

part of the program's execution. In order to use enough data so that an accurate representation of the $S(\hat{S},t)$ function is obtained and still use a relatively small amount of CPU time, the fraction of $|G_h|$ used should be the amount having $|G_h| > 1$. Plots for more than one axis can be generated in the same run.

QDATA generates $Q(\hat{S},t)$ or $S(\hat{S},t)$ values for a symmetry operator and writes the data to a file. This is the program that one would use in conjunction with the interactive graphics programs 3DDRAW and 2DDRAW, or some other graphics program (such as PICSURE). For operators that have two translational coordinates (i.e., rotation axes), the values are calculated for a square grid having points every .01 in each direction and ranging from 0.0 to 0.5 inclusive. For operators that have one translational coordinate (i.e., a symmetry plane), the values are calculated for points between 0.0 and 0.5 inclusive, with a step size of .001.

CHAPTER 5. APPLICATION OF HYPAD TO THE SOLUTION OF KNOWN AND UNKNOWN STRUCTURES

Any automated Patterson-based technique should be able to satisfy certain minimum requirements. These requirements should include the ability to solve structures, and applicability to structures of any symmetry type, both centrosymmetric and non-centrosymmetric.

The heavy atom method for the determination of phases is based upon the assumption that the heavy atom(s) present in the structure make up a significant portion of the overall electron density. A general rule of thumb is that if $Z_H^2 > \mu \Sigma Z_L^2$ holds, where Z_H is the atomic number of the heavy atom, Z_L are the atomic numbers of the light atoms, and μ is in the range of ½ to 1; then the heavy atom method is likely to succeed. This condition is typically met by organo-metallic compounds since they usually contain a central metal atom surrounded by an organic ligand. However, there are many organo-metallics that fail to fall into this category. In these cases, phase predictions based solely upon the metal atom are usually erroneous and subsequent electron density maps are of little use. Clearly, these cases require the positions of several atoms be determined and not just the position of the metal atom.

HYPAD Solution of $Cr(CO)_3 \cdot C_6H_6$

This compound represents a very simple organo-metallic, containing one metal atom per molecule. It also meets the criterion for the heavy-atom method easily (μ = 1.1). This makes it a very attractive first test for the reciprocal space methods. The HYPAD solution of this molecule turned out to be extremely straightforward.

An ORTEP drawing of $Cr(CO)_3 \cdot C_6H_6$ (BCT) is given in Figure 5.1. Extinction conditions on the observed intensities failed to uniquely determine the space group, but did indicate the presence of a 2_1-screw axis parallel to **b**. Since the compound crystallizes in the monoclinic system, there are two possible space group choices, $P2_1$ (non-centrosymmetric) and $P2_1/m$ (centrosymmetric).

A vector corresponding to a Cr-C vector was used as the shift vector for a weighted superposition. The Fourier transform coefficients of the superposition map were determined and the $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions were calculated for the mirror plane, the 2_1-screw axis and the inversion operator. Figures 5.2 and 5.3 show plots of the $Q(\hat{S},t)$ function for symmetry planes perpendicular to **b** and the 2_1-screw axis parallel to **b**, respectively. The contour layers in Figure 5.3 are color-coded, with the contours every 10 percent of the range from 0 to the maximum $Q(\hat{S},t)$ value on the map. The colors and their percentage levels

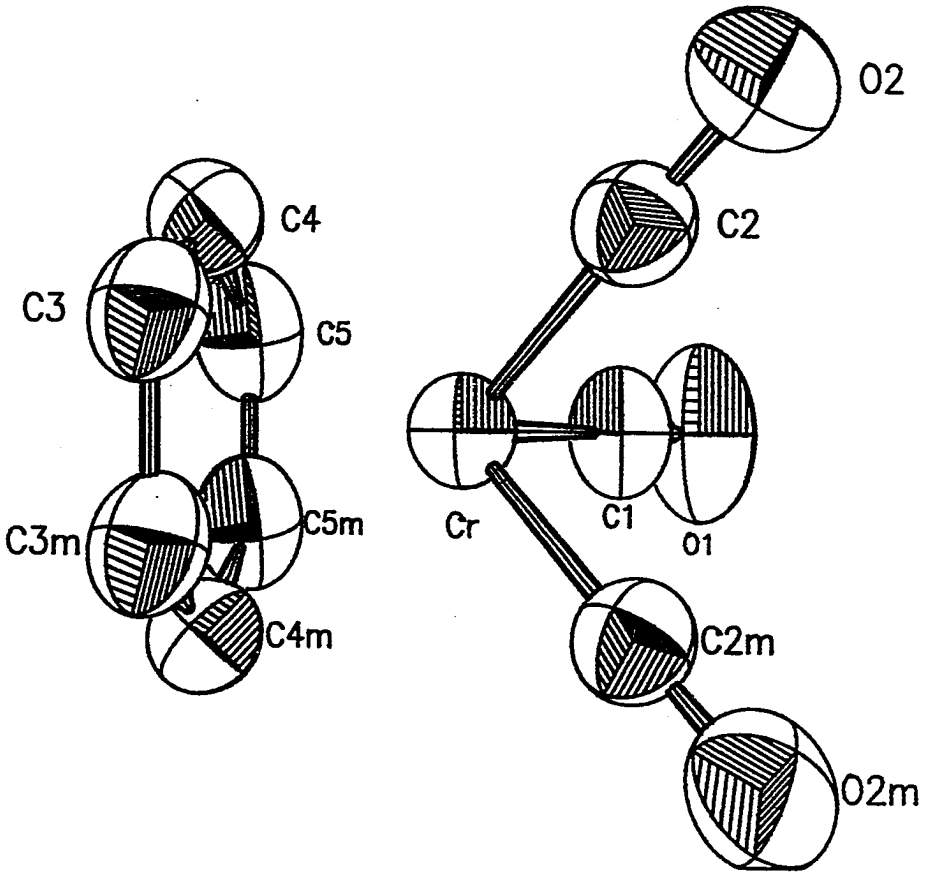


Figure 5.1. ORTEP drawing for $Cr(CO)_3 \cdot C_6H_6$

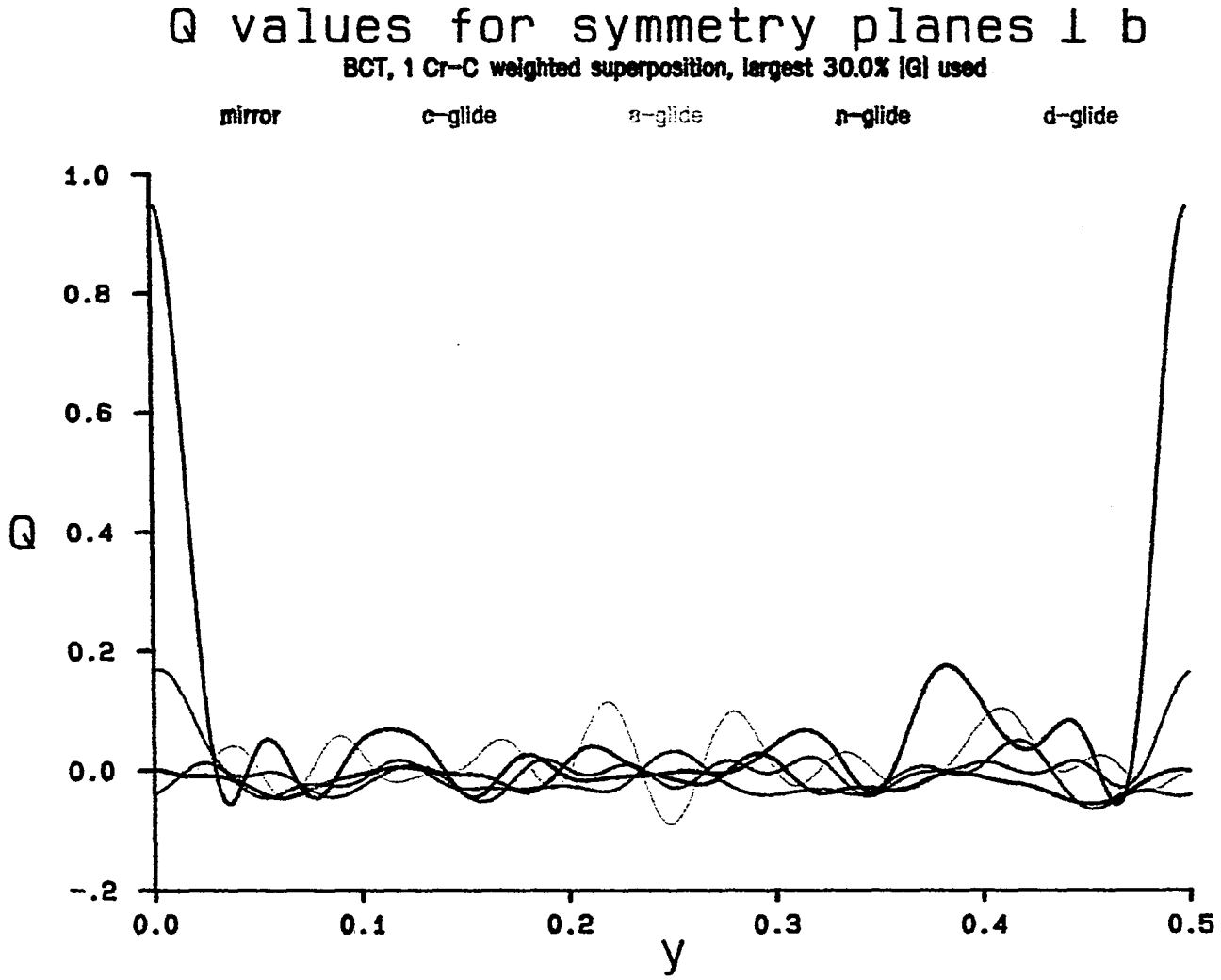


Figure 5.2. $Q(\hat{S},t)$ values for symmetry planes $\perp$ b for BCT

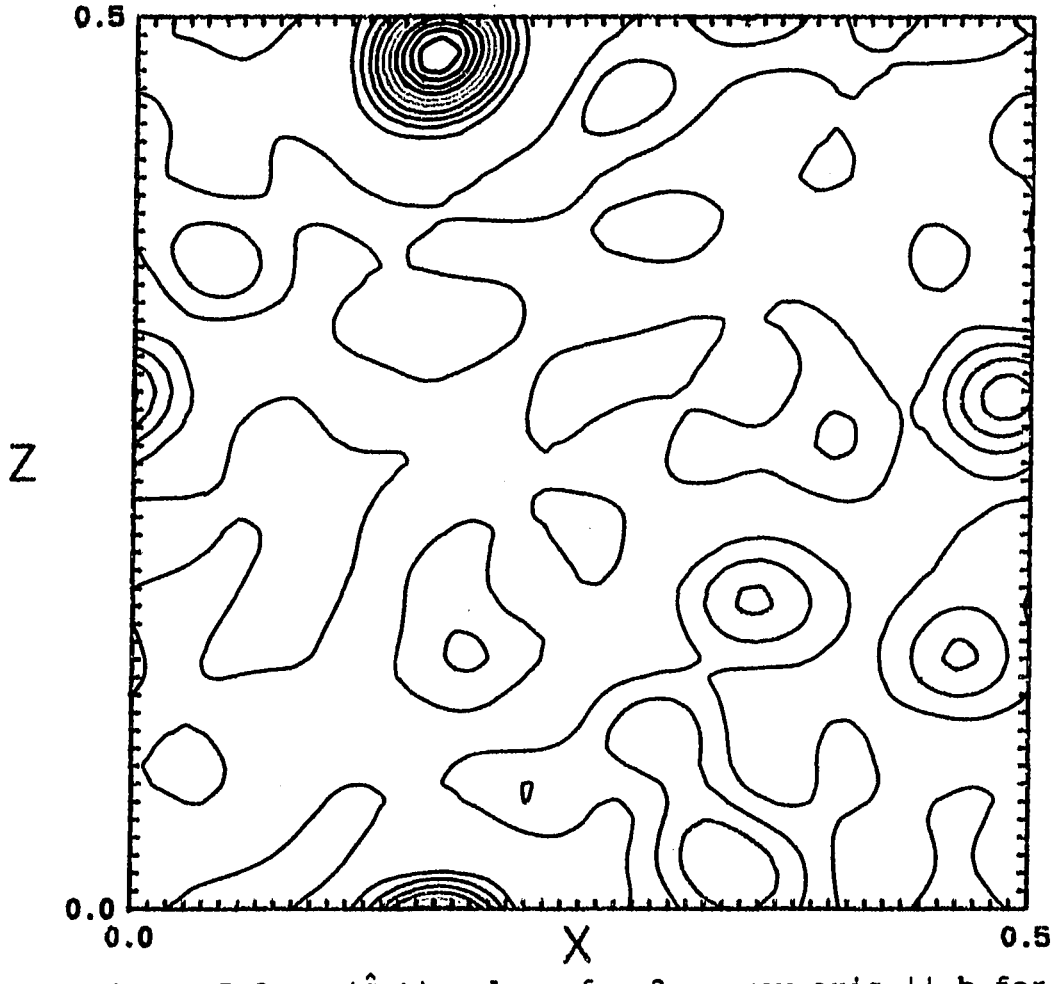


Figure 5.3. $Q(\hat{S},t)$ values for 2_1-screw axis || b for BCT

Table 5.1. Symmetry Operator Locations for $Cr(CO)_3 \cdot C_6H_6$

symmetry operator	t[a]	$Q(\hat{S},t)$	$S(\hat{S},t)$
-x,1/2+y,-z	(0.1681, 0.0, 0.4784)	0.8407	0.0168
x,-y,z	(0.0, 0.0009, 0.0)	0.9544	0.0044
-x,-y,-z	(0.1681, 0.2515, 0.4782)	0.7902	0.0234

[a]Translations are given in fractions of a unit cell.

are: red (0%, 10%), orange (20%, 30%), yellow (40%, 50%), green (60%, 70%), and blue (80%, 90%). Consistent translations, listed in Table 5.1, were found for all three operators, indicating that the space group was $P2_1/m$. It turns out that the Cr atom lies on the mirror plane (as does one of the carbonyls).

Therefore, the large peak at the origin in the $Q(\hat{S},t)$ function of the mirror in Figure 5.2 is from the Cr image. There is a smaller peak at 0.38 which is the location of the mirror plane in the C image. It is interesting to note that the ratio of the values of the $Q(\hat{S},t)$ function at these two points is approximately the value of the weight applied to the superposition. There is only one large peak in the plot for the 2_1-screw axis, and it is the location of the 2_1-screw axis in the Cr image.

Since the origin in the space group lies on the inversion center, the G_h were transformed using the inversion center location, according to equation (3.2). The

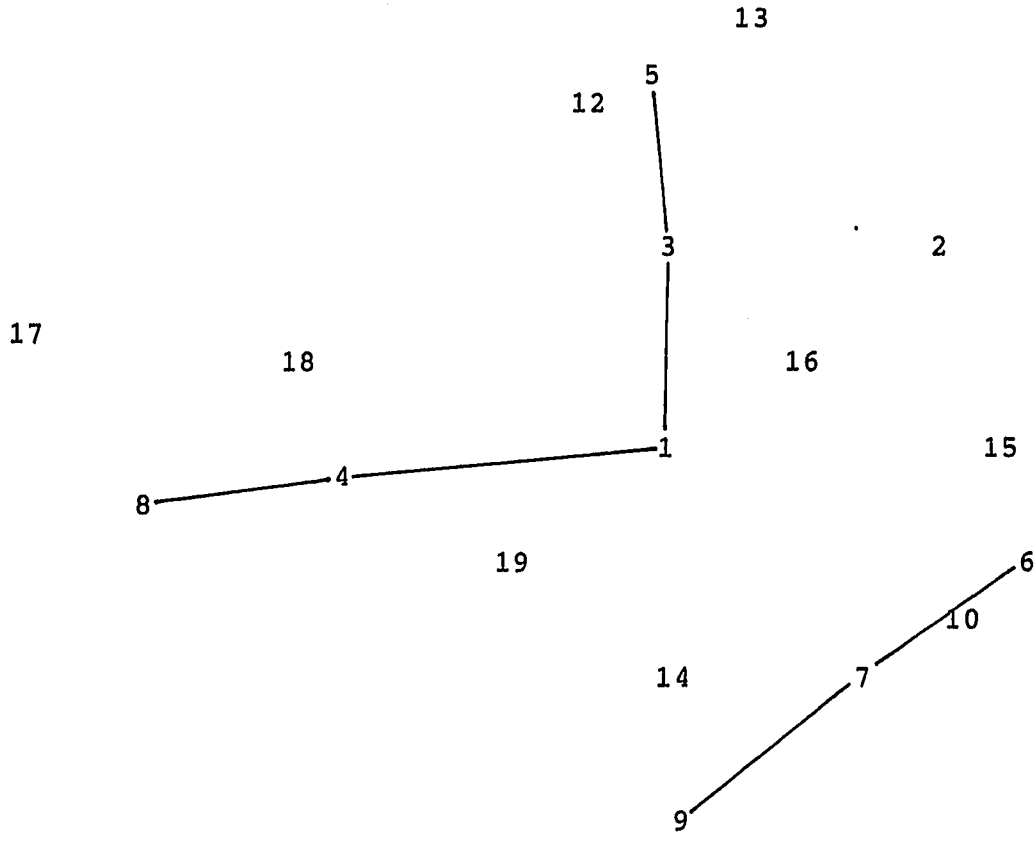


Figure 5.4. HYPAD-generated projection of $Cr(CO)_3 \cdot C_6H_6$

Table 5.2. HYPAD Atomic Coordinates[a] ($x10^4$) for $Cr(CO)_3 \cdot C_6H_6$

atom	#	pk.ht.	x	y	z
Cr	1	505	1692	2500	4755
O1	8	90	-1971	2500	1050
O2	5	96	4140	642	2700
C1	4	108	-520	2500	2398
C2	3	111	3131	1396	3572
C3	6	96	3234	1854	7962
C4	7	95	1281	1260	7270
C5	9	81	-739	1852	6534

[a]Atomic coordinates are given as fractions of the unit cell.

Table 5.3. Refined Atomic Coordinates[a] ($x10^4$) for $Cr(CO)_3 \cdot C_6H_6$

atom	x	y	z
Cr	1682(1)[b]	2500	4755(1)
O1	-1917(7)	2500	1005(6)
O2	4094(6)	668(3)	2744(4)
C1	-523(9)	2500	2446(7)
C2	3167(6)	1359(3)	3548(5)
C3	3202(6)	1868(3)	7967(5)
C4	1215(7)	1231(3)	7271(5)
C5	-753(6)	1858(4)	6585(6)

[a]Atomic coordinates are given as fractions of the unit cell.

[b]Estimated standard deviations for the refined coordinates are given in parentheses for the least significant digit.

Table 5.4. Comparative bond distances for $Cr(CO)_3 \cdot C_6H_6$

atoms	refined(Å)	HYPAD(Å)	Δ(Å)
Cr - C1	1.831(5)[a]	1.85	0.02
Cr - C2	1.833(5)	1.78	-.05
Cr - C3	2.231(3)	2.25	0.02
Cr - C4	2.218(4)	2.20	-.02
Cr - C5	2.207(4)	2.20	-.01
C1 - O1	1.144(6)	1.13	-.01
C2 - O2	1.143(4)	1.25	0.09
C3 - C4	1.405(5)	1.37	-.03
C4 - C5	1.389(5)	1.40	0.01

[a]Estimated standard deviations for the refined distances are given in parentheses for the least significant digit.

Table 5.5. Comparative bond angles for $Cr(CO)_3 \cdot C_6H_6$

atoms	refined(°)	HYPAD(°)	Δ(°)
C2 - Cr - C1	89.07(16)[a]	89.0	-.1
C2 - Cr - C2m	86.75(21)	86.7	0.0
Cr - C1 - O1	179.26(46)	175.2	-3.9
Cr - C2 - O2	179.30(30)	178.4	-.9
C3 - C4 - C5	120.09(34)	123.5	3.4
C4 - C3 - C3m	120.03(21)	118.7	-1.3
C4 - C5 - C5m	119.87(23)	117.8	-2.1

[a]Estimated standard deviations for the refined angles are given in parentheses for the least significant digits.

phases of 75 of the largest E_h, representing 11.4% of the total number of E_h, were set by directly assigning the phases of the corresponding G_h. The E_h were then refined using the EG relationship and an electron density map was calculated. Figure 5.4 shows a projection of the electron density map onto the least-squares plane as generated by HYPAD. (The bonds in this projection were drawn in by hand.) The complete molecule is clearly identifiable from this picture. The HYPAD atomic positions, along with the peak number and peak heights are given in Table 5.2 and the refined atomic positions are given in Table 5.3. All of the unique non-hydrogen atoms are found in the top 9 peaks on the map. Tables 5.4 and 5.5 compare the refined and HYPAD-generated bond distances and bond angles for this compound. All of the bond distances in the HYPAD molecule are within 0.1Å of the refined distances and the HYPAD bond angles are within 4° of the refined angles.

It is encouraging to note that even with incomplete initial knowledge of the space group, this structure can be solved with very little user input, and that the correct space group is obtained during the procedure.

HYPAD Solution of $FeP_2OC_{32}H_{31}I$

This is a significantly more complicated compound, and was the first acentric structure studied using HYPAD.

Crystals of the compound were obtained from John Nelson's group (Department of Chemistry, University of Nevada-Reno). It was originally solved by another member of our group, Cathy Day, only after a very tedious Patterson analysis. As shown in Figure 5.5, $FeP_2OC_{32}H_{31}I$ (FEI) is a very complex molecule, the type that it would be hoped that HYPAD would be useful in solving.

This compound crystallizes in the orthorhombic system. The extinction conditions failed to uniquely determine the space group, but did indicate the presence of a c-glide plane perpendicular to a and a n-glide plane perpendicular to c. This leaves two possible space group choices: $Pc2_1n$ (non-centrosymmetric) and Pcmn (centrosymmetric). Direct methods failed to provide a solution in either of the possible space groups.

A vector corresponding to a Fe-P bond distance was chosen as the shift vector for a weighted superposition. The Fourier transform coefficients were determined and the $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions for all of the possible symmetry elements in the two space group choices were calculated. The best translations, and their corresponding $Q(\hat{S},t)$ and $S(\hat{S},t)$ values are listed in Table 5.6. Consistent translations were found for the c-glide plane perpendicular to a, the 2_1-screw axis parallel to b and the n-glide plane perpendicular to c, indicating that the space group was $Pc2_1n$. Plots of the $Q(\hat{S},t)$ function for symmetry planes

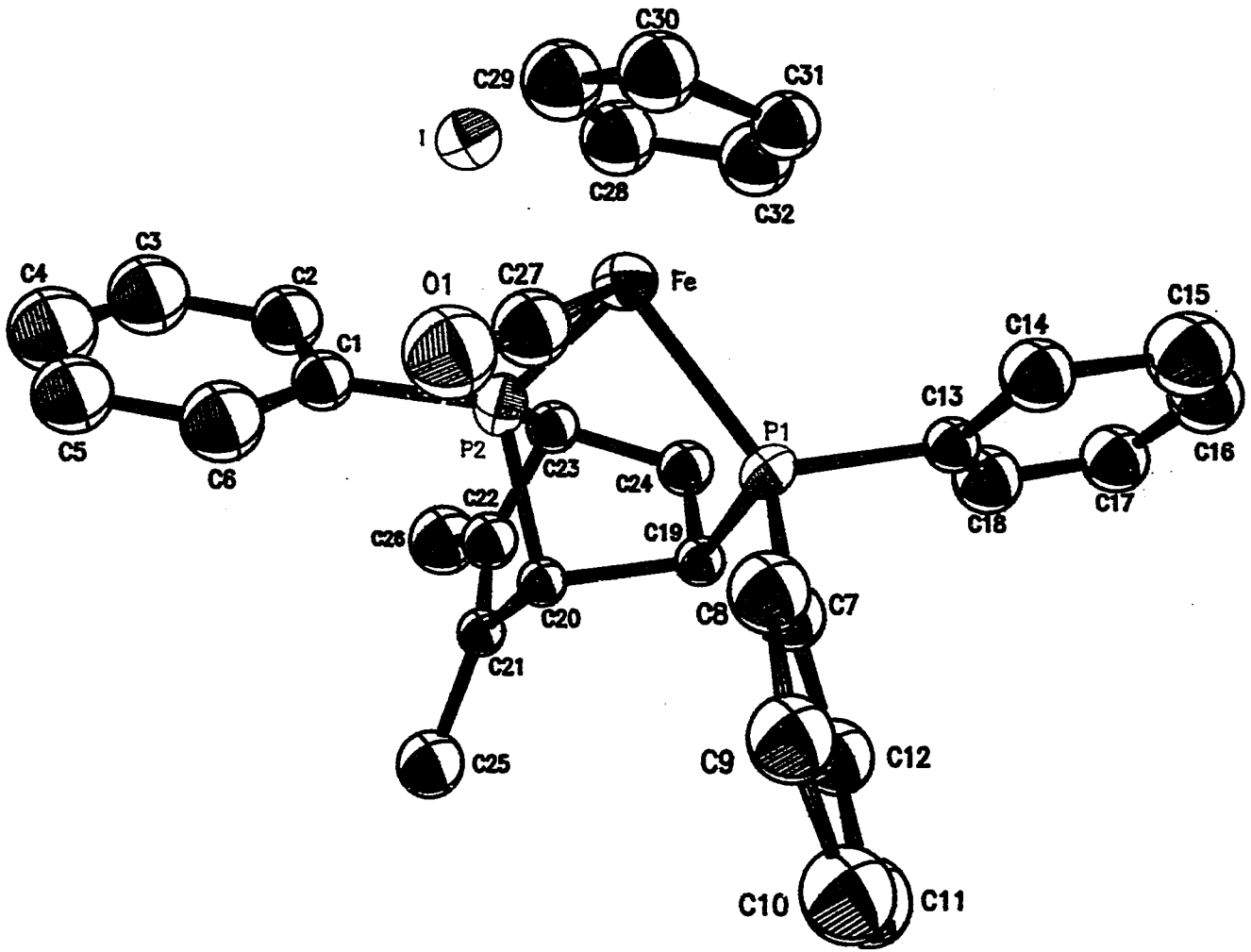


Figure 5.5. ORTEP drawing for $FeP_2OC_{32}H_{31}I$

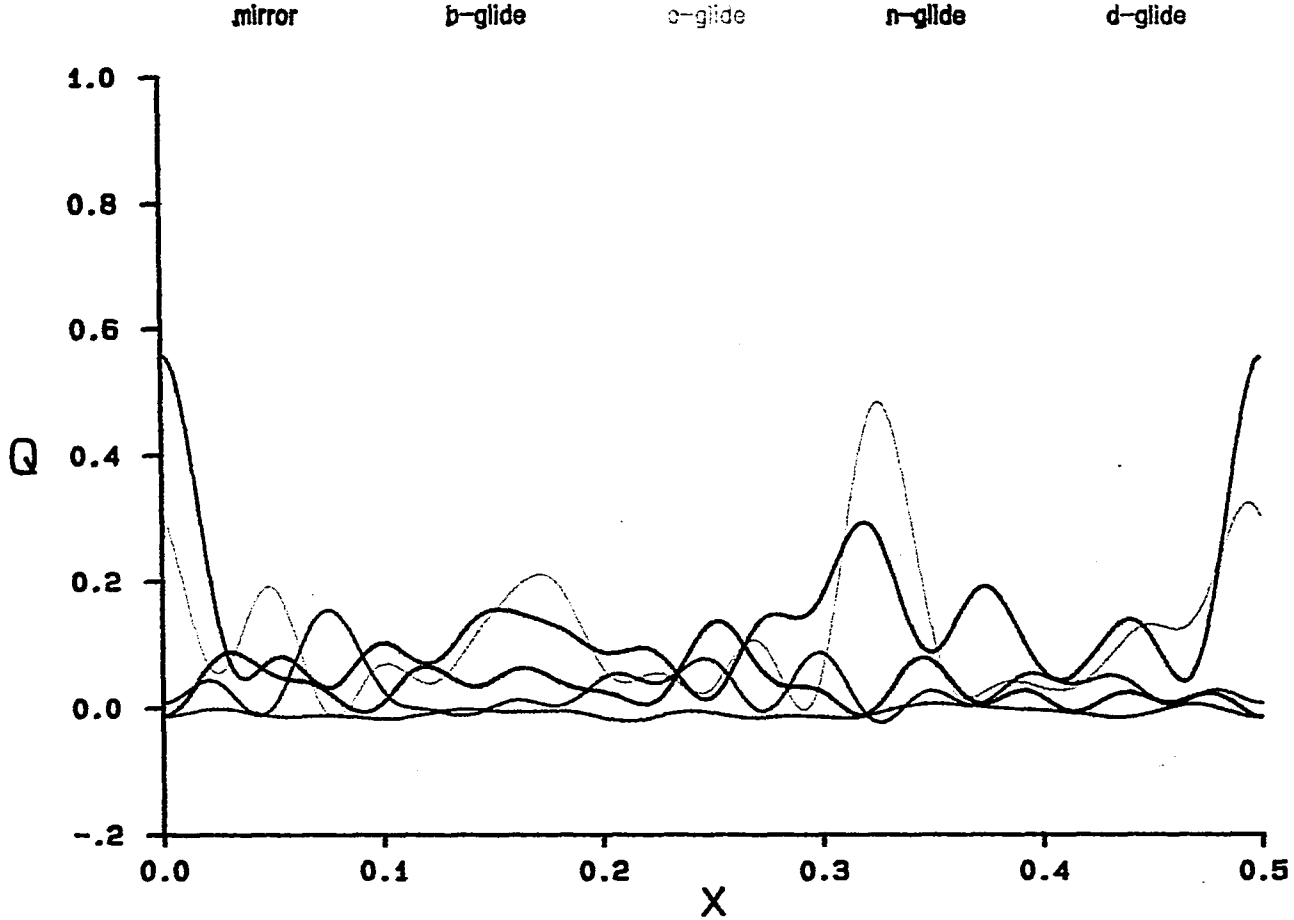


Figure 5.6. $Q(\hat{S},t)$ values for symmetry planes $\perp$ a for FEI

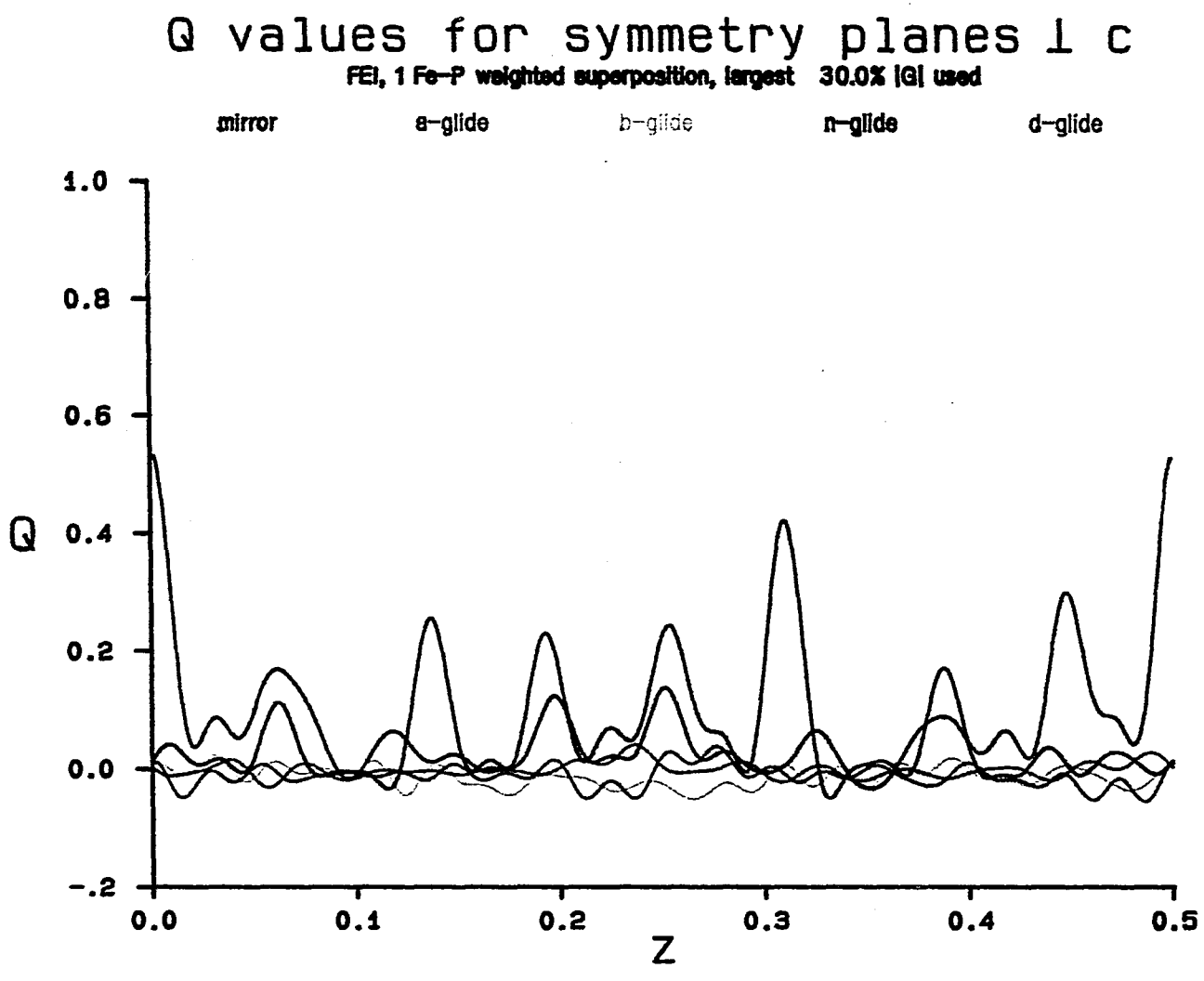


Figure 5.7. $Q(\hat{S},t)$ values for symmetry planes $\perp$ c for FEI

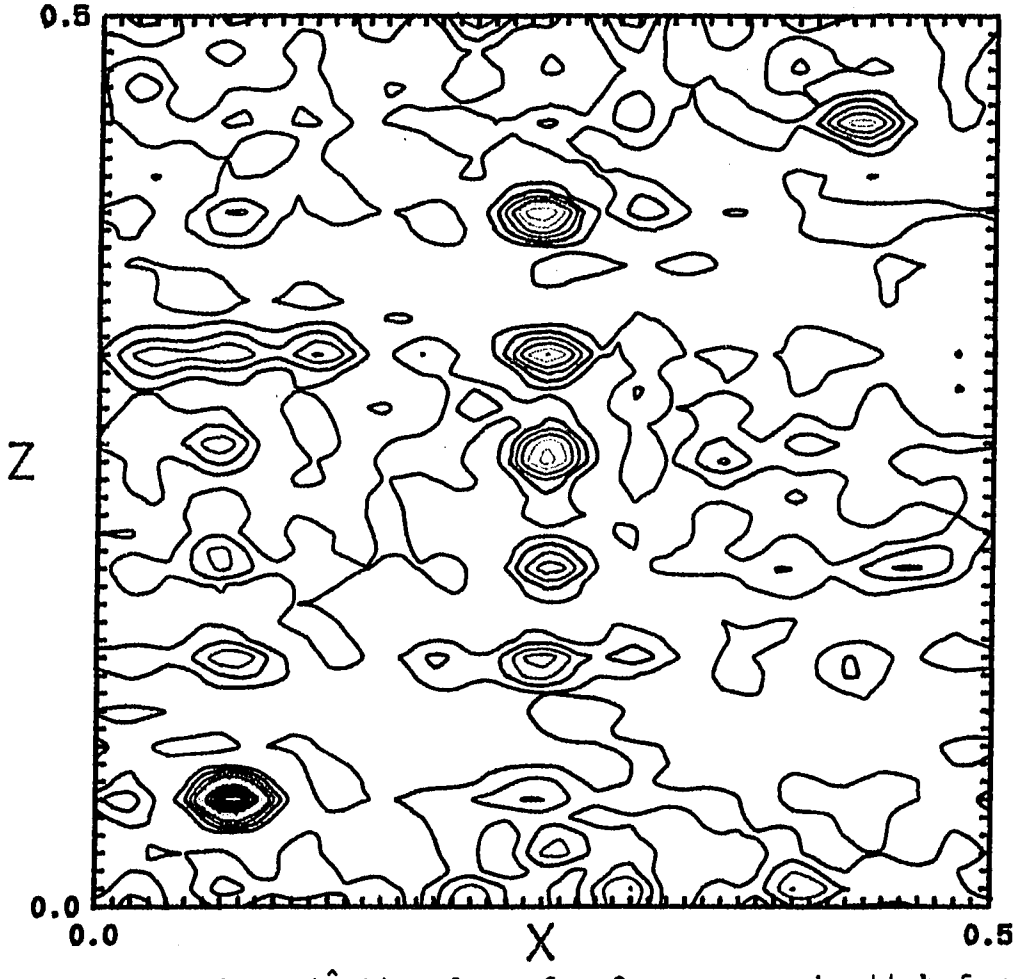


Figure 5.8. $Q(\hat{S},t)$ values for 2_1-screw axis || b for FEI

Table 5.6. Symmetry Operator Locations for $FeP_2OC_{32}H_{31}I$

symmetry operator	t[a]	$Q(\hat{S},t)$	$S(\hat{S},t)$
x+1/2,-y,-z	(0.0, 0.2548, 0.3106)	0.3273	0.0888
-x,y+1/2,-z	(0.0762, 0.0, 0.0604)	0.4191	0.0761
-x,-y,z+1/2	(0.3287, 0.0046, 0.0)	0.3467	0.0874
-x,y,z+1/2	(0.3260, 0.0, 0.0)	0.4967	0.0703
x,-y,z	(0.0, 0.1885, 0.0)	0.3130	0.0919
x+1/2,y+1/2,-z	(0.0, 0.0, 0.3104)	0.4274	0.0749
-x,-y,-z	(0.4100, 0.4845, 0.4745)	0.8301	0.0203

[a]Translations are given in fractions of a unit cell.

perpendicular to a and c are shown in Figures 5.6 and 5.7, respectively, and for the 2_1-screw axis parallel to b in Figure 5.8.

There is a sizable peak in the $Q(\hat{S},t)$ function for the c-glide plane at approximately 0.32, which corresponds to the location of the glide plane in the Fe image. It turns out that the P atom used in the superposition vector lies almost on the glide plane. Thus, the peak at the origin of the $Q(\hat{S},t)$ function indicates the location of the glide plane in the P image. As in the case of $Cr(CO)_3 \cdot C_6H_6$, the ratio of the $Q(\hat{S},t)$ values for these peaks is close to the weight used in the superposition.

The $Q(\hat{S},t)$ function for the n-glide plane perpendicular to c has several peaks, the largest of which, at approximately 0.31, is the location of the glide plane in the Fe image. The next highest peak is the location of the

glide plane in the P image. Again, the ratio of the $Q(\hat{S},t)$ values for these two peaks is approximately the weight used for the superposition.

There is one peak in the $Q(\hat{S},t)$ function for the 2_1-screw axis parallel to b, at approximately x = 0.07, z = 0.06, which is obviously much higher than the other peaks. This is the location of the screw axis in the Fe image. The next highest peak, at approximately x = 0.25, z = 0.39, is location of the 2_1-screw axis in the P image.

Another interesting feature in this map is the line of peaks having x = 0.25. If this line of the 2_1-screw axis map is plotted on the same diagram as the n-glide plane perpendicular to c, as in Figure 5.9, an interesting feature is observed. There is a very strong correlation between the peaks in the two curves. Essentially every place that there is a peak in one curve, there is a peak in the other curve.

As can be seen from the $Q(\hat{S},t)$ functions for the symmetry planes, there is still quite a bit of pseudo-symmetry remaining from the Patterson map in the superposition map. This pseudo-symmetry can interact with the actual symmetry of the unit cell to produce still more pseudo-symmetry. However, the general location of this "interaction" pseudo-symmetry can be predicted.

Let us assume that there is still pseudo-symmetry remaining from the Patterson map in the form of a mirror plane perpendicular to a. In the Patterson map, this mirror

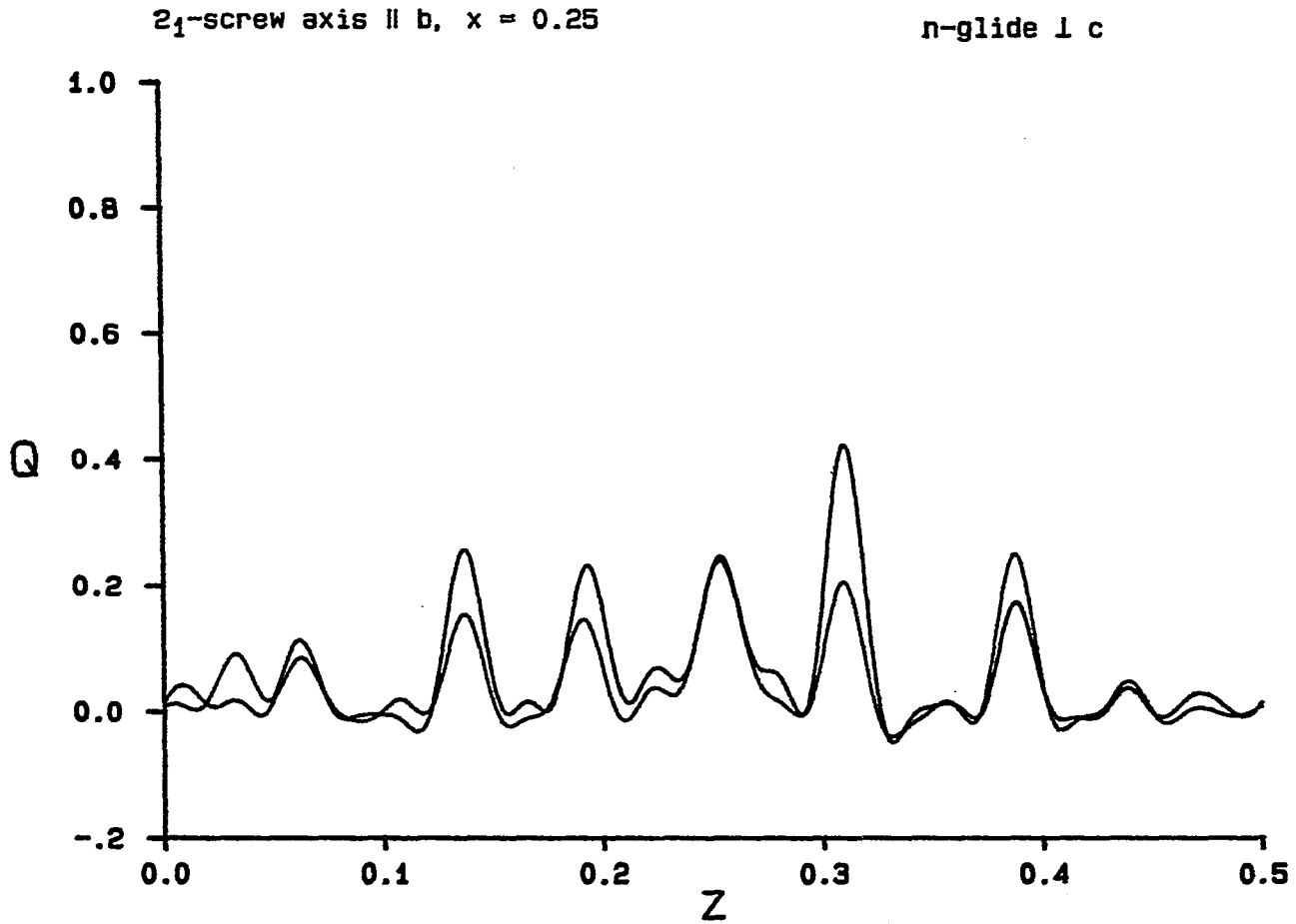


Figure 5.9. $Q(\hat{S},t)$ values for n-glide plane $\perp$ c and 2_1-screw axis || b for FEI

plane lies at x = 0, and so a point at x,y,z in the superposition map generates at pseudo-point at -x,y,z. Because of the n-glide plane perpendicular to c, located at t_z, a peak at x,y,z will have a symmetry-related peak at $1/2+x,1/2+y,2t_z-z$. This symmetry-related peak will generate, by the pseudo-mirror, a pseudo-peak at $1/2-x,1/2+y,2t_z-z$. The peak at x,y,z is "related" to the pseudo-peak at $1/2-x,1/2+y,2t_z-z$ by a pseudo-2_1-screw axis parallel to b. This pseudo-2_1-screw axis is located at (0.25, 0.0, t_z). Thus, the two peaks in the $Q(\hat{S},t)$ function at (0.25, 0.0, 0.14) and (0.25, 0.0, 0.31) are the locations of pseudo-symmetry elements caused by the n-glide plane in the P image and the Fe image, respectively. The n-glides should also produce pseudo-2_1-screw axes parallel to a at (0.0, 0.25, 0.31) and (0.0, 0.25, 0.14). Examination of Table 5.6 shows that the "best" translation for the 2_1-screw axis is at (0.0, 0.255, 0.311). There is also a peak at (0.0, 0.253, 0.138).

Using similar reasoning, the c-glide plane perpendicular to a will produce pseudo-2_1-screw axes parallel to c at (t_x, 0.0, 0.0). According to Table 5.6, the c-glide plane is located at (0.326, 0.0, 0.0). This should produce a pseudo-2_1-screw axis at approximately (0.326, 0.0, 0.0) and examination of Table 5.6 shows that the "best" translation for the 2_1-screw axis is at (0.329, 0.005, 0.0). The c-glide plane in the P image is located at (0.0, 0.0, 0.0)

which should produce a pseudo-2_1-screw axis at (0.0, 0.0, 0.0). There is a peak in the $Q(\hat{S},t)$ function for the 2_1-screw axis at (-.003, 0.003, 0.0).

In certain cases, rotation (and screw) axes can produce pseudo-glide planes. A point at x,y,z, operated on by a 2_1-screw axis parallel to b located at (0.25, 0.0, t_z), will produce a point at 1/2-x,1/2+y,2t_z-z. Assuming a mirror perpendicular to a remains from the Patterson, a pseudo-peak at 1/2+x,1/2+y,2t_z-z is generated which is "related" to the original peak at x,y,z by a pseudo-n-glide perpendicular to c at (0.0, 0.0, t_z). Since the location of the 2_1-screw axis parallel to b in the P image is located at (0.25, 0.0, 0.39), the peak at (0.0, 0.0, 0.388) in the $Q(\hat{S},t)$ function for the n-glide is the result of the interaction of pseudo-symmetry with the 2_1-screw axis in the P image.

Since there is apparently a large amount of pseudo-symmetry remaining in the superposition map, it is reasonable to expect that the G_h would not be a good approximation to the E_h. When the G_h are transformed and the phases of the some of the largest E_h assigned and then refined using the EG relationship, the electron density map produced is uninterpretable. However, the fact that the locations of the symmetry elements are known was used to determine the locations of the symmetry-related P atoms. The vectors between the Fe atom and the four symmetry-related P atoms were then used in a multiple weighted

superposition. The list of peaks in the superposition map was reduced using the program TRIM. The 400 highest peaks in the list were checked for n-glide-related peaks in the list (using a tolerance for matching of 0.6Å), then the reduced list was checked for c-glide-related peaks, and the new reduced list so obtained was checked for 2_1-screw-related peaks. The Fourier transform coefficients of this new superposition map were determined and the $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions for the three symmetry operators, c-glide perpendicular to a, n-glide perpendicular to c, and 2_1-screw axis parallel to b, were calculated. Plots of the $Q(\hat{S},t)$ function for symmetry planes perpendicular to a and c and the 2_1-screw axis parallel to b are shown in Figure 5.10, Figure 5.11 and Figure 5.12, respectively.

As can be seen in these plots, the amount of pseudo-symmetry in the superposition map has been significantly reduced. The G_h were transformed according to equation (3.2) and the phases of some of the largest E_h assigned by direct transfer. The phases of those E_h with $|E_h| > 1.2$ were refined using the EG relation and an electron density map calculated. Figure 5.13 shows the projection of the electron density map onto the least-squares plane. A sizable part of the molecule can be determined by examination. The HYPAD atomic positions, along with the peak number and peak heights are given in Table 5.7, and the refined atomic positions are given in Table 5.8. A

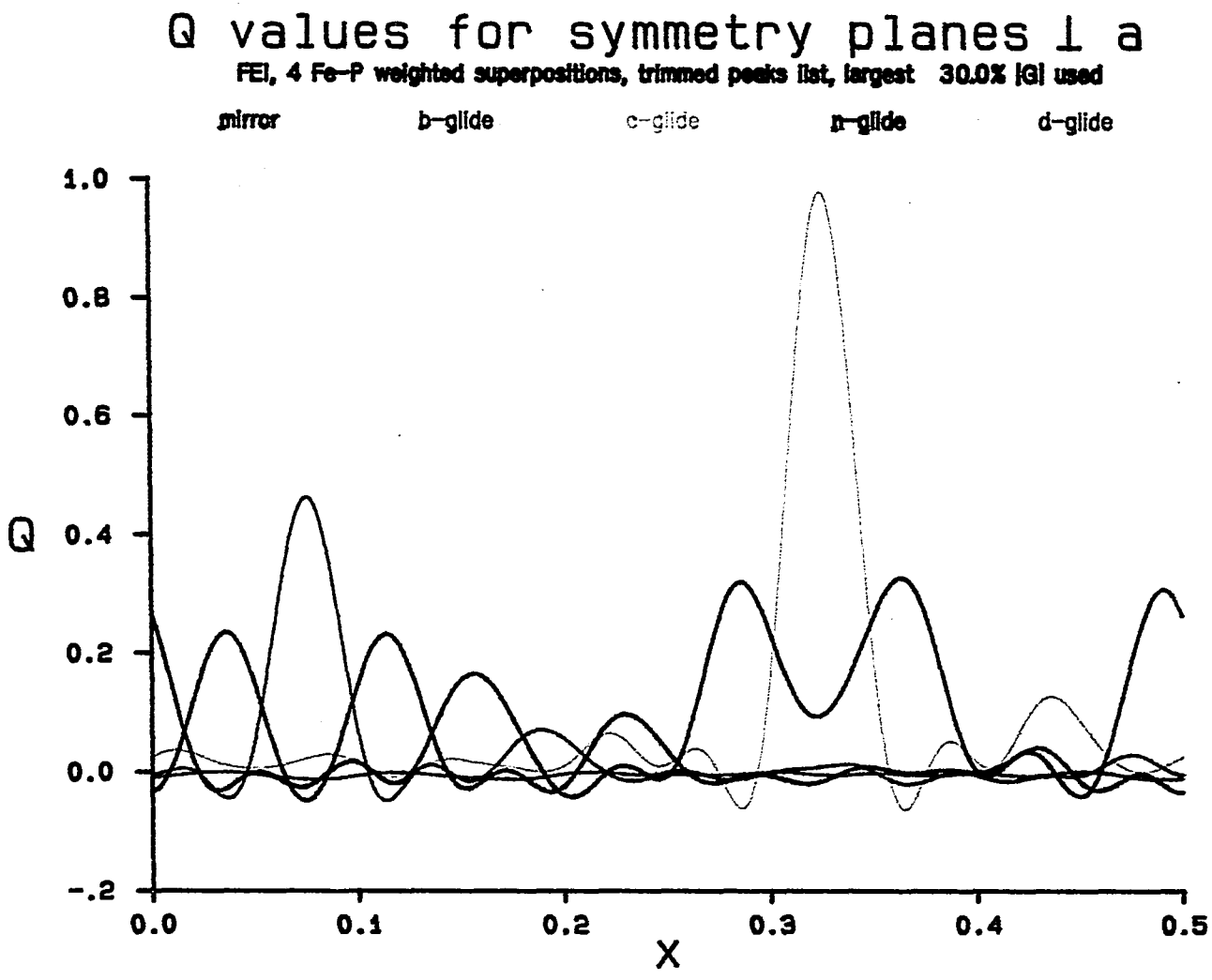


Figure 5.10. $Q(\hat{S},t)$ values for symmetry planes $\perp$ a for FEI

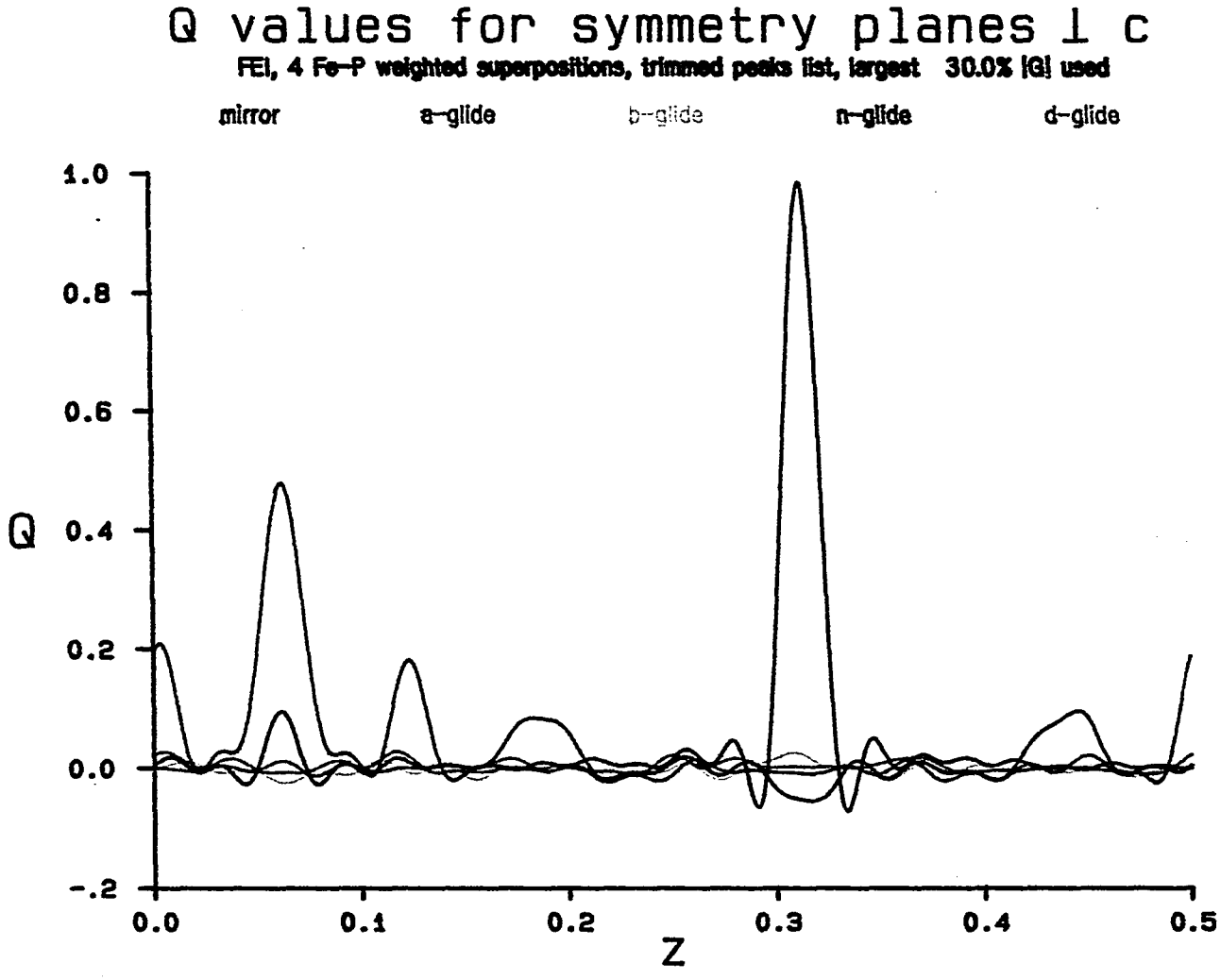


Figure 5.11. $Q(\hat{S},t)$ values for symmetry planes $\perp$ c for FEI

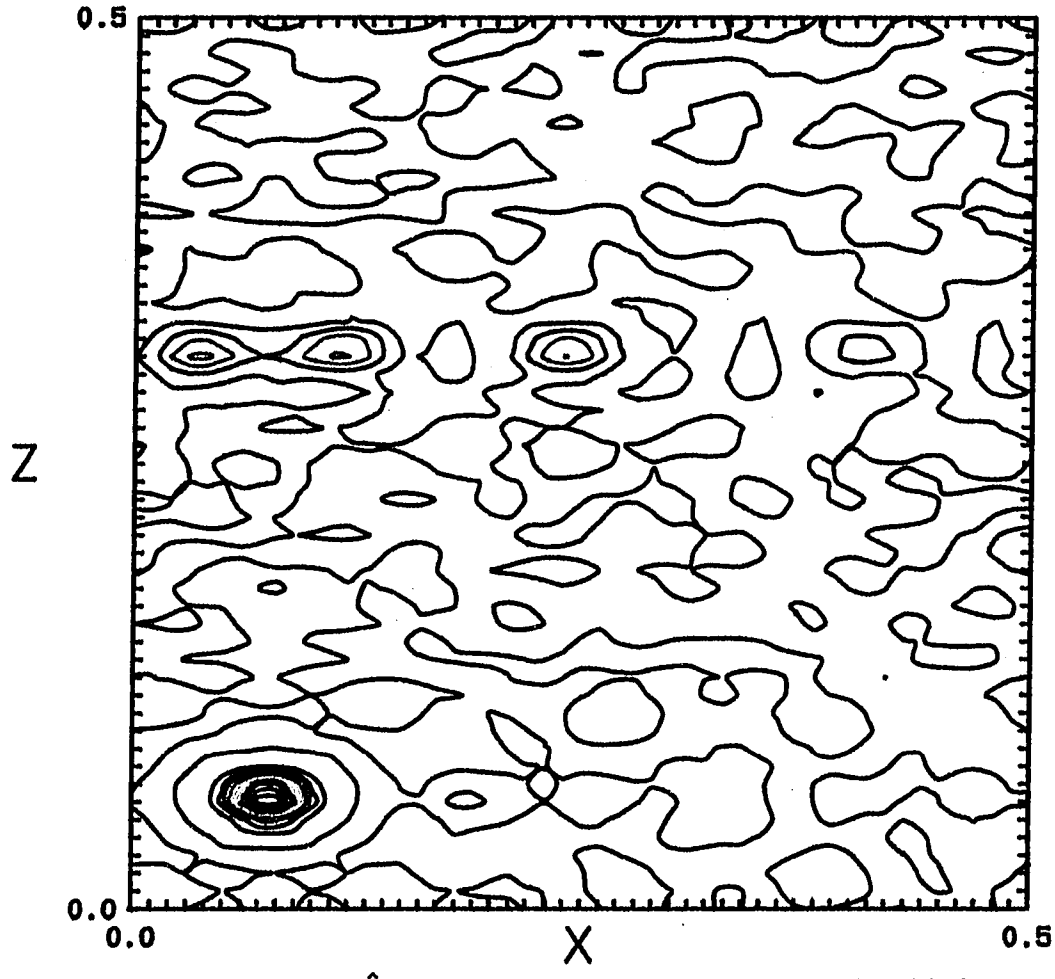


Figure 5.12. $Q(\hat{S},t)$ values for 2_1-screw axis || b for FEI

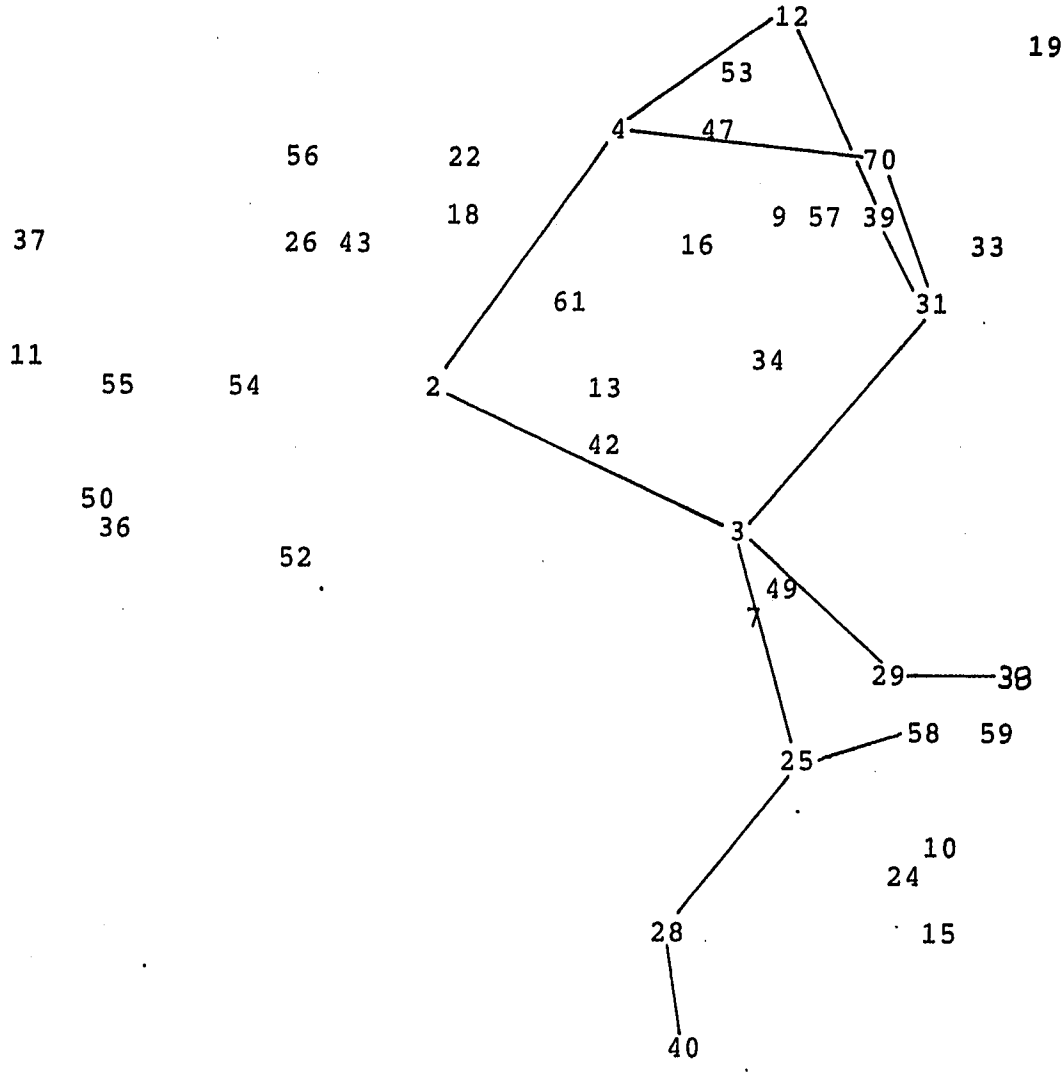


Figure 5.13. HYPAD-generated projection of $FeP_2OC_{32}H_{31}I$

Table 5.7. HYPAD Atomic Coordinates[a] ($x10^4$) for $FeP_2OC_{32}H_{31}I$

atom	#	pk.ht.	x	y	z
I	1	359	2072	3141	5015
Fe	2	144	-4223	5008	4399
P1	3	87	-2892	3941	4695
P2	4	77	-2460	5425	3858
C23	12	36	-914	5764	4388
C13	25	23	-2838	3459	5748
C14	28	23	-4028	3159	5997
C7	29	23	-2867	2847	4228
C19	31	22	-1029	4279	4411
C8	38	20	5846	2619	3911
C24	39	20	-785	5134	4995
C15	40	20	-4078	2978	6688
C18	58	16	-1825	3587	6194
C20	70	15	-1155	4552	3634

[a]Atomic coordinates are given as fractions of the unit cell.

Table 5.8. Refined Atomic Coordinates[a] ($x10^4$) for $FeP_2OC_{32}H_{31}I$

atom	x	y	z
I	2059(1)[b]	3175(2)	5058(1)
Fe	-4235(3)	5000	4429(2)
P1	-2878(6)	3965(4)	4714(3)
P2	-2467(6)	5440(4)	3853(3)
C23	-1022(21)	5838(12)	4369(12)
C13	-2829(22)	3608(12)	5658(12)
C14	-3993(21)	3306(16)	5917(12)
C7	-3007(22)	3011(13)	4231(10)
C19	-1177(20)	4312(12)	4506(11)
C8	-4187(23)	2774(14)	3920(12)
C24	-855(18)	513(12)	4903(11)
C15	-3975(25)	2998(15)	6695(13)
C18	-1643(23)	3717(15)	6113(14)
C20	-1272(20)	4630(12)	3651(11)

[a]Atomic coordinates are given as fractions of the unit cell.

[b]Estimated standard deviations for the refined coordinates are given in parentheses for the least significant digit.

Table 5.9. Comparative bond distances for $FeP_2OC_{32}H_{31}I$

atoms	refined(Å)	HYPAD(Å)	Δ(Å)
Fe - P1	2.240(7)[a]	2.26	0.02
Fe - P2	2.173(7)	2.13	-.04
P1 - C7	1.815(22)	2.00	0.19
P1 - C13	1.876(22)	2.09	0.21
P1 - C19	1.821(21)	2.00	0.18
P2 - C20	1.826(21)	1.98	0.15
P2 - C23	1.839(22)	1.90	0.06
C7 - C8	1.358(31)	1.45	0.09
C13 - C14	1.327(30)	1.36	0.03
C13 - C18	1.425(32)	1.31	-.12
C14 - C15	1.514(33)	1.30	-.21
C19 - C20	1.656(29)	1.50	-.16
C19 - C24	1.582(28)	1.79	0.21

[a]Estimated standard deviations for the refined distances are given in parentheses for the least significant digit.

comparative table of bond distances can be found in Table 5.9.

A least-squares refinement of the atom positions, with fixed isotropic thermal parameters, obtained from HYPAD resulted in a residual index (R factor) of 16.8%. This represents an excellent start! The remainder of the molecule, except hydrogen atoms, was found in the first electron density map calculated from these results.

These results show a that a good initial model for the structure of a fairly complex compound can be derived from the superposition map. The bond distances in the HYPAD-gernerated fragment are within 0.21Å of the refined bond distances. The errors in the distances of bonds containing C atoms are less than ten times the standard deviation of the refined distances. Considering the fact that the nominal resolution in the map was on the order of 0.26 - 0.31Å and that the only small portion of the E_h were used in generating the electron density map, this result is quite satisfactory. These results also illustrate one of the major difficulties of superposition techniques, that of pseudo-symmetry. However, even with the presence of a large amount of pseudo-symmetry, HYPAD was able to determine the correct symmetry and the locations of the symmetry elements. Although, in this case, the single-superposition map does not give a good phase set for the E_h, a method for obtaining a good phase set does present itself. Use of the locations

of the symmetry operators to find more vectors to use for additional superpositions is a trivial procedure, and, as this case shows, it can have tremendous rewards.

HYPAD Solution of $Ni(C_6O_3H_9)_2Cl_2$

This compound (NIT) - shown in Figure 5.14 - is a slightly more complex organo-metallic compound than $Cr(CO)_3 \cdot C_6H_6$. It was originally solved using the "heavy-atom" method by Bill Jenson's group (Department of Chemistry, South Dakota State University). The HYPAD solution of this compound turned out to be very routine.

This compound crystallizes on the monoclinic crystal system. The extinction conditions on the intensities uniquely indicated that the space group was the centrosymmetric choice $P2_1/c$. A vector corresponding to a Ni-O vector was chosen as the shift vector and an unweighted superposition was done. The Fourier transform coefficients of the superposition map were determined and the $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions were calculated for the 2_1-screw axis parallel to **b**, the c-glide plane perpendicular to **b** and the inversion operator. Plots of the $Q(\hat{S},t)$ function for the 2_1-screw axis and symmetry planes perpendicular to **b** are shown in Figures 5.16 and 5.15, respectively. Table 5.10 lists the translations for the operators, as well as the

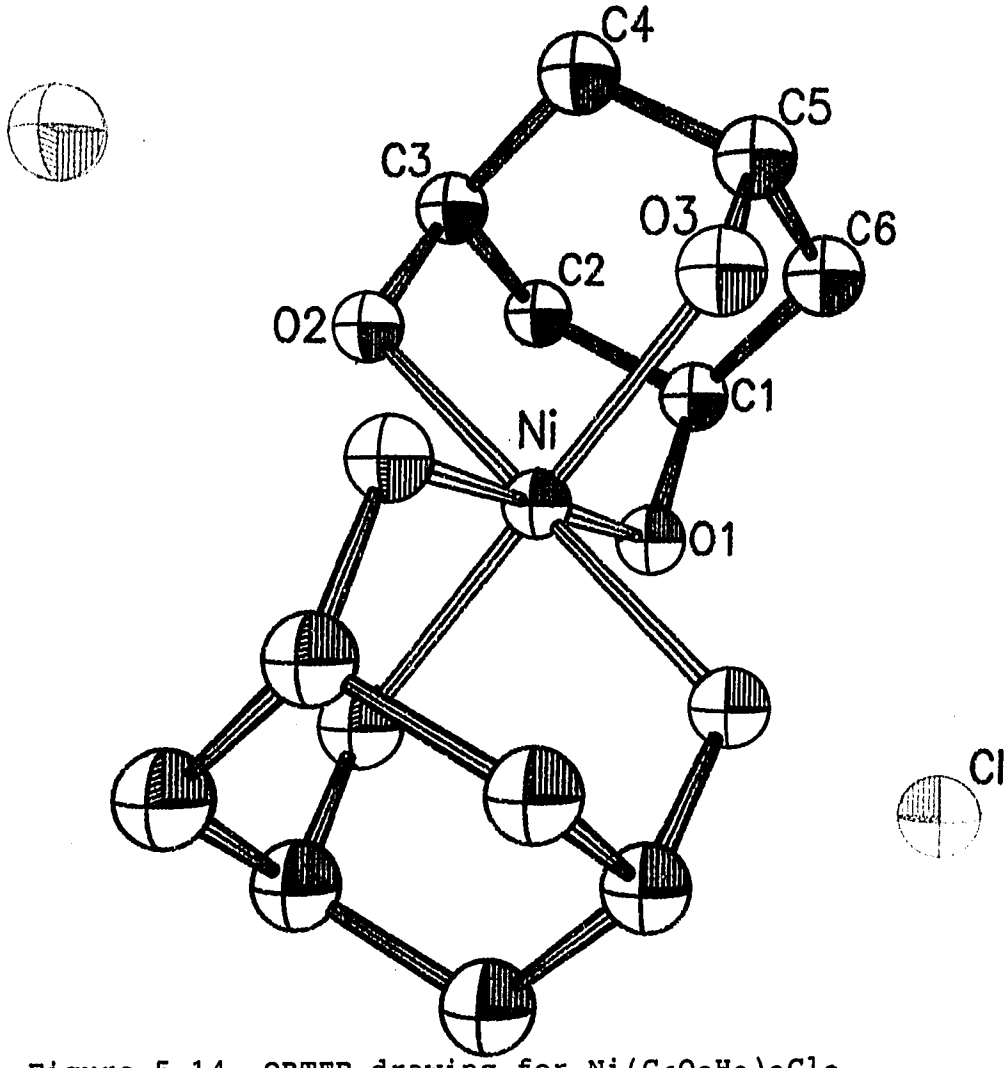


Figure 5.14. ORTEP drawing for $Ni(C_6O_3H_9)_2Cl_2$

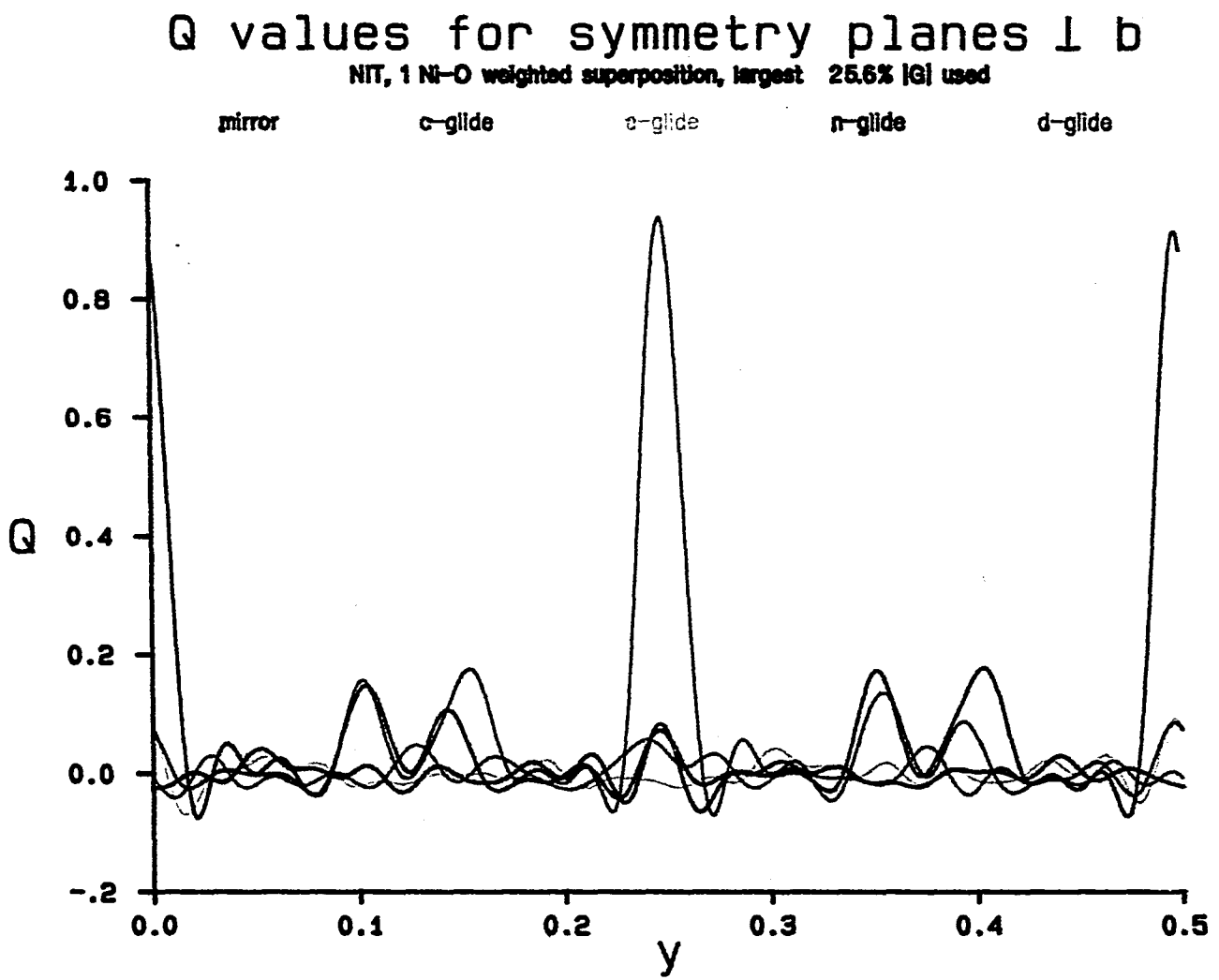


Figure 5.15. $Q(\hat{S},t)$ values for symmetry planes $\perp$ b for NIT

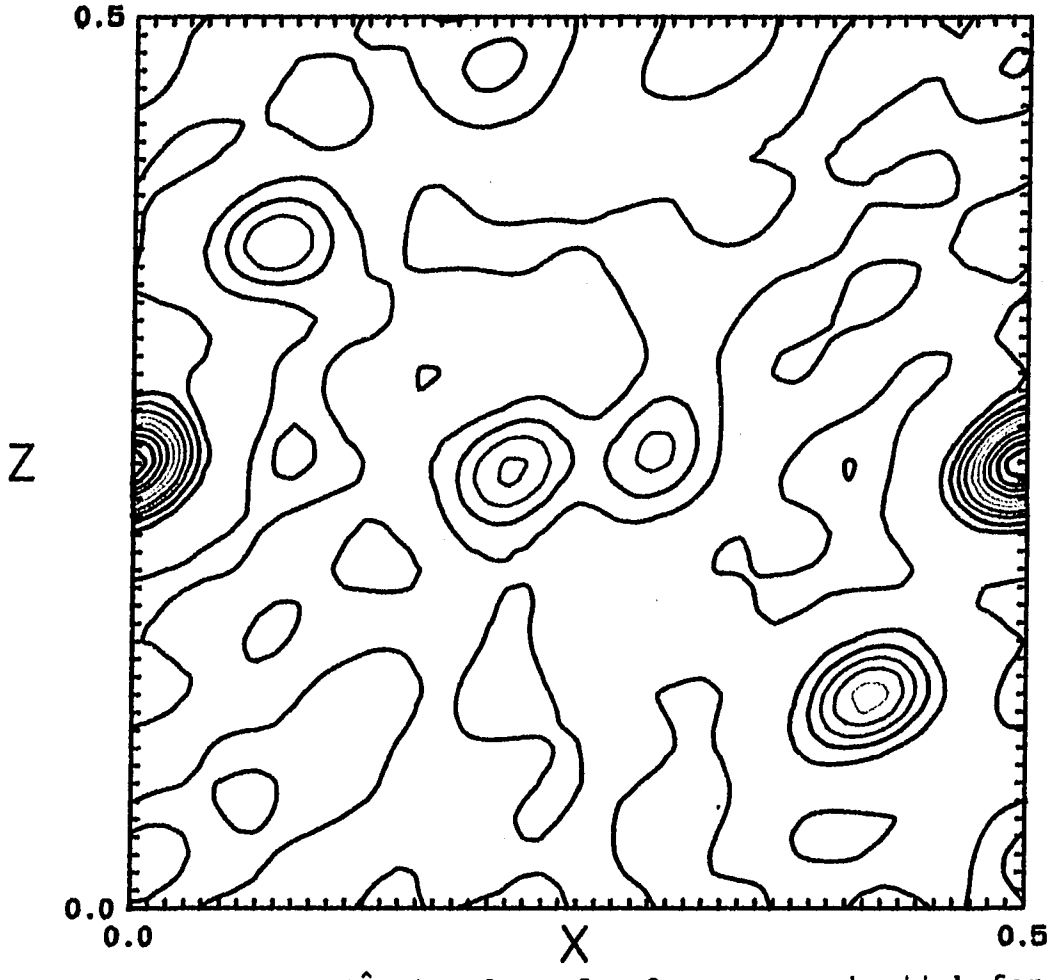


Figure 5.16. $Q(\hat{S},t)$ values for 2_1-screw axis || b for NIT

Table 5.10. Symmetry Operator Locations for $Ni(C_6O_3H_9)_2Cl_2$

symmetry operator	t[a]	$Q(\hat{S},t)$	$S(\hat{S},t)$
-x,1/2+y,-z	(-.0024, 0.0, 0.2493)	0.8845	0.0140
x,-y,z+1/2	(0.0, 0.2475, 0.0)	0.9343	0.0073
-x,-y,-z	(-.0024, -.0029, -.0003)	0.8317	0.0213

[a]Translations are given in fractions of a unit cell.

values of the $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions for the translations.

There is essentially one peak in the $Q(\hat{S},t)$ function for the c-glide plane, and it corresponds to the location of the glide plane in the Ni image. Similarly, there is one large peak in the $Q(\hat{S},t)$ function for the 2_1-screw axis. For this compound, the Ni atom lies on an inversion center. Thus, all vectors between the Ni atom and all other atoms in the cell (except the other Ni atom) are multiple vectors.

According to superposition theory, there should be 2m images in the superposition map, where m is the multiplicity of the superposition vector. The fact that a weighted superposition was done only suppresses the O images relative to the Ni images, and does not eliminate these images. The vector used was of multiplicity two (the vector from the Ni atom to the O atom is the same as the vector from the inversion-related O atom to the Ni), so there should be four images in the superposition map. However, two of the

images, those with Ni at the tail of the shift vectors, exactly superimpose on top of one another, which leaves a result of three images.

The peaks in the $Q(\hat{S},t)$ function for the c-glide plane at 0.16 and 0.35 are the locations of the glide plane in the O images. The large peak at the origin of the $Q(\hat{S},t)$ function for the mirror plane indicates that there is still a sizable amount of pseudo-symmetry left in the superposition map. The $Q(\hat{S},t)$ function for the 2_1-screw axis has smaller peaks at (0.08, 0.0, 0.38) and (0.42, 0.0, 0.12) which are the locations of the 2_1-screw axis in the O images.

Since the Ni atom lies on an inversion center, and by convention, the origin of the electron density map in a centrosymmetric space group is taken to be on an inversion center, there is no need to transform the G_h according to equation (3.2). The G_h that were symmetry-related were averaged. The phases of 196 of the largest E_h were set by direct transfer of the phases of the corresponding G_h. The E_h were refined using the EG relationship and an electron density map was calculated. Figure 5.17 shows the projection of the electron density map onto the least-squares plane. The molecule is clearly identifiable in this figure. The Cl atoms are present in the electron density map, but are not included in the projection because they are not bound to any part of the structure. All of the unique

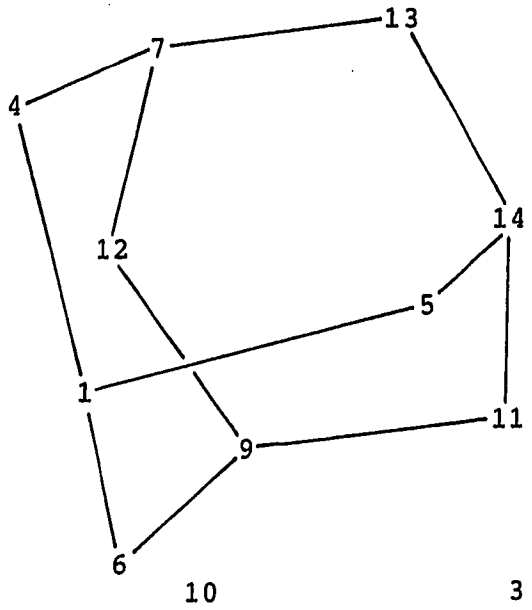


Figure 5.17. HYPAD projection of $Ni(C_6O_3H_9)_2Cl_2$

Table 5.11. HYPAD Atomic Coordinates[a] (x10^4) for $Ni(C_6O_3H_9)_2Cl_2$

atom	#	pk.ht.	x	y	z
Ni	1	529	-9	30	10
Cl	2	217	-4931	-1441	-2259
O1	5	92	1576	147	2535
O2	4	94	-1933	-817	940
O3	6	90	1902	-962	-260
C1	14	47	2184	-525	3543
C2	13	47	265	-956	3735
C3	7	68	-935	-1418	2210
C4	12	50	491	-2002	1351
C5	9	56	2469	-1582	1180
C6	11	52	3736	-1112	2850

[a]Atomic coordinates are given as fractions of the unit cell.

Table 5.12. Refined Atomic Coordinates[a] (x10^4) for $Ni(C_6O_3H_9)_2Cl_2$

atom	x	y	z
Ni	0	0	0
Cl	-4927(2)[b]	-1462(7)	-2265(1)
O1	1591(4)	173(2)	2495(3)
O2	-1877(4)	-818(2)	945(3)
O3	1877(5)	-987(2)	-202(3)
C1	2270(6)	-553(3)	3579(5)
C2	307(6)	-998(3)	3802(5)
C3	-888(6)	-1447(3)	2194(5)
C4	533(7)	-2009(3)	1460(5)
C5	2495(7)	-1576(3)	1211(5)
C6	3775(7)	-1120(3)	2819(5)

[a]Atomic coordinates are given as fractions of the unit cell.

[b]Estimated standard deviations for the refined coordinates are given in parentheses for the least significant digit.

Table 5.13. Comparative bond distances for $Ni(C_6O_3H_9)_2Cl_2$

atoms	refined(Å)	HYPAD(Å)	Δ(Å)
Ni - O1	2.048(4)[a]	2.05	0.00
Ni - O2	2.056(4)	2.06	0.00
Ni - O3	2.029(5)	2.06	0.03
O1 - C1	1.457(8)	1.35	-.11
O2 - C3	1.466(8)	1.44	-.03
O3 - C5	1.470(8)	1.50	0.03
C1 - C2	1.537(10)	1.47	-.07
C1 - C6	1.532(10)	1.58	0.05
C2 - C3	1.533(9)	1.48	-.05
C3 - C4	1.518(10)	1.59	0.07
C4 - C5	1.526(11)	1.49	-.04
C5 - C6	1.518(10)	1.59	0.07

[a]Estimated standard deviations for the refined distances are given in parentheses for the least significant digit.

Table 5.14. Comparative bond angles for $Ni(C_6O_3H_9)_2Cl_2$

atoms	refined(°)	HYPAD(°)	Δ(°)
O1 - Ni - O2	86.55(17)[a]	84.9	-1.7
O1 - Ni - O3	90.64(18)	89.9	0.7
O2 - Ni - O3	87.79(18)	88.3	0.5
Ni - O1 - C1	119.37(35)	121.8	2.4
Ni - O2 - C3	119.71(38)	118.3	-1.4
Ni - O3 - C5	120.59(38)	118.1	-2.5
O1 - C1 - C2	107.04(15)	107.5	0.5
O1 - C1 - C6	111.28(51)	112.6	1.3
C2 - C1 - C6	111.58(56)	113.2	1.6
C1 - C2 - C3	113.72(53)	116.3	2.6
O2 - C3 - C2	107.07(49)	108.2	1.1
O2 - C3 - C4	110.33(50)	107.4	-2.9
C2 - C3 - C4	112.32(58)	113.7	1.4
C3 - C4 - C5	114.24(56)	112.8	-1.4
O3 - C5 - C4	107.03(56)	108.0	1.
O3 - C5 - C6	109.88(54)	108.7	-1.2
C4 - C5 - C6	112.16(56)	115.1	2.9
C5 - C6 - C1	114.38(59)	109.3	-5.1

[a]Estimated standard deviations for the refined angles are given in parentheses for the least significant digits.

non-hydrogen atoms in the molecule are found in the top 14 peaks. The HYPAD atomic positions, peak number and peak heights are given in Table 5.11 and the refined atomic positions are given in Table 5.12. Tables 5.13 and 5.14 compare the bond distances and bond angles obtained from the HYPAD and refined positions. Again, the HYPAD-generated bond distances are within 0.11Å of the refined distances and the HYPAD-generated bond angles are generally within 3° of the refined angles. Considering the fact that the nominal resolution in the map was 0.25 - 0.41Å, these results are excellent.

The symmetry averaging of the G_h is important as it removes many peaks in the electron density map which are "related" to actual peaks by pseudo-symmetry. Without symmetry averaging, the unique atoms are found in the top 19 peaks in the map. However, many of the extraneous peaks are pseudo-related to actual peaks by a mirror or a 2-fold axis, neither of which is present in the space group $P2_1/c$. Many of these pseudo-peaks are more prominent than the actual peaks. When the G_h are averaged, only one pseudo-related peak remains, a pseudo-chlorine, at approximately ½ the peak height of the actual chlorine atom.

HYPAD Solution of $(NHC_5H_5)_2I_{10}$

This compound formed as a side product of the reaction that produced crystals of $(NHC_5H_5)SbI_4$ (see Appendix B). One of more interesting physical characteristics of this compound is that it forms extremely long needle-shaped crystals, several centimeters in length. It was initially thought that these crystals were an alternate crystalline form of the antimony compound. However, the structure, shown in Figure 5.18, indicates otherwise.

This compound crystallizes in the orthorhombic crystal system. The extinction conditions upon the observed intensities only indicated the presence of a b-glide plane perpendicular to c. Thus, there were three possible space groups: Pmmb (centrosymmetric), P2mb (non-centrosymmetric) and $Pm2_1b$ (non-centrosymmetric).

Direct methods failed to provide a solution in any of the three possible space groups. Attempts to analyze the Patterson map by hand were fruitless. A vector which seemed a likely candidate for an Sb-I vector was used for an unweighted superposition, since the weighting factor for this vector is almost unity. The Fourier transform coefficients of the superposition map were calculated. The locations of all of the possible symmetry elements in the three space groups are listed in Table 5.15, along with $Q(\hat{S},t)$ and "mirrored" $S(\hat{S},t)$ values. Translations

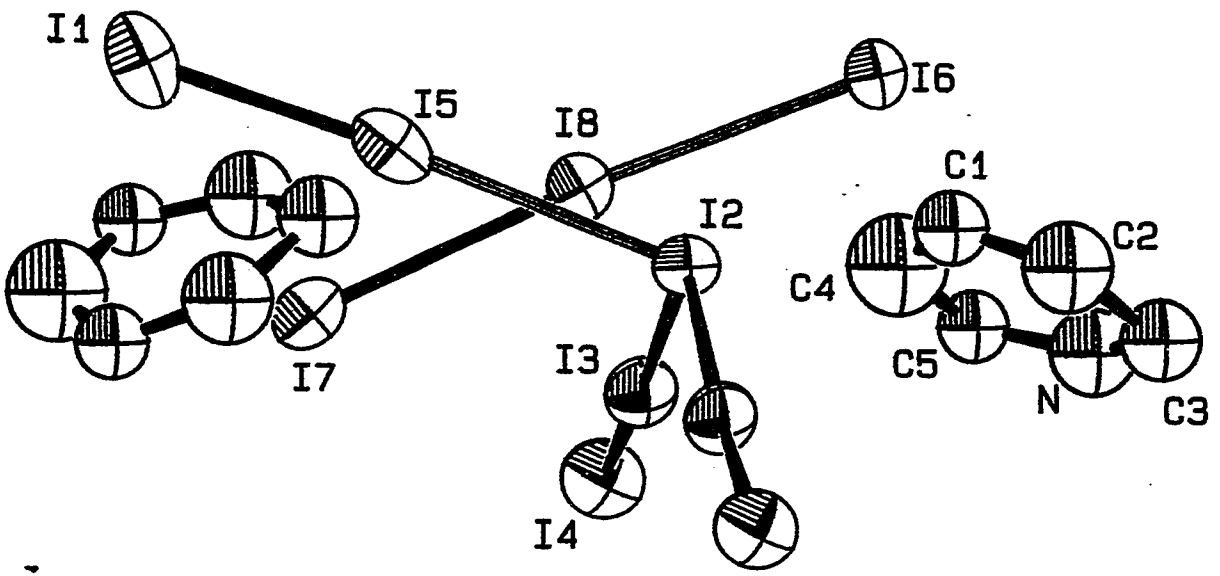


Figure 5.18. ORTEP drawing for $(NHC_5H_5)_2I_{10}$

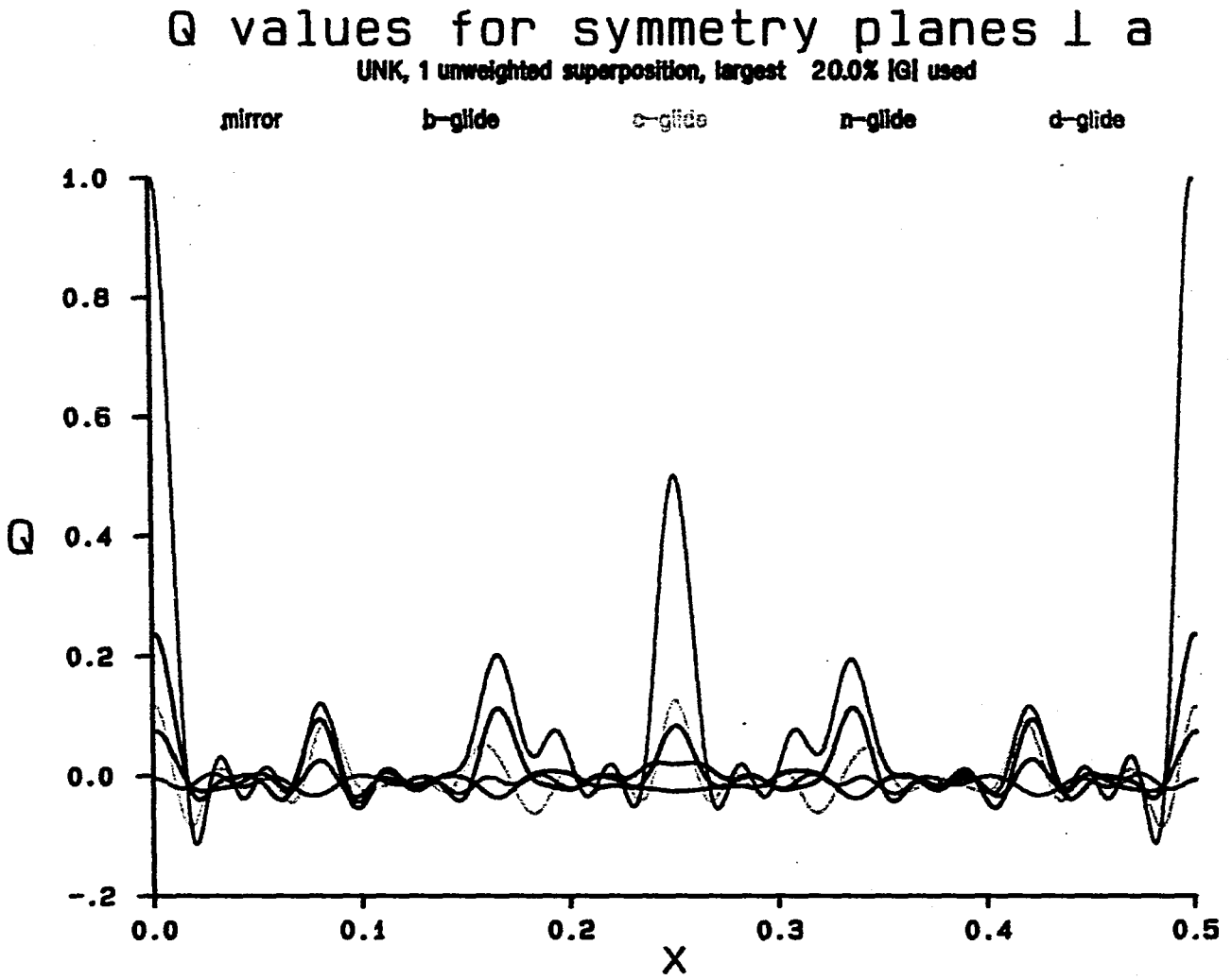


Figure 5.19. $Q(\hat{S},t)$ values for symmetry planes $\perp$ a for UNK

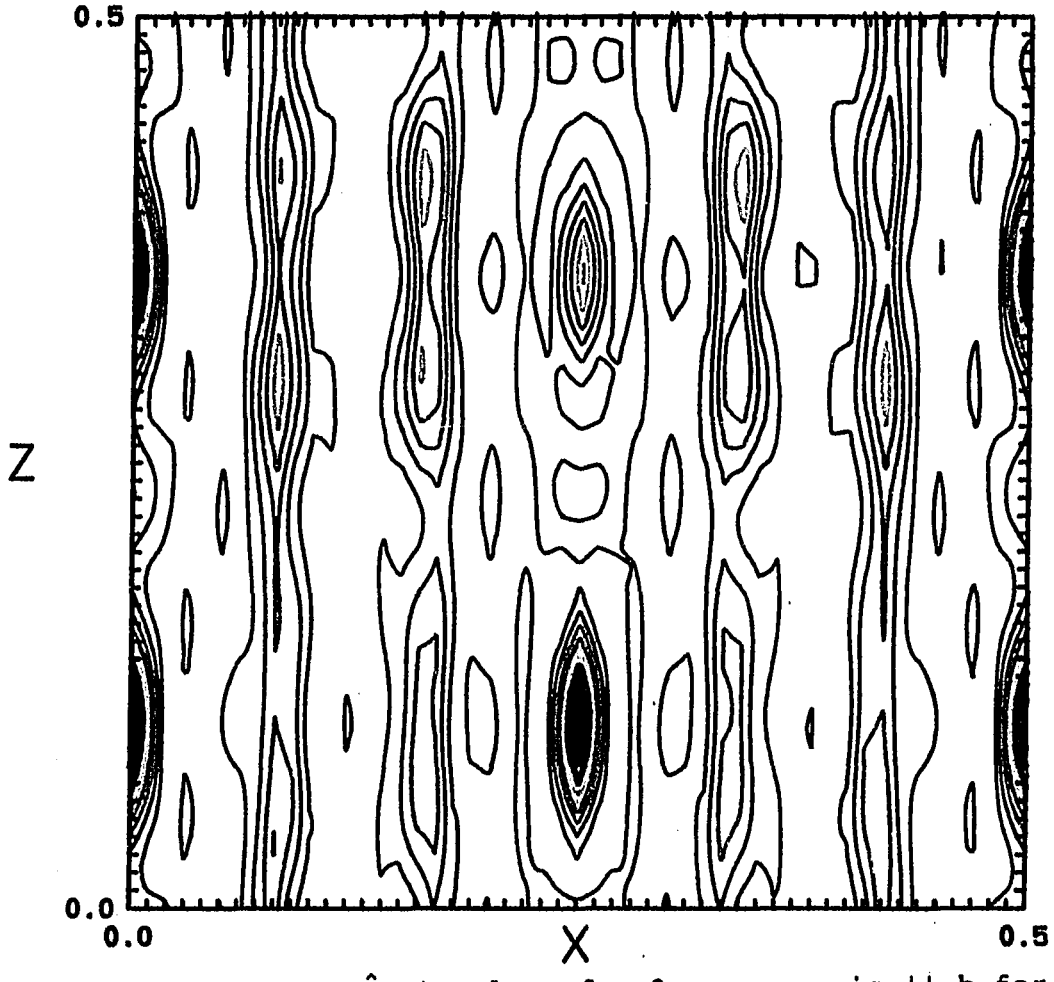


Figure 5.20. $Q(\hat{S},t)$ values for 2_1-screw axis || b for UNK

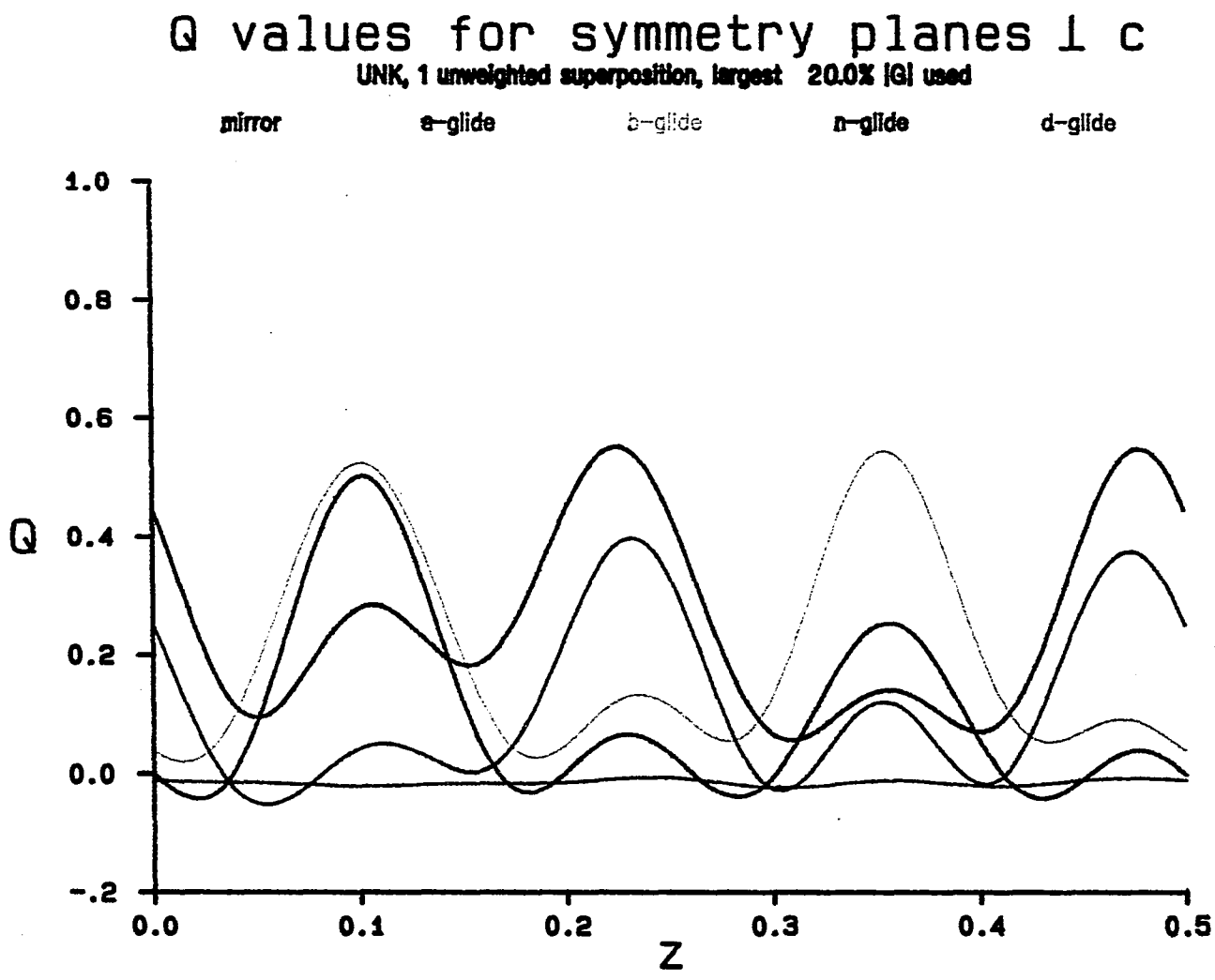


Figure 5.21. $Q(\hat{S},t)$ values for symmetry planes $\perp$ c for UNK

Table 5.15. Symmetry Operator Locations for $(NHC_5H_5)_2I_{10}$

symmetry operator	t[a]	$Q(\hat{S},t)$	$S(\hat{S},t)$
x,-y,-z	(0.0, 0.4220, 0.3517)	0.9772	0.0030
-x,y+1/2,z	(0.0, 0.0, 0.1011)	0.5232	0.0037
-x,-y,z	(0.0, 0.1728, 0.0)	0.5459	0.0047
-x,y,z	(0.2500, 0.0, 0.0)	0.5018	0.0001
x,-y,z	(0.0, 0.1728, 0.0)	0.5465	0.0046
x,y+1/2,-z	(0.0, 0.0, 0.1011)	0.5236	0.0036
-x,-y,-z	(0.0000, 0.4220, 0.3517)	0.9766	0.0031

[a]Translations are given in fractions of a unit cell.

consistent with the b-glide plane perpendicular to c were found only for the mirror perpendicular to a and the 2_1-screw axis parallel to b, indicating that the correct space group was $Pm2_1b$. The $Q(\hat{S},t)$ functions are plotted for the mirror perpendicular to a, the 2_1-screw axis parallel to b and the b-glide plane perpendicular to c in Figures 5.19, 5.20, 5.21, respectively. For this space group, the origin is located on the 2_1-screw axis. The positions of the ends of the shift vector relative to the location of the 2_1-screw axis were calculated and used as input to a least-squares refinement. The rest of the molecule was found by successive structure factor and electron density map calculations.

Although the structure was not immediately evident from the superposition map, it is important to note that the space group and the location of the symmetry elements could

be obtained. Even in a worst-case scenario, with an unweighted superposition and a very vague idea as to what types of atoms are in the structure, one can at least get a "foot in the door" and continue with a least-squares refinement.

HYPAD Solution of $Cu(C_6O_3H_9)Cl_2$

This compound (CUT), shown in Figure 5.22, was expected to be the Cu analogue to $Ni(C_6H_9O_3)_2Cl_2$; however its structure turned out to be different. It was originally solved by Bill Jenson's group (Department of Chemistry, South Dakota State University) using direct methods. Although the position of the Cu atom can be obtained from the Patterson map, it does not provide enough phase information to solve the structure via the "heavy-atom" method.

This compound crystallizes in the monoclinic crystal system. The extinction conditions on the intensities indicated the presence of a c-glide plane perpendicular to b and C-centering. There are two possible space group choices: Cc (non-centrosymmetric) and C2/c (centrosymmetric).

A vector corresponding to a Cu-O vector was chosen to use as a superposition shift vector. Selection of an appropriate vector was complicated by the fact that the

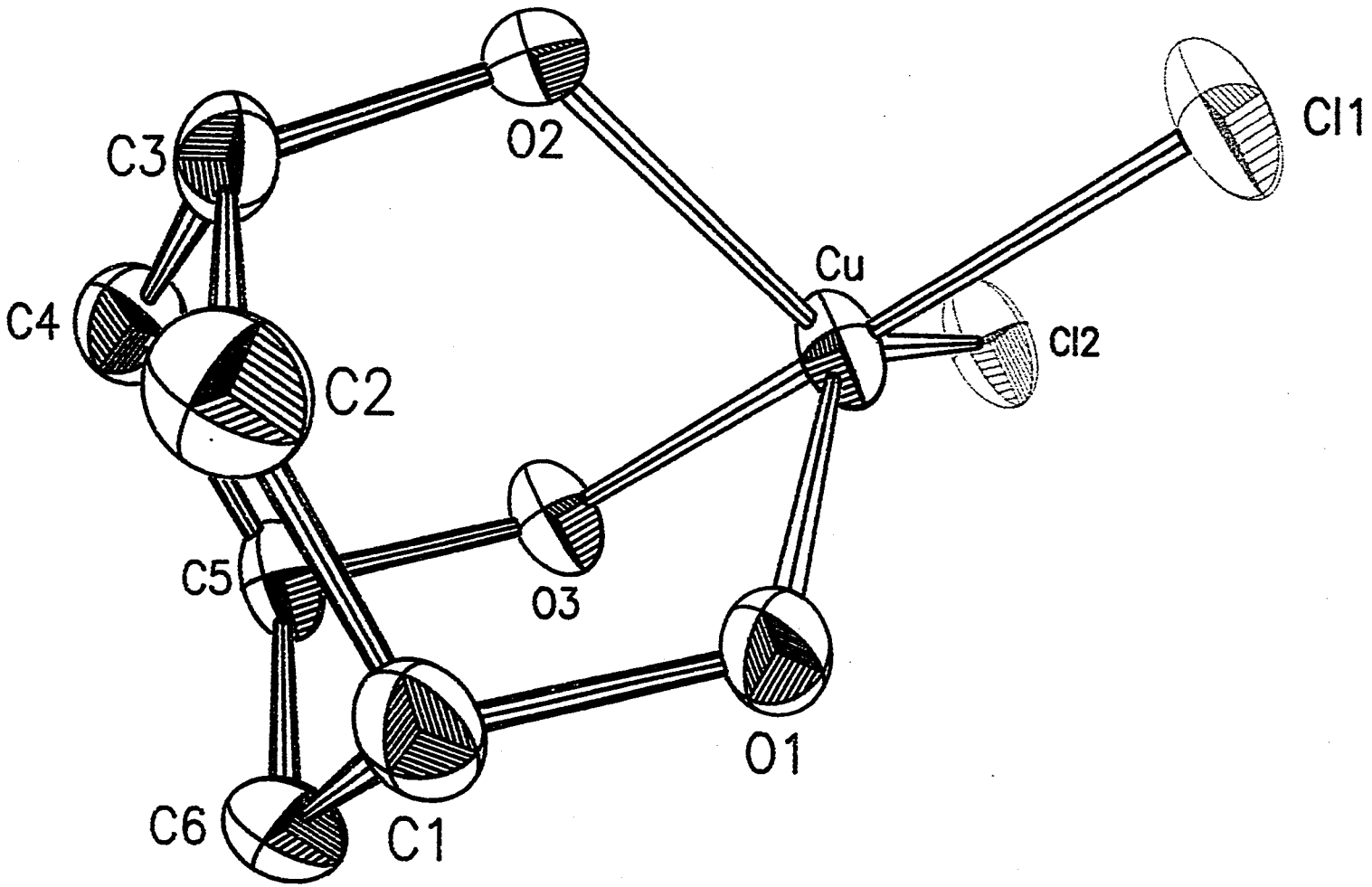


Figure 5.22. ORTEP drawing for $Cu(C_6O_3H_9)Cl_2$

Cu-Cl bond distances are comparable to the Cu-O bond distances. A weighted superposition was done and the Fourier transform coefficients of the map were calculated. $Q(\hat{S},t)$ and $S(\hat{S},t)$ functions for all of the possible symmetry elements in both space groups were calculated and the best translation choices are listed in Table 5.16. A plot of the $Q(\hat{S},t)$ values for symmetry planes perpendicular to **b** is shown in Figure 5.23.

There is still a large amount of pseudo-symmetry remaining as indicated by the large peaks in the $Q(\hat{S},t)$ functions of the mirror and a-glide planes, as well as the large Q values for the rotation-type axes and the inversion center translations. The locations of the 2-fold axis, the 2_1-screw axis (generated from the 2-fold axis by the C-centering operation), and the inversion centers are not consistent with the location of the c-glide plane, indicating that the space group is Cc. The other symmetry elements are pseudo-symmetry from the Patterson map which has symmetry C2/m.

There are two large peaks in the c-glide plane $Q(\hat{S},t)$ function. The peak at 0.18 is the location of the glide plane in the Cu image. The other large peak, at 0.32 is the result of pseudo-symmetry. The peak at 0.07 is the location of the glide plane in the O image and the peak at 0.43 is the result of pseudo-symmetry. The C-centering operation

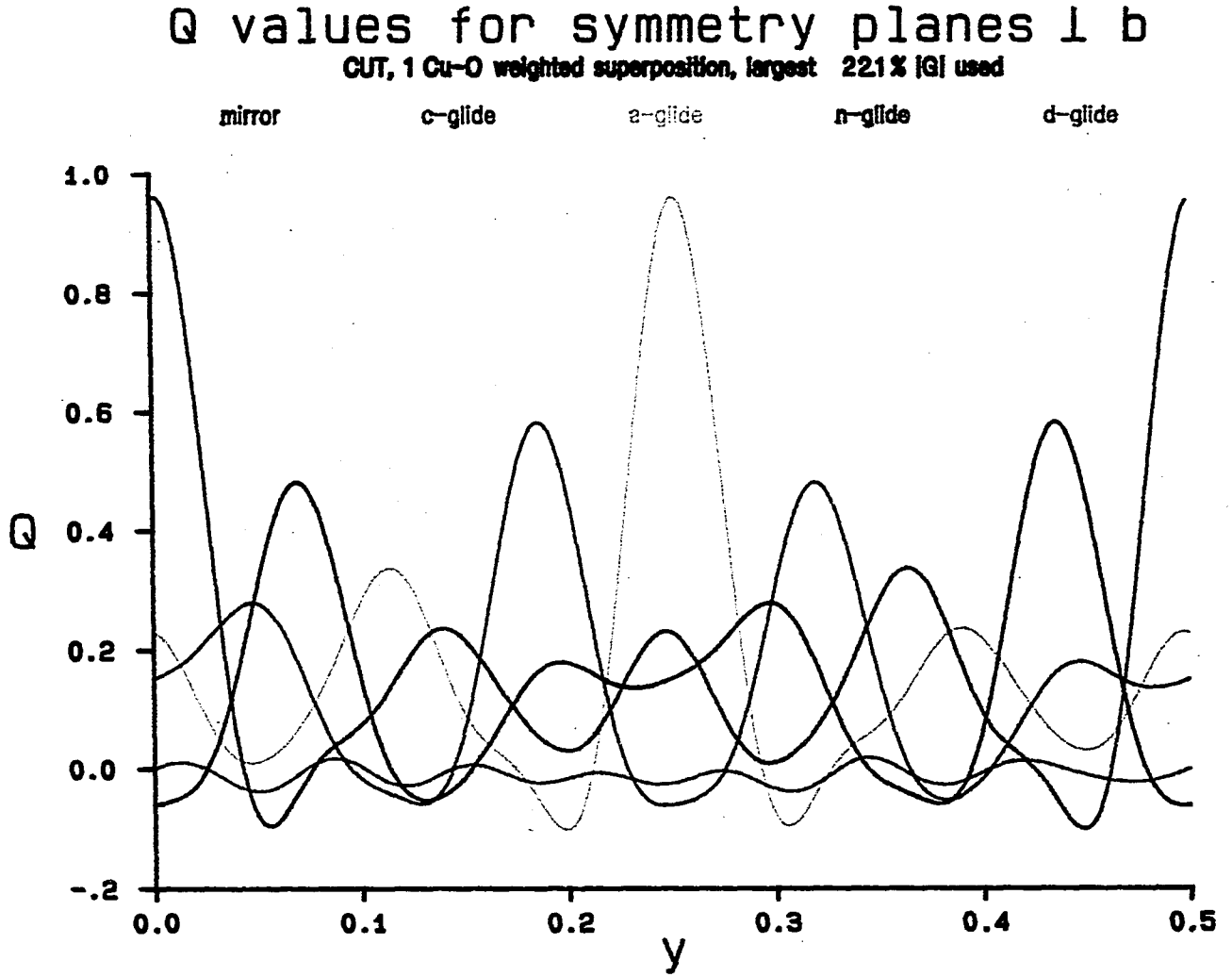


Figure 5.23. $Q(\hat{S},t)$ values for symmetry planes $\perp$ b for CUT

Table 5.16. Symmetry Operator Locations for $Cu(C_6O_3H_9)Cl_2$

symmetry operator	t[a]	Q(Ŝ,t)	S(Ŝ,t)
-x,y,-z	(-.0005, 0.0, 0.0012)	0.9004	0.0201
-x,y+1/2,-z	(0.2495, 0.0, 0.0012)	0.9004	0.0201
x,-y,z+1/2.	(0.0, 0.1861, 0.0)	0.5820	0.0907
1/2+x,-y,1/2+z.	(0.0, .4361, 0.0)	0.5820	0.0907
-x,-y,-z	(-.0005, 0.0025, 0.0013)	0.8554	0.0293
-x,-y,-z	(0.2495, 0.2525, 0.0011)	0.8554	0.0293

[a]Translations are given in fractions of a unit cell.

generates an n-glide plane one quarter of the unit cell away from the c-glide plane, which is clearly indicated in the plot.

Even though there is a large amount of pseudo-symmetry remaining in the superposition map, when the G_h are transformed, the phases of the E_h are refined using the EG relationship and an electron density map is calculated, the Cu atom, both Cl atoms and all three O atoms are all found in the largest 9 peaks on the map. The Cu atom and the Cl atoms are the three largest peaks in the electron density map. The three other large peaks in map are pseudo-related to the Cl atoms and one of the oxygen atoms.

Two additional weighted superpositions were carried out using the Cu-Cl vectors. The list of peaks in the superposition map was reduced using TRIM. The 300 highest peaks in the list were checked for c-glide-related peaks in the list (using a tolerance of 0.6Å). The Fourier transform

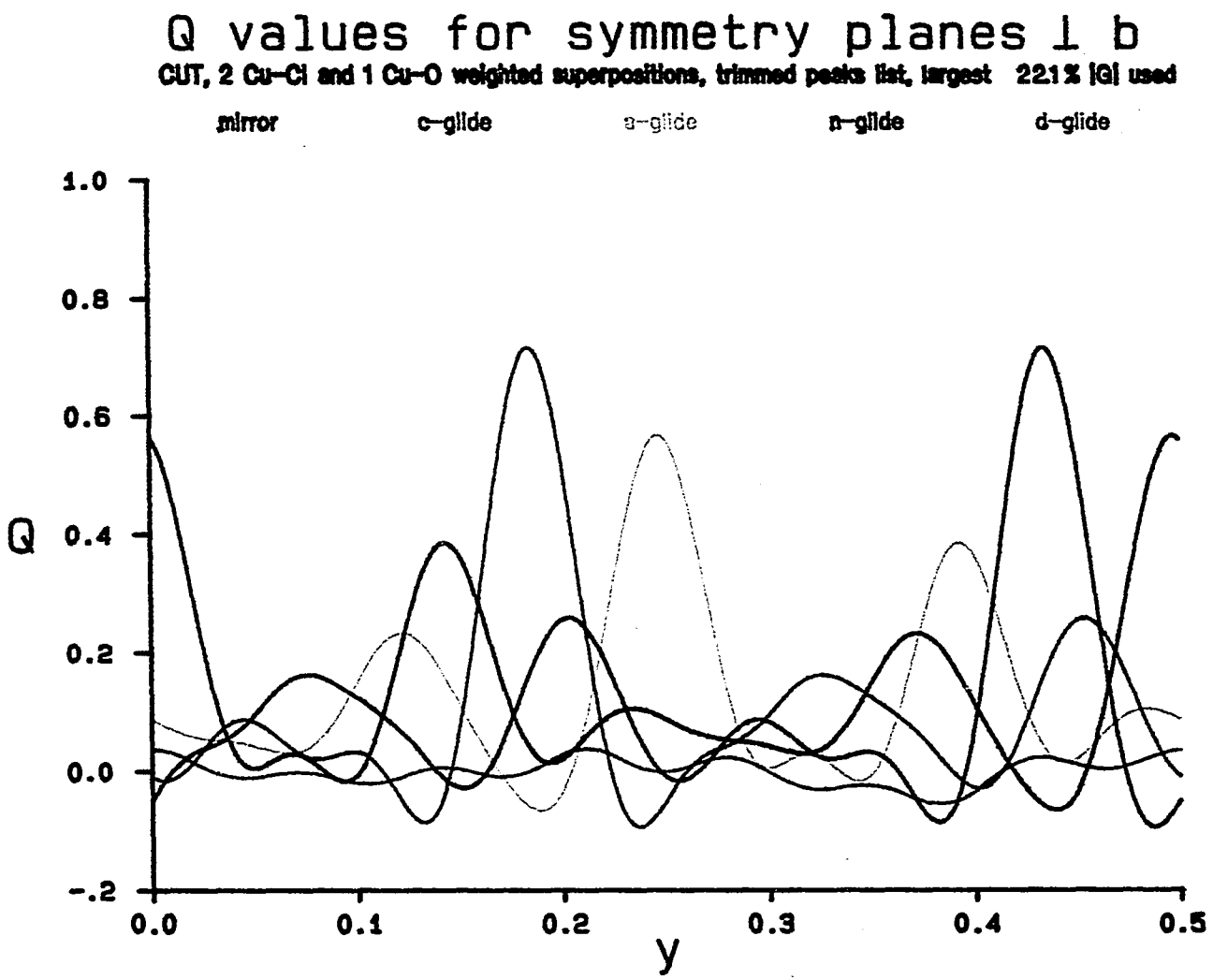


Figure 5.24. $Q(\hat{S},t)$ values for symmetry planes $\perp$ b for CUT

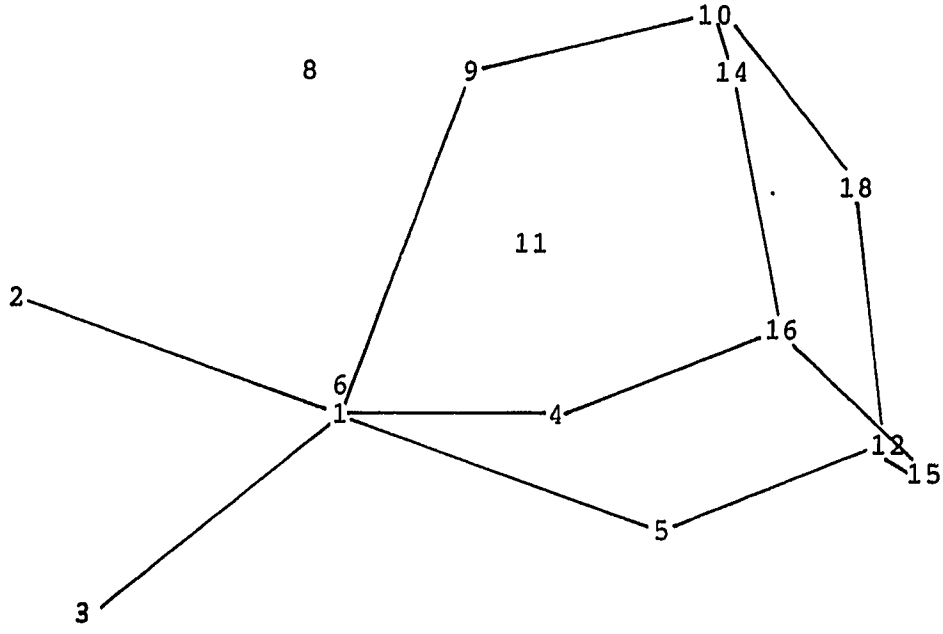


Figure 5.25. HYPAD projection of $Cu(C_6O_3H_9)Cl_2$

Table 5.17. HYPAD Atomic Coordinates[a] (x10^4) for $Cu(C_6O_3H_9)Cl_2$

atom	#	pk.ht.	x	y	z
Cu	1	264	-18	1879	-1
Cl1	2	111	1563	276	818
Cl2	3	107	-1355	580	-2499
O1	4	59	470	1870	2192
O2	9	39	988	4355	782
O3	5	58	-1426	3430	-672
C1	16	24	338	3396	2779
C2	14	25	1303	4747	3026
C3	10	37	1045	5650	1954
C4	18	23	-505	6256	369
C5	12	27	-1407	4869	160
C6	15	24	-908	4090	1674

[a]Atomic coordinates are given as fractions of the unit cell.

Table 5.18. Refined Atomic Coordinates[a] (x10^4) for $Cu(C_6O_3H_9)Cl_2$

atom	x	y	z
Cu	0 [b]	1892(1)	0
Cl1	1593(2)	225(4)	869(3)
Cl2	-1242(2)	575(4)	-2406(3)
O1	579(6)	1797(8)	2297(7)
O2	1017(5)	4366(8)	829(7)
O3	-1344(6)	3395(9)	-681(7)
C1	528(9)	3475(13)	2945(11)
C2	1372(8)	4872(14)	3214(11)
C3	945(9)	5703(12)	1692(11)
C4	-363(9)	6335(13)	557(12)
C5	-1209(8)	4963(12)	262(11)
C6	-788(9)	4162(13)	1788(12)

[a]Atomic coordinates are given as fractions of the unit cell.

[b]Estimated standard deviations for the refined coordinates are given in parentheses for the least significant digit.

Table 5.19. Comparative bond distances for $Cu(C_6O_3H_9)Cl_2$

atoms	refined(Å)	HYPAD(Å)	Δ(Å)
Cu - Cl1	2.243(3)[a]	2.22	-.02
Cu - Cl2	2.232(3)	2.30	0.07
Cu - O1	2.151(6)	2.10	-.05
Cu - O2	2.126(6)	2.12	-.01
Cu - O3	1.949(7)	2.05	0.10
O1 - C1	1.464(11)	1.39	-.07
O2 - C3	1.454(11)	1.61	0.16
O3 - C5	1.493(11)	1.41	-.08
C1 - C2	1.497(14)	1.62	0.12
C1 - C6	1.538(14)	1.47	-.07
C2 - C3	1.521(13)	1.21	-.31
C3 - C4	1.518(14)	1.78	0.26
C4 - C5	1.476(14)	1.59	0.11
C5 - C6	1.520(13)	1.46	-.06

[a]Estimated standard deviations for the refined distances are given in parentheses for the least significant digit.

Table 5.20. Comparative bond angles for $Cu(C_6O_3H_9)Cl_2$

atoms	refined(°)	HYPAD(°)	Δ(°)
Cl1 - Cu - Cl2	92.86(11)[a]	94.3	1.4
Cl1 - Cu - O1	92.15(19)	96.2	4.0
Cl1 - Cu - O2	91.59(18)	90.7	-.9
Cl1 - Cu - O3	177.84(20)	177.8	0.0
Cl2 - Cu - O1	143.15(18)	141.7	-1.5
Cl2 - Cu - O2	130.80(17)	129.9	-.9
Cl2 - Cu - O3	89.28(20)	87.9	-1.4
O1 - Cu - O2	85.49(23)	86.7	1.2
O1 - Cu - O3	86.30(25)	82.5	-3.8
O2 - Cu - O3	86.79(26)	87.7	0.9
Cu - O1 - C1	119.10(50)	122.3	3.2
Cu - O2 - C3	119.93(51)	117.3	-2.6
Cu - O3 - C5	124.02(51)	128.9	4.9
O1 - C1 - C2	109.27(78)	98.6	-10.7
O1 - C1 - C6	109.00(74)	112.9	3.9
C2 - C1 - C6	112.33(78)	114.7	2.4
C1 - C2 - C3	114.32(78)	122.5	8.2
O2 - C3 - C2	110.87(72)	109.9	-1.0
O2 - C3 - C4	106.75(72)	91.0	-15.8
C2 - C3 - C4	112.17(80)	116.9	4.7
C3 - C4 - C5	114.45(80)	113.3	-1.2
O3 - C5 - C4	110.01(73)	100.5	-9.5
O3 - C5 - C6	107.03(69)	106.3	-.7
C4 - C5 - C6	113.09(78)	112.8	-.3
C5 - C6 - C1	112.69(78)	117.8	5.1

[a]Estimated standard deviations for the refined angles are given in parentheses for the least significant digits.

coefficients of this new superposition map were determined. A plot of the $Q(\hat{S},t)$ function for symmetry planes perpendicular to **b** is shown in Figure 5.24. As can be seen in this plot, the amount of pseudo-symmetry has been reduced significantly. The G_h were transformed and the phases of the E_h that were refined using the EG relationship were used to calculate an electron density map. Figure 5.25 shows the projection of the electron density map onto the least-squares plane. All of the non-hydrogen atoms were found in the largest 18 peaks in the map. The HYPAD atomic positions, peak number and peak height are given in Table 5.17 and the refined positions are given in Table 5.18. The bond distances and bond angles obtained from the HYPAD and the refined positions are listed in Tables 5.19 and 5.20, respectively.

CHAPTER 6. CONCLUSION

I hope that the results presented in Chapter 5 successfully show that HYPAD is a viable alternative to the existing automatic phase determining techniques. We have shown that for simple organo-metallics, complete structural solutions can be obtained and for more complex compounds, a sizable part of the structure can be obtained. It is encouraging to note that if the space group is not uniquely defined by the extinction conditions, HYPAD also appears to be able to correctly determine the space group.

Although the presence of spurious peaks in the superposition map can cause erroneous phase predictions, centrosymmetric structures seem to be less susceptible to this than do non-centrosymmetric structures. This is not unexpected, since there are only two phase choices for the structure factors of centrosymmetric structures. It is very unlikely that the spurious peaks will be in the correct orientation to cause the phase to be shifted from 0° to 180° (or 180° to 0°). It is more likely that the phase would be shifted by several degrees, say 20 or 30°, which could have a significant effect on the EG refinement in the non-centrosymmetric case.

There is a growing need for a procedure that can work as an alternative to direct methods. Very often, when direct methods fails, it is the result of a poor starting set of

phases. The user who is inexperienced in the use of direct methods is faced with a dilemma when the first attempt fails. Usually, there is no obvious choice as to what changes to make in the starting set or in the phases initially assigned to these reflections. It has been shown that a good initial set of phases can be obtained from the transformed Fourier coefficients of the superposition map, and that these phases, when refined via the EG relation, can produce, in most cases, either the structure or a sizable fragment thereof.

Furthermore, if HYPAD fails in its first attempt, other superposition vectors can be chosen by using the location of the symmetry elements to locate symmetry-related peaks. The superposition map obtained using these vectors will generally provide a good starting set of phases.

CHAPTER 7. FUTURE WORK

It should be realized that HYPAD is still in the developmental stage. A number of the structures selected up to this point were chosen to provide a means of testing the computer code rather than to fully explore the the capabilities of the method. There are still areas of the hybrid Patterson superposition-direct methods philosophy that need to be examined. It is to be fully expected that there will be structures that resist solution by this method. It would be a great benefit, then, to examine these types of structures to try to pinpoint at which aspect(s) of the current methodology difficulties are encountered.

One of the more obvious difficulties that is encountered is the presence of multiple images. As has been shown, weighted superpositions reduce this problem, but does not eliminate it. Since the $\xi(t)$ and $\xi'(t)$ functions, as well as the $Q(\hat{S},t)$ function, can be used to determine the location of one image relative to another, it would seem that it should be possible to use the G_h from an unweighted superposition to gain information about a single image, but very little work has been done in this area.

In many of the structures which seem to be unsolvable, the space group symmetry is questionable. Patterson, and to a lesser extent, superposition maps contain additional symmetry which can further cloud the issue. While the

results obtained concerning the determination of space group symmetry are very encouraging, the user must be conscious of the effects of pseudo-symmetry. The ultimate goal, then, would be the automatic determination of the space group by the program without user input. Obviously, if this automation could be done, then the rest of the procedure could follow.

The routine could automatically shift the phases of the G_h, assign and then refine the phases of the E_h. A quantity which would indicate how good the phases are, such as the residual index, would be calculated and depending upon its value, one of two choices would be made. If the phases were considered "good", the refined phases would be reintroduced into the EG or Σ_2 relation to produce a new refined set of phases. The iterations of phase refinement would continue until the phase set has converged. If the phases were considered "poor", the routine would use the locations of the symmetry elements to automatically calculate the positions of atoms that are symmetry-related to the atoms in the shift vector used in the original superposition. These interatomic vectors would then be used as shift vectors in multiple superpositions. New Fourier coefficients would be calculated and the process restarted, like the approach used in the solution of $FeP_2OC_{32}H_{31}I$ described in Chapter 5.

Ideally, the procedure would be automated to the point that the user would only need to indicate the peak to be

used for the first superposition, and the procedure would take over, making all of the decisions, and outputing the solution. In fact, it might be possible to have the routine make the choice of the first superposition vector using the distance of the peak from the origin and the peak height as criteria for its decision.

One always has to worry that the EG relation may produce phases for the E_h that resemble the superposition map more than they resemble the phases of the electron density map. It may turn out that it is best if both the EG and Σ_2 refinements are carried out in tandem. In the early stages of refinement, the phases obtained from the EG relationship would have more weight than those obtained from the Σ_2 relation, but as the refinement proceeded, the EG phases would decrease in importance and the Σ_2 phases would increase in importance.

APPENDIX A. KEYWORD OPTIONS IN HYPAD

All of the user-created input files that are not created by the CHES.CAT procedure use a keyword format for the imput of parameters into the individual programs in HYPAD. This type of approach allows the user to easily see which parameters will be used, and greatly facilitates modification of the input files Keywords for the individual programs are listed below.

Only the first three letters of each keyword are required. In the following lists, the keywords are given in uppercase letters and the type of value for the keyword (if any) is indicated by one of three codes: n => integer number, x => real (floating point) number, c => character string. Some of the keywords have default values that the program will use if the keyword is missing; these are indicated in []'s. Except in certain cases, it is possible to have more than one keyword (and its accompanying value) on a line. The keyword and its value(s) must be on the same line. Special cases when only one keyword can appear on a line are indicated in the description of the appropriate keyword.

GCALC.INP

PEAKS n - the number of peaks to be read in from xxxSUPR.PKS and used to calculate the G_h. This number must be less than or equal to 200, and if it is 0, then the entire superposition map will be read in and transformed.
(Map option is not available at this time)

2THETA x - the maximum 2θ value to use in calculating the G_h.

WAVELENGTH x - the wavelength of the radiation used.

LAUE n - code indicating the Laue symmetry. The Laue symmetries and their corresponding codes are:

1 => triclinic, -1 symmetry
2 => monoclinic, b-unique, 2/m symmetry
3 => monoclinic, c-unique, 2/m symmetry
4 => orthorhombic, mmm symmetry
5 => tetragonal, 4/m symmetry
6 => tetragonal, 4/mmm symmetry
7 => trigonal, -3 symmetry
8 => trigonal, -3m symmetry
9 => hexagonal, 6/m symmetry
10 => hexagonal, 6/mmm symmetry
11 => cubic, -3m symmetry

12 => cubic, m3m symmetry

LATTICE c - code indicating the centering of the lattice.

P => Primitive

A => A-centered

B => B-centered

C => C-centered

I => I-centered (body centering)

F => F-centered (face centering)

R => R-centered (rhombohedral)

UNKNOWN - space group is unknown. If this keyword is included anywhere in the input file, the space group will be considered unknown, even if the KNOWN keyword appears later in the file.

KNOWN - space group is known (inclusion of this is optional)

LOCSYM.INP

2THETA x x - the minimum and maximum 2θ values to use. All G_h in this range will be used in the calculations.

LAUE n - code indicating the Laue symmetry. See above for the appropriate codes.

WEIGHTS c - indicates the weighting scheme to be used in calculating $S(\hat{S},t)$. "c" must be one of the choices given below, no abbreviations are allowed.

UNIT => $|G_h G_{h'}|$ weights

E(h) => $|E_h G_h G_{h'}|$ weights

FRACTIONG x - the fraction of the G_h to be used in calculating $Q(\hat{S},t)$ and $S(\hat{S},t)$.

NSUPERPOS n - the number of superpositions that were done.

PLOT c - indicates that graphics output is desired. The commands required to draw the plots of the $Q(\hat{S},t)$ values for the symmetry planes will be written out.

HP => HP-GL commands

UNWTSUPPOS - indicates that unweighted superpositions were done and that the "mirrored" form of $S(\hat{S},t)$ will be calculated.

WTSUPERPOS - indicates that weighted superpositions were done (inclusion of this is optional as it is the default condition).

SPECIAL - If the user suspects that the tail atom of the superposition (i.e., the one at the origin) actually lies on a special position, then this option should be used. The figure-of-merit approach sometimes

passes over a valid translation choice because the program thinks that it is a result of the superpostion failing to completely remove the Patterson symmetry. This option will save the two "best" translations for later analysis.

SYMMETRY - indicates that the rest of the file contains symmetry operations. Therefore, it MUST be the last keyword in the list. It must stand alone (i.e., do not include a symmetry operator on the same line because the operator will be ignored and do not include it on the same line with other keywords as they will also be ignored).

The remaining lines are formatted (A1,A26), and contain the variables COPT and CSYMM (if COPT is a #, then the following line is free-format).

COPT - a special option indicator
(blank) => no special option
* => space group symmetry operator check is desired, only works for symmetry planes

=> the translation for which the symmetry operator is to be evaluated follows. The next line is a free format line containing the x,y,z coordinates of the translation.

CSYMM - the symmetry operator in alpha-numeric form.

MULTIM.INP

2THETA x x - the minimum and maximum 2θ values to use. All G_h in this range will be used in the calculations.

FRACTIONG x - the fraction of the G_h to be used in calculating $\xi(t)$ and $\xi'(t)$.

FOURIER.INP

MAP n - indicates which type of Fourier map should be calculated. Map type 3 is the usual choice for the FOUR routine. The map types and their codes are:

1 => Patterson, using $|F_h|$

2 => Patterson, using $|F_h|^2$

3 => Electron density, using $F_h(obs)$

4 => Difference electron density, using $F_h(obs) - F_h(calc)$

5 => Electron density, using $F_h(calc)$

6 => Electron density, using $A_h(calc)*F_h(obs)^2$ and $B_h(calc)*F_h(obs)^2$

7 => Electron density, using $A_h(calc)*F_h(obs)$ and $B_h(calc)*F_h(obs)$

ORIENT n - the map orientation. HYPAD requires that orientation 2 be used.

	across	down	layers
1 =>	x	y	z
2 =>	x	z	y
3 =>	y	x	z
4 =>	y	z	x
5 =>	z	x	y
6 =>	z	y	x

FIELDWIDTH n - the number of columns of printed output for each grid point.

2 => 2 columns, base 10

3 => 3 columns, base 10

4 => 4 columns, base 10

6 => 2 columns, base 36 (0-9,A-Z)

7 => 3 columns, base 36

LOWVALUE n - the minimum value printed on the map

LAUE n - Laue symmetry of the crystal. NOTE!!! This is a different Laue code than what is used in the rest of the package.

1 => -1 symmetry, triclinic

2 => 2/m symmetry, monoclinic

3 => mmm symmetry, orthorhombic

4 => -3 symmetry, trigonal

5 => -3m symmetry, trigonal

6 => 4/m symmetry, tetragonal

7 => 4/mmm symmetry, tetragonal

8 => 6/m symmetry, hexagonal

9 => 6/mmm symmetry, hexagonal

10 => m3 symmetry, cubic

11 => m3m symmetry, cubic

BLANKS n - the number of blank lines between printed output lines

GRID n n n - the number of grid points along the 3 crystal axes. The first number of the number of grids along a, the second is the number of grids along b, and the last is the number of grids along c. The allowed values are 0, 16, 32, 64, 128.

LAYER n n - the minimum and maximum layers to be calculated. The minimum must be $\geq$ 0 and the maximum must be $\leq$ the number of grids in the layer direction.

OFFSET n n - the number of grids to shift the origin of the map across and down. These are usually 0.

SCALE x - the scale factor. 2/volume gives electron density maps in units of electrons/$Å^3$.

F000 x - the value of F(000) on the same scale as the map, otherwise set it equal to zero.

ACROSS c - fraction of the unit cell across the page.

QRTR => 1/4 of the cell

HALF => 1/2 of the cell

FULL => full cell

DOWN c - fraction of the unit cell down the page.

QRTR => 1/4 of the cell

HALF => 1/2 of the cell

FULL => full cell

OUTPUT c - the type of output desired. BOTH is the usual choice. For the rest of the FOUR routine to work, the binary file must be created.

PRTR => line printer only

DISK => binary file only

BOTH => both line printer and binary file

CENTRO - indicates a centrosymmetric cell.

NONCENTRO - indicates a non-centrosymmetric cell.

LATTICE c - indicates the lattice symmetry

P => Primitive

A => A-centered

B -> B-centered

C => C-centered

F => Face-centered

I => Body-centered

R => Rhombohedral

SYMMETRY n - the number of symmetry elements, excluding the center of symmetry and the lattice symmetry operations. This must be the last keyword in the list, as the program assumes that all lines following this contain symmetry operators in alpha-numeric form (X,1/2+Y,Z), one per line. This keyword must also be on a line by itself, any other keywords on this line will be ignored.

PRESEARCH.INP

The first line of the file must be the title (80 characters maximum). It is primarily used as a header for the SEARCH output. The following lines in the file contain keywords or symmetry elements.

ATOMS n	- the number of atoms in the asymmetric unit, ≥ 1.
ANOMALOUS n	- the number of types of atoms with significant anomalous scattering effects. Usually, these are atoms whose atomic numbers are greater than 9.
CELL x x x x x x	- the unit cell lengths a, b, c (in Å) and the angles alpha, beta, gamma (in degrees).
SYMMETRY n	- the number of symmetry elements, excluding the center of symmetry and the lattice symmetry operations. This must be the last keyword in the list, as the program assumes that all lines following this contain symmetry operators in alpha-numeric form (X,1/2+Y,Z), one per line. This keyword must also be on a line by itself, any other keywords on this line will be ignored.

SEARCH.INP

This routine is part of MULTAN80. The list of keywords are included here as a convenience. If the first card is

blank, defaults will be assumed throughout.

PEAKS n - n is the number of peaks to be searched for in the map. The default is (11*n)/9 where n is the number of independent atoms (excluding hydrogens). If two clusters are within 2.8Å of each other, the default is automatically increased to (3*n)/2.

ATOMS n - n is the number of additional atoms input by the user to be included in the interpretation of the map (number of peaks + number of additional atoms must be less than 200). See below.

NOJOIN - requests that no interpretation of the peaks be done. The program prints a list of the highest peaks and then stops. This condition is set automatically if no chemical bonds can be found by the program.

PROJECTIONS n - the number of orthogonal projections of each cluster to be plotted (to a maximum of three). The default is to always plot the least-squares projection, and to plot the projection orthogonal to the least- and most-squares planes if the cluster is cylindrical or spherical. n equal to zero

suppresses all plotting.

DOUT x - interpeak distances less than x Å are tabulated. [2.4]

DMIN x - minimum allowed bond length in the interpretation, in Å. [1.1]

DMAX x - maximum allowed bond length in the interpretation, in Å. [1.95]

AMIN x - minimum allowed bond angle in the interpretation, in °. [85]

AMAX x - maximum allowed bond angle in the interpretation, in °. [145]

METAL n - the number of heavy atoms in the asymmetric unit. When performing an interpretation of a cluster, the program will assume that the n highest peaks correspond to these heavy atoms. They will not be included in the interpretation, and all peaks within the maximum bond length will be marked as spurious. This is useful when the compound contains a metal atom ionically bonded to the organic part of the molecule. Using this facility will prevent spurious peaks caused by diffraction ripples around the metal being interpreted as atoms.

MOLE - this is specified if molecular

connectivity is to be input by the user (see below).

END - this terminates the free format read and must be present if any keyword has been specified.

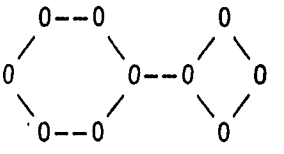
If 'ATOMS' is specified, the card following the one containing 'END' has the run time format for reading the n sets of atomic coordinates (e.g., (3F10.5)). Then follow n cards with the atomic coordinates in that format. This facility will normally be used when a difference Fourier is being interpreted. The inclusion of the subtracted atoms ensures that they are included in the bond-lengths and angles calculations, and also that they appear in the lineprinter picture of the structure.

If 'MOLE' is specified, free format cards after the 'END' keyword, (and after the atom cards above if they are present), specify the molecular connectivity. The bond codes allow the program to compare the molecular fragment it finds with that input here. They are not used by the program in its search for a fragment and are therefore optional.

The bond codes for any molecule are found as follows:

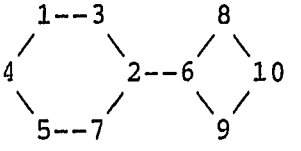
(a) Draw the molecule or fragment whose chemical structure (not necessarily geometry) is known.

e.g.,



(b) Number the atoms in any order

e.g.,



(c) Specify the connections to each atom as follows:

1 3 4/2 3 6 7/3 1 2/4 1 5/5 4 7/6 2 8 9/7 2 5/8 6 10/9 6 10/10 8 9//

which means atom 1 is bonded to atoms 3 and 4; 2 is bonded to 3, 6 and 7, etc.

Redundant information may be removed from the bond sequence

e.g., 1 3 4/2 3 6 7/4 5/5 7/6 8 9/8 10/9 10//

or 1 3 4/2 3 6 7/5 4 7/8 6 10/9 6 10//

in which each bond is specified only once.

(d) Type any of these descriptions in free format, ending the information with //. The maximum number of atoms in the molecule is 200. Up to five molecules can be input in this way and the complete data is terminated with ///

QPLANES.INP

It is possible for the program to generate more than one plot per run (simply by having more than one SYMMETRY

keyword), but all non-SYMMETRY keywords after the first SYMMETRY keyword will be ignored (i.e., if the first plot is drawn in color with solid lines, so will the rest of them).

2THETA x x - the minimum and maximum 2θ values. Only the G_h which lie in this range will be used to calculate $Q(\hat{S},t)$. [min=0.0, max=50.0]

FRACTIONG x - the fraction of the largest $|G_h|$ to be used in calculating $Q(\hat{S},t)$. [1.0]

COLOR - indicates colors are to be used to differentiate between the symmetry planes. This is the default choice and inclusion of this keyword implies that the curves will be drawn using solid lines. The colors and the corresponding symmetry planes are:

mirror => red (pen 2)
b/c/a -glide => green (pen 3)
c/a/b -glide => yellow (pen 4
n-glide => blue (pen 5)
d-glide => orange (pen 6)

BLACK - indicates that all curves are to be drawn as black lines. Inclusion of this keyword implies that the the curves will be drawn using different linestyles.

LINETYPES - indicates that the curves will be drawn

using different line types. The line types and the corresponding symmetry planes are:

mirror => solid
b/c/a -glide => dashed
c/a/b -glide => dash-do
n-glide => long dash-short dash
d-glide => long dash-short dash-short dash

SOLID - indicates that the curves will be drawn using solid lines. This is the default choice.

SYMMETRY c - the axis to which the symmetry planes are perpendicular.

QROTLINE.INP

2THETA x x - the minimum and maximum 2θ values. Only the G_h which lie in this range will be used to calculate $Q(\hat{S},t)$. [min=0.0, max=50.0]

FRACTIONG x - the fraction of the largest $|G_h|$ to be used in calculating $Q(\hat{S},t)$. [1.0]

SYMMETRY c - the symmetry operator for which the $Q(\hat{S},t)$ values are to be calculated. [-X,Y,-Z]

FIX c - indicates which coordinate is to be kept fixed. [Z if rotation axis is parallel to a or b, Y if parallel to c]

LOCATION x - indicates the value of the fixed coordinate. [0.25]

SHIFT x - indicates how much the varying coordinate will be shifted. This option is usually not used. [0.0]

QSPLOT.INP

It is possible for the program to generate more than one plot per run (simply by having more than one SYMMETRY keyword), but all non-SYMMETRY keywords after the first SYMMETRY keyword will be ignored (i.e., if the first plot uses 30% of the largest $|G_h|$, so will the rest of them).

2THETA x x - the minimum and maximum 2θ values. Only the G_h which lie in this range will be used to calculate $Q(\hat{S},t)$ and $S(\hat{S},t)$. [min=0.0, max=50.0]

FRACTIONG x - the fraction of the largest $|G_h|$ to be used in calculating $Q(\hat{S},t)$ and $S(\hat{S},t)$. [1.0]

SYMMETRY c - the symmetry operator for which the $Q(\hat{S},t)$ and $S(\hat{S},t)$ values are to be

calculated. [-X,Y,-Z]

NORMAL - indicates that the normal $S(\hat{S},t)$ expression (equation (3.9)) will be used. Inclusion of this keyword is optional as it is the default.

MIRROR - indicates that the "mirrored" form of the $S(\hat{S},t)$ expression (equation (3.21)) will be used.

QDATA.INP

It is possible for the program to generate more than one data set per run (simply by having more than one SYMMETRY keyword), but all non-SYMMETRY keywords after the first SYMMETRY keyword will be ignored (i.e., if the first plot uses 30% of the largest $|G_h|$, so will the rest of them).

2THETA x x - the minimum and maximum 2θ values. Only the G_h which lie in this range will be used in the calculation. [min=0.0, max=50.0]

FRACTIONG x - the fraction of the largest $|G_h|$ to be used in the calculation. [1.0]

QFUNCTN - indicates that $Q(\hat{S},t)$ values for the symmetry operator will be calculated. Inclusion of this keyword is optional as

it is the default.

SFUNCTN - indicates that $S(\hat{S},t)$ values for the symmetry operator will be calculated.

NORMAL - indicates that the normal $S(\hat{S},t)$ expression (equation (3.9)) will be used. This keyword is ignored unless SFUNCTN is also specified.

MIRROR - indicated that the "mirrored" form of the $S(\hat{S},t)$ expression (equation (3.21)) will be used. This keyword is ignored unless SFUNCTN is also specified.

SYMMETRY c - the symmetry operator for which the values are to be calculated. [-X,Y,-Z]

Sample Input Files

Sample input files for two of the programs, LOCSYM and FOUR, are listed below. This were used in the solution of $Cr(CO)_3 \cdot C_6H_6$ (Chapter 5).

LOCSYM.INP

```
2THETA 0.0 50.0
LAUE 2
WEIGHTS UNIT   FRACTIONG 0.300
NSUPERPOS 1    WTSUPERPOS
```

```
SYMMETRY
 -X,Y+1/2,-Z                diad on b
*X,-Y,Z                     inv. diad (mirror) on b
 -X,-Y,-Z                   inversion center
```

FOURIER.INP

```
CENTRO   MAP 3   ORIENT 2
OUTPUT BOTH   GRID 32 64 32
ACROSS FULL   DOWN FULL   LAYER 0 32
FIELDWIDTH 4   BLANKS 0   LOWVALUE 1
LAUE 2   SCALE 0.45634   F000 0.000   OFFSET 0 0
LATTICE P
SYMMETRY 2
X,Y,Z
-X,1/2+Y,-Z
```

APPENDIX B. CRYSTAL STRUCTURE DETERMINATIONS

Although the major emphasis of this research has been involved with the development of a new structural solution strategy, crystallographic studies of several compounds have been undertaken. This Appendix contains detailed discussions of these studies.

Structure Determination of $(NHC_5H_5)SbI_4$

Introduction

The crystal structure of pyridinium antimony(III) tetraiodide, $(NHC_5H_5)SbI_4$, has been determined as part of an ongoing series of investigations into halo-coordinated antimony compounds. Previously, the structures of chloro- and bromo-coordinated single- and mixed-valence state antimony anions in the presence of various cationic species had been determined using X-ray diffraction. In these structures, the antimony ions coordinated to the halogens exist at the center of octahedral units or units best described as fragments of an octahedron[38-50]. In some of the compounds, the trihalide and halide ions, as well as the halogen molecule, coexist in the crystal with the antimony-centered units, forming a bridge between the separated antimony-halide units[41,43,44,47]. However, this compound,

like its chloro- and bromo-analogues[39,46], forms infinite chains involving the SbI_4^- anion via halide linkages.

Experimental

Crystal Data $(C_5H_5NH)SbI_4$, M = 709.48, orthorhombic, a = 25.156(6), b = 13.960(2), c = 7.838(1) Å, V = 2752.5 $Å^3$, space group Pbca from systematic absences 0kl:k=2n+1, h0l:l=2n+1 and hk0:h=2n+1, D_c = 3.42 g cm^{-3}, Z = 8, μ(Mo) = 108.6 cm^{-1}.

Preparation 3.48 g of SbI_3 (s) were dissolved in 16.92 ml of concentrated HI (aq). 1.0 ml of pyridine was added to the solution producing a thick black precipitate. The mixture was vacuum filtered and the precipitate was dried overnight in a desiccator using 40 ml of concentrated H_2SO_4 as the desiccant. The precipitate was recrystallized in hot concentrated HI (aq), allowed to cool overnight and then placed in an ice bath for several hours until orange crystals had formed.

Crystallographic Studies For data collection, a crystal of approximate dimensions 0.20 x 0.20 x 0.15 mm was mounted in a 0.2 mm diameter thin-walled capillary tube, sealed with wax, and attached to a standard goniometer head. The crystal was centered on a four-circle DATEX X-ray diffractometer. Four preliminary ω-oscillation photographs were taken at various ϕ settings. The approximate positions of 12 reflections were selected from photographs and used as

input to an automatic indexing program[51]. The resulting reduced cell and reduced cell scalars indicated a primitive orthorhombic lattice. The predicted layer spacings for this cell were observed, within experimental error, on three axial ω-oscillation photographs.

Intensity data were collected at room temperature (20°C) using Mo K_α (λ = 0.71069 Å) radiation on the above mentioned diffractometer equipped with a graphite monochromating crystal and interfaced to a VAX 730 computer in a real-time mode. A total of 2846 intensities were measured, corresponding to reflections in octants hkl and $h\bar{k}\bar{l}$, using an ω-step scan technique with a scan half-width of 0.5 degree. The data were collected to a maximum $\sin\theta/\lambda$ of 0.5; an additional shell of data with $0.5 < \sin\theta/\lambda \leq 0.6$ was collected but was found to be quite weak on the average, and was not used in the final refinement of the structure (see below). As a check on the stability of the crystal, the intensities of seven standard reflections were measured every 50 reflections during data collection. No significant decay was observed.

Accurate unit cell parameters were obtained by carefully centering 25 independent reflections, having $16° \leq 2\theta \leq 33°$, on the diffractometer and inputing the $\pm 2\theta$ values of these reflections into a least-squares fitting program. Intensity data were corrected for absorption (T_{min}/T_{max} = 0.045/0.082), as well as for Lorentz-polarization effects.

Of the 1886 reflections measured having $\sin\theta/\lambda \leq 0.5$, the 1039 that had $I \geq 3\sigma(I)$ were considered to be observed. The symmetry equivalent reflections were averaged ($R_{av} = 0.049$), yielding 607 reflections which were used in the refinement of the structure.

Solution and Refinement

The position of the antimony atom was determined from analysis of a sharpened Patterson map. The positions of the remaining non-hydrogen atoms were determined from successive structure factor[52] and electron density map calculations[53]. The positional and isotropic thermal parameters of the antimony and iodine atoms were initially refined using a block-matrix least-squares procedure. The anisotropic thermal parameters of the antimony and iodine atoms, the isotropic thermal parameters of the pyridine ring, and the positional parameters of all the non-hydrogen atoms were refined to their final value using a full-matrix least-squares procedure, minimizing the function $\Sigma w(|F_o|-|F_c|)^2$ to a conventional residual index of $R = (\Sigma||F_o|-|F_c||)/\Sigma|F_o| = 0.039$ and a weighted residual index of $R_w = [\Sigma w(|F_o|-|F_c|)^2/\Sigma w|F_o|^2]^{1/2} = 0.039$, where $w = 1/\sigma^2(F)$. The positions of the hydrogen atoms were calculated and not refined. A secondary extinction correction was applied and gave a coefficient of $\gamma = 5.303 \times 10^{-6}$. The final positional and averaged isotropic thermal parameters, with

their standard deviations, of the SbI_4^- anion and the pyridinium ring are given in Tables B.1 and B.2, respectively. The final anisotropic thermal parameters of the SbI_4^- anion are given in Table B.3, and bond distances and angles are given in Table B.4. Refinement was also carried out with inclusion of data with $\sin\theta/\lambda > 0.5$, but did not give improved standard deviations of the parameters ($R = 0.049$).

Discussion

The environment of the antimony atom in $(C_5H_5NH)SbI_4$ is shown in Figure B.1; infinite chains are formed similar to those found in the chloro and bromo analogues, even though the crystal system of the iodo compound is orthorhombic and the chloro and bromo compounds crystallize in the monoclinic system. A view of several unit cells is shown in Figure B.2. Each antimony lies at the center of a distorted octahedral structure, with all of the I-Sb-I bond angles within five degrees of 90°. The Sb-I3 and Sb-I4 bond distances, 2.809(4) Å and 2.828(4) Å respectively, are slightly longer than the 2.67Å Sb-I bond distances found in SbI_3 by electron diffraction techniques[38]. I1 and I2 are bridging iodines, connecting neighboring SbI_4^- moieties to form infinite chains via somewhat asymmetric bridges. The shorter Sb-I1 and Sb-I2 bridging distances are 3.021(4) and

Table B.1. Positional[a] ($x10^4$) and Averaged Isotropic Thermal ($Å^2$, $x10^3$) Parameters for $(C_5H_5NH)SbI_4$

atom	x	y	z	U_{av}[b]
Sb	1268(1)[c]	1614(1)	-7978(3)	57
I1	1954.6(9)	1700(2)	-4820(3)	73
I2	565.1(9)	1641(2)	-11129(3)	68
I3	1976(1)	245(2)	-9290(3)	84
I4	579(1)	231(2)	-6544(3)	84

[a]Positional parameters are given as fractions of a unit cell.

[b]$U_{av} = (U_{11} + U_{22} + U_{33})/3$. $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[c]Estimated standard deviations are given in parentheses for the least significant digits.

Table B.2. Positional[a] ($x10^4$) and Isotropic Thermal ($Å^2$, $x10^3$) Parameters for the Cation in $(C_5H_5NH)SbI_4$

atom	x	y	z	U[b]
N	1083(13)[c]	-791(22)	-12564(37)	89(11)
C2	1567(17)	-975(28)	-13148(53)	89(12)
C3	1669(16)	-1945(28)	-13146(53)	87(12)
C4	1332(18)	-2605(25)	-12670(50)	88(12)
C5	861(19)	-2328(33)	-12073(60)	119(17)
C6	746(15)	-1398(25)	-11984(48)	77(11)

[a]Positional parameters are given as fractions of a unit cell.

[b]$T = \exp[-8\pi^2U(\sin\theta/\lambda)^2]$.

[c]Estimated standard deviations are given in parentheses for the least significant digits.

Table B.3. Anisotropic Thermal Parameters[a] ($Å^2$, $x10^3$) for the Anion in $(NHC_5H_5)SbI_4$

atom	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
Sb	51(1)[b]	53(1)	67(2)	2(2)	-1(1)	0(2)
I1	63(2)	73(1)	83(2)	16(2)	14(1)	12(2)
I2	59(2)	69(2)	75(2)	12(2)	8(1)	3(2)
I3	66(2)	80(2)	107(2)	18(2)	4(2)	29(2)
I4	69(2)	78(2)	104(2)	-14(2)	-1(2)	-28(2)

[a] $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[b] Estimated standard deviations are given in parentheses for the least significant digit.

Table B.4. Selected Distances and Angles for $(C_5H_5NH)SbI_4$

Distances (Å)			
Sb-I1	3.021(4)*	Sb-$I1_a$	3.257(3)
Sb-I2	3.037(3)	Sb-$I2_b$	3.340(3)
Sb-I3	2.809(4)	Sb-I4	2.828(4)
I2-N1	3.807(31)		
Angles (deg)			
I1-Sb-$I1_a$	91.80(9)	I1-Sb-$I2_b$	85.31(9)
I1-Sb-I3	87.94(10)	I1-Sb-I4	92.96(10)
I2-Sb-$I1_a$	86.52(9)	I2-Sb-$I2_b$	92.05(8)
I2-Sb-I3	94.60(10)	I2-Sb-I4	88.56(10)
I3-Sb-$I1_a$	89.60(10)	I3-Sb-I4	93.99(10)
I4-Sb-$I2_b$	90.06(10)	$I1_a$-Sb-$I2_b$	86.92(8)
N1-H1-I2	153.07		

Subscripts refer to equivalent positions relative to the reference molecule.

(a)	x, 1/2 - y, z - 1/2
(b)	x, 1/2 - y, 1/2 - z

*Estimated standard deviations are given in parentheses for the least significant digits.

Table B.5. Positional[a] ($x10^4$) and Isotropic Thermal ($Å^2$, $x10^3$) Parameters for the Hydrogen Atoms in $(C_5H_5NH)SbI_4$

atom	x	y	z	U[b]
H1	993	5015	-7562	63.3
H2	1843	5411	-8618	63.3
H3	2060	7175	-8614	63.3
H4	1429	8391	-7676	63.3
H5	577	7907	-6630	63.3
H6	560	6180	-6526	63.3

[a]Hydrogen atoms were not refined.

[b]$T = \exp[-8\pi^2 U(\sin\theta/\lambda)^2]$.

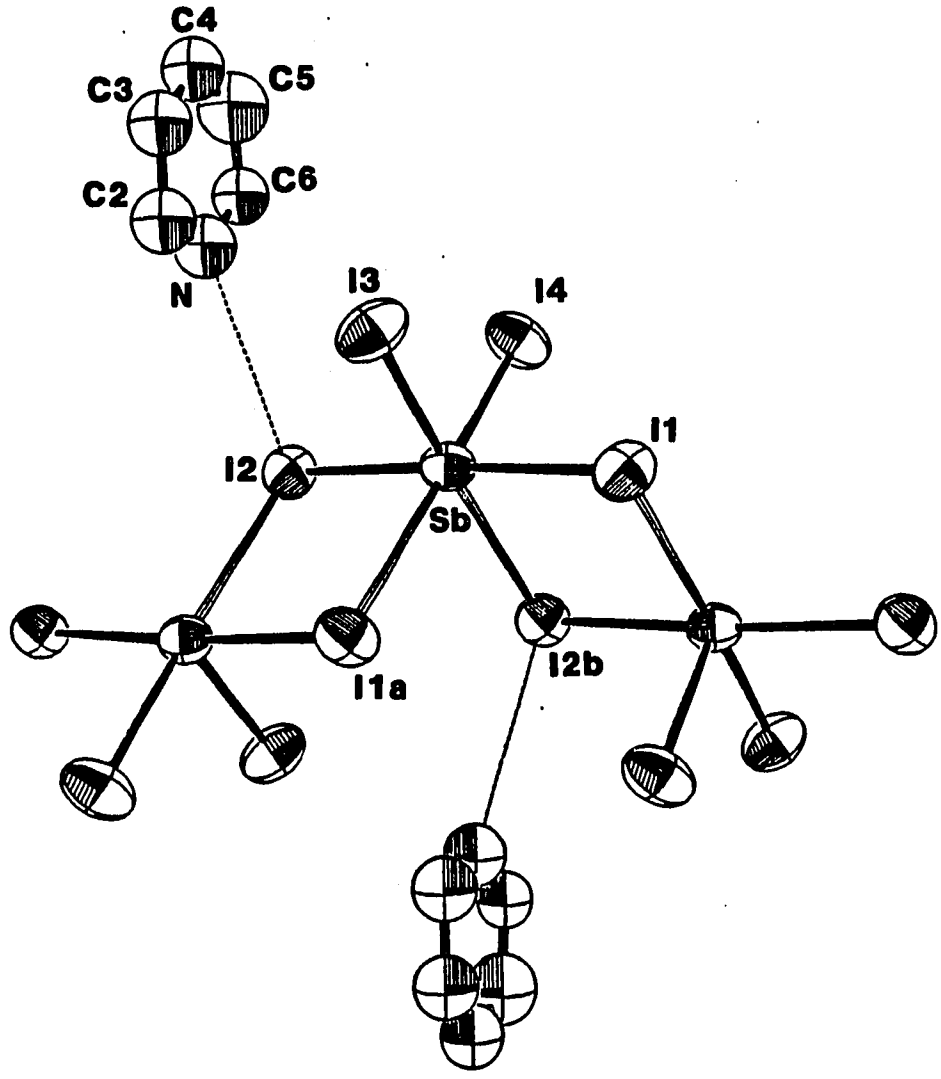


Figure B.1. View showing the immediate environment around the antimony atom in the structure of $(NHC_5H_5)SbI_4$. The bridging iodines are I1 and I2

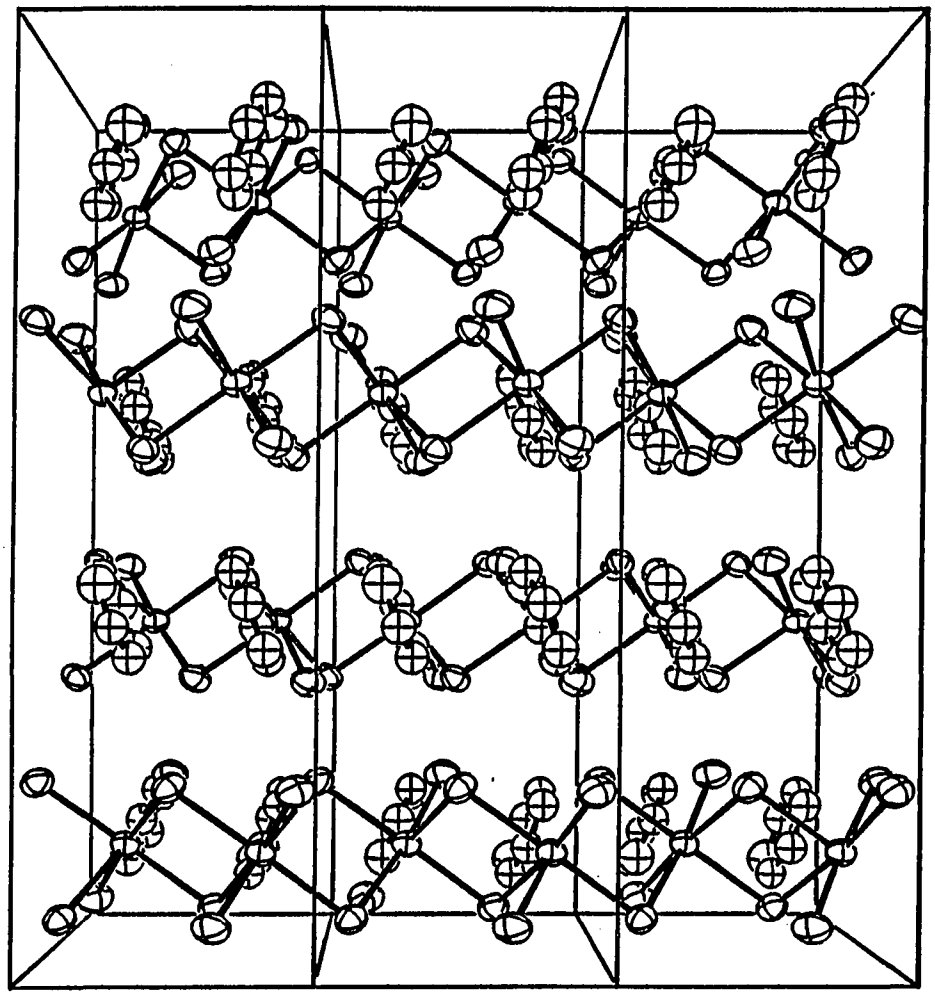


Figure B.2. Unit cell view of $(NHC_5H_5)SbI_4$ showing the infinite chain in the z-direction. In the figure, c is horizontal, a is vertical and b is perpendicular to the plane of the paper

3.037(3) Å, respectively, while the longer bridging distances are 3.257(3) and 3.340(3) Å. All are much shorter than the sum of the van der Waals radii, 4.35 Å. Examination of the differences in bond length between the long and short bridge bonds in the three antimony-halogen compounds that form infinite chains shows the following: chloro, $\Delta = 0.496$; bromo, $\Delta = 0.408$; and iodo, $\Delta = 0.269$ Å. Thus, there appears to be a clear trend toward less asymmetry in the bridge as the size of the halogen increases.

Because antimony halides often form as mixed-valence compounds, and it is difficult to distinguish pyridine from pyridinium by X-ray diffraction, one might speculate as to whether the antimony exists solely in the trivalent state or in a mixture of the tri- and pentavalent states. Using the approximation to the cation valence state given by Zachariasen[54], and taking the Sb-I bond distance of 2.67Å in SbI_3 as the distance corresponding to a unit bond strength distance, the charge on the antimony is calculated to be 2.94, implying that the antimony is in the trivalent state. Also, in charge-transfer complexes, where the metal ion exists in a mixture of tri- and pentavalent states, the crystals appear jet black under reflected light. However, the crystal under examination appears orange under reflected light, which further suggests that the antimony is not in a mixed-valent state. Finally, the fact that the reaction

mixture was an acidic medium suggests that the ring is the pyridinium ion as opposed to the neutral pyridine molecule. Thus, we can conclude that the antimony is in a trivalent state.

The Sb-I2 bridging distances are slightly longer than the Sb-I1 bridging distances, which is most likely a result of interactions between the pyridinium ion and I2. The nitrogen-hydrogen bond in the pyridinium ion points towards I2, and the I2-N1 distance of 3.807(31)Å is shorter than the nitrogen-iodine van der Waals distance of 3.85Å, suggesting that there is hydrogen bonding occurring between the I2 and the pyridinium ring.

It is interesting to note that Zachariasen's method predicts Sb-I bond strengths that agree quite well with the bond strengths that would be expected. Intuitively, the Sb-I bond strengths in the SbI_4^- anion would be expected to be 3/4 of that found in SbI_3. The calculated bond strengths range from 0.70 to 0.77 (Sb-I1: 0.48 + 0.29 = 0.77, Sb-I2: 0.46 + 0.24 = 0.70, Sb-I3: 0.75, Sb-I4 = 0.72). Because of the hydrogen bonding between I2 and the pyridinium ring, the Sb-I2 bond would be expected to be the weakest, and the calculation indicates that this is indeed the case.

Examination of the isotropic temperature factors of the pyridinium ion atoms shows that they are quite large. This is consistent with the observed fall off in intensity of

data having $\sin\theta/\lambda > 0.5$, since thermal vibration effects are more pronounced at large $\sin\theta/\lambda$ values.

Structure Determination of $PO_4C_{24}H_{25}$

Experimental

Crystal Data $PO_4C_{24}H_{25}$, M = 408.43, monoclinic, a = 16.189(5), b = 14.186(2), c = 19.135(8) Å, β = 114.97(3)°, V = 4138.7 $Å^3$, space group $P2_1/a$ from systematic absences h01:h=2n and 0k0:k=2n, D_c = 1.311 g cm^{-3}, Z = 8, μ(Mo) = 1.543 cm^{-1}.

Crystallographic studies Crystals of this compound were prepared in Dr. Verkade's group (Department of Chemistry, Iowa State University). For data collection, a crystal of approximate dimensions 0.1 x 0.1 x 0.1 mm was mounted in a 0.2 mm diameter thin-walled capillary tube, sealed with wax, and attached to a standard goniometer head. The crystal was aligned on a four-circle DATEX X-ray diffractometer. Four preliminary ω-oscillation photographs were taken at various ϕ settings. The approximate positions of 15 reflections were selected from the photographs and used as input to an automatic indexing program[51]. The resulting reduced cell and reduced cell scalars indicated a primitive monoclinic lattice. The predicted layer spacings for this cell were observed, within experimental error, on three axial ω-oscillation photographs.

Intensity data were collected at room temperature (20°C) using Mo K_α (λ = 0.70169 Å) radiation on the above mentioned diffractometer equipped with a graphite monochromating crystal and interfaced to a VAX 370 computer in a real-time mode. A total of 4768 intensities were measured, corresponding to reflections in octants hkl and $\bar{h}$kl, using an ω-step scan technique with a scan half-width of 0.5 degree. The data were collected to a maximum $\sin\theta/\lambda$ of 0.6. As a check on the stability of the crystal, the intensities of seven standard reflections were measured every 50 reflections during data collection. No significant decay was observed.

Accurate unit cell parameters were obtained by carefully centering ten independent high-angle reflections on the diffractometer noted above and inputing the $\pm 2\theta$ values of these reflections into a least-squares fitting program. Intensity data were corrected for Lorentz-polarization effects, but not for absorption due to the low value of μ. Of the 4768 reflections measured, the 2808 that had $I \geq 3\sigma(I)$ were considered to be observed. The symmetry equivalent reflections were averaged (R_{av} = 0.028), yielding 2649 reflections which were used in the refinement of the structure.

Solution and Refinement

The positions of the non-hydrogen atoms were determined as follows: Patterson superpositions were carried out using (i) the a-glide phosphorus-phosphorus Harker vector, followed by (ii) the 2_1 phosphorus-phosphorus Harker vector. The resulting superposition map showed a quartet of large peaks, which were taken to be phosphorus atoms. Further analysis showed a second quartet of peaks, also quite large in the original Patterson, which had the same relationships amongst themselves as did the initial quartet of peaks. This implied the presence of two parallelograms of phosphorus atoms, having the vectors that were chosen for the superpositions in common. From these two quartets, the positions of two non-symmetry-related phosphorus atoms were determined. The positions of the remaining non-hydrogen atoms were determined from successive structure factor[52] and electron density map calculations[53].

The initial positional and isotropic thermal parameters were refined using a block-matrix least-squares procedure. The positional and anisotropic thermal parameters were refined to their final values using a full-matrix least-squares procedure, minimizing the function $\Sigma w(|F_o|-|F_c|)^2$ to a conventional residual index of $R = \Sigma(|(|F_o|-|F_c|)|)/\Sigma|F_o| = 0.047$ and a weighted residual index of $R_w = [\Sigma w(|F_o|-|F_c|)^2/\Sigma w|F_o|^2]^{1/2} = 0.047$, where $w = 1/\sigma^2(|F_o|)$. The positions of approximately two-thirds of the hydrogens were

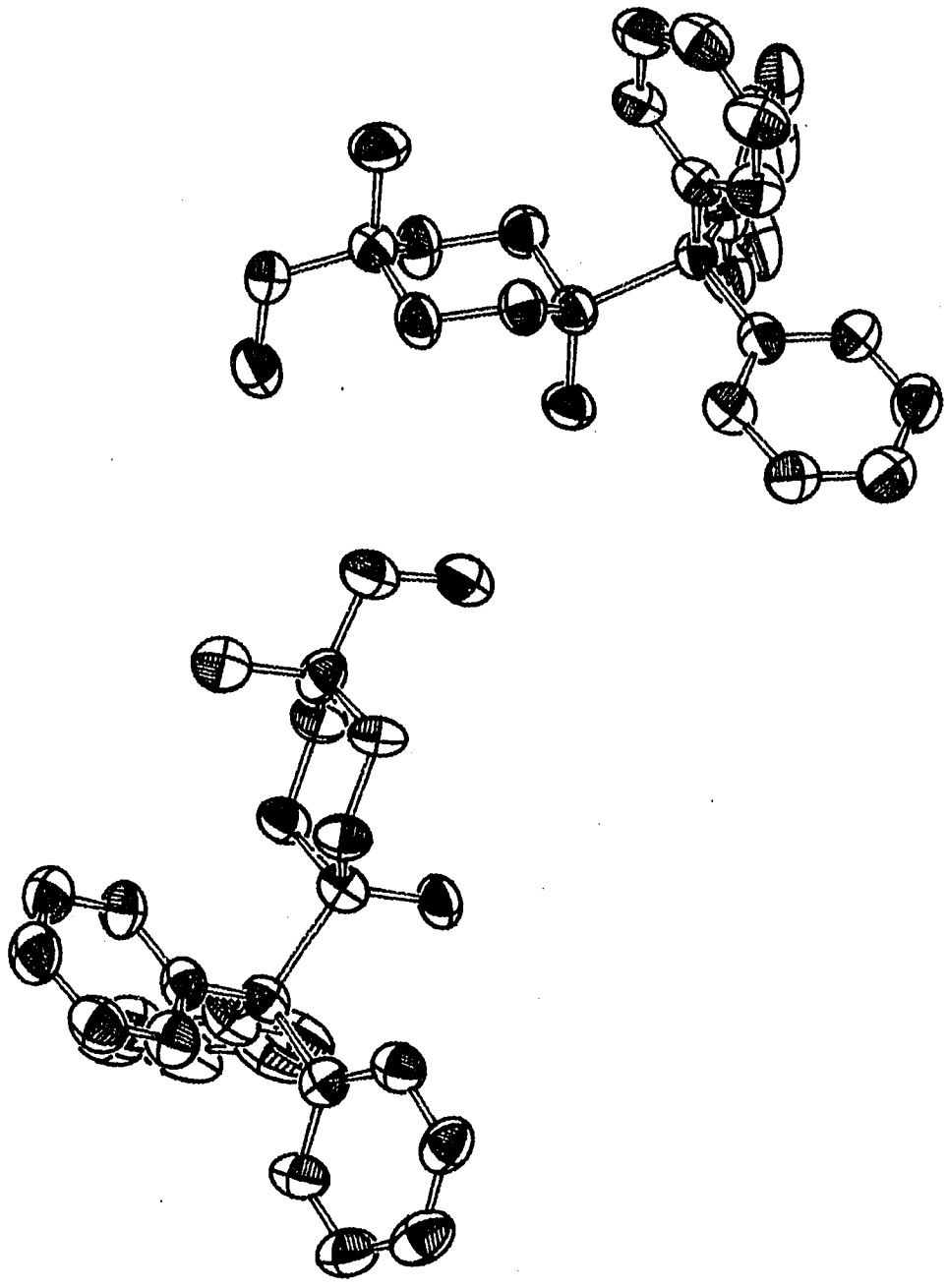


Figure B.3. ORTEP drawing for $PO_4C_{24}H_{25}$

Table B.6. Positional[a] ($x10^4$) and Averaged Isotropic Thermal ($Å^2$, $x10^3$) Parameters for $PO_4C_{24}H_{25}$

atom	x	y	z	U_{av}[b]
P1	9142(1)[c]	2327(1)	9657(1)	42
O1	9508(3)	2711(3)	9140(2)	52
O2	8674(3)	1336(3)	9378(2)	45
O3	8399(3)	2953(3)	9719(2)	48
O4	6707(3)	2536(4)	7402(2)	67
C1	9942(4)	2157(5)	10684(3)	40
C2	7857(4)	1370(4)	8669(4)	48
C3	7588(4)	3015(4)	8996(4)	49
C4	7183(4)	2041(5)	8740(4)	46
C5	6415(4)	2159(5)	7931(4)	55
C6	6806(5)	1644(5)	9287(4)	71
C11	9626(4)	1396(4)	11076(3)	41
C12	10140(5)	596(5)	11420(4)	53
C13	9845(6)	-75(5)	11785(4)	66
C14	9048(5)	29(5)	11837(4)	63
C15	8526(5)	813(6)	11494(4)	62
C16	8805(4)	1487(5)	11110(4)	50
C21	10830(4)	1839(5)	10683(4)	43
C22	10849(5)	1175(5)	10159(4)	51
C23	11656(5)	806(5)	10214(4)	55
C24	12443(5)	1149(5)	10797(5)	62
C25	12411(5)	1817(6)	11318(4)	57
C26	11622(5)	2162(5)	11268(4)	54
C31	10041(4)	3149(5)	11054(4)	45
C32	9868(4)	3293(6)	11706(4)	54
C33	9947(5)	4201(7)	12023(5)	65
C34	10173(6)	4948(6)	11690(5)	67
C35	10398(6)	4811(6)	11070(5)	69
C36	10311(5)	3907(5)	10756(4)	60
Pa	9621(1)	2326(2)	5366(1)	46
O1a	10497(3)	2680(4)	5885(3)	49
O2a	9384(3)	1346(3)	5637(3)	51

[a]Positional parameters are given as fractions of the unit cell.

[b]$U_{av} = (U_{11} + U_{22} + U_{33})/3$. $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[c]Estimated standard deviations are given in parentheses for the least significant digit.

Table B.6. (Continued)

atom	x	y	z	U_{av}
O3a	8854(3)	3002(3)	5322(3)	49
O4a	9358(3)	2542(4)	7635(3)	72
C1a	9430(5)	2156(5)	4342(4)	39
C2a	9260(5)	1395(5)	6358(4)	51
C3a	8724(5)	3061(5)	6042(4)	52
C4a	8550(5)	2088(6)	6296(5)	50
C5a	8528(5)	2198(7)	7093(4)	72
C6a	7641(5)	1718(6)	5743(5)	72
C11a	9241(6)	3155(6)	4004(5)	50
C12a	8453(6)	3417(6)	3381(5)	62
C13a	8299(7)	4357(8)	3120(5)	87
C14a	8940(10)	5037(8)	3472(7)	106
C15a	9742(8)	4779(8)	4054(6)	93
C16a	9890(7)	3851(7)	4315(5)	72
C21a	10269(5)	1783(5)	4300(5)	42
C22a	10455(6)	2005(6)	3672(4)	59
C23a	11175(6)	1624(7)	3600(5)	67
C24a	11730(6)	1005(7)	4150(6)	69
C25a	11541(5)	762(6)	4765(5)	58
C26a	10820(5)	1126(6)	4844(4)	51
C31a	8658(5)	1461(5)	3931(4)	42
C32a	8776(5)	630(6)	3601(5)	53
C33a	8083(7)	11(6)	3240(5)	62
C34a	7260(6)	203(7)	3204(5)	60
C35a	7142(6)	1023(6)	3553(5)	58
C36a	7842(6)	1665(6)	3921(4)	48

Table B.7. Anisotropic Thermal Parameters[a] ($Å^2$, $x10^3$) for the Non-Hydrogen Atoms in $PO_4C_{24}H_{25}$

atom	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
P1	43(1)[b]	45(1)	37(1)	-3(1)	14.6(9)	-3(1)
O1	56(3)	59(3)	41(3)	-10(3)	24(2)	1(2)
O2	47(3)	42(3)	44(3)	-4(2)	13(3)	-2(2)
O3	44(3)	53(3)	47(3)	1(2)	15(3)	-7(2)
O4	59(3)	101(4)	41(3)	-10(3)	12(3)	6(3)
C1	36(4)	50(5)	33(4)	7(4)	11(3)	2(4)
C2	48(5)	55(5)	40(4)	0(4)	9(4)	-12(4)
C3	49(4)	44(5)	54(5)	13(4)	7(4)	-4(4)
C4	46(4)	51(5)	42(4)	1(4)	18(4)	-2(4)
C5	39(4)	83(6)	44(4)	-6(4)	13(4)	-4(5)
C6	64(6)	88(6)	62(5)	-22(5)	37(5)	-9(5)
C11	55(5)	36(5)	34(4)	-5(4)	18(4)	-1(3)
C12	61(6)	43(4)	56(5)	2(4)	26(5)	9(4)
C13	87(7)	50(6)	60(6)	-3(5)	39(5)	2(4)
C14	86(6)	51(5)	52(5)	-13(5)	32(5)	0(4)
C15	69(6)	69(6)	49(5)	-13(5)	31(5)	-3(5)
C16	57(5)	55(5)	39(5)	1(4)	24(4)	-2(4)
C21	45(5)	44(5)	39(5)	-10(4)	14(4)	3(4)
C22	57(5)	50(5)	46(5)	2(4)	22(4)	1(4)
C23	56(5)	55(5)	55(5)	2(4)	27(4)	2(4)
C24	59(7)	63(6)	65(6)	11(5)	30(5)	25(4)
C25	48(5)	65(6)	59(5)	-7(4)	17(4)	9(5)
C26	44(4)	64(5)	54(5)	-6(4)	21(4)	-2(4)
C31	47(5)	47(5)	42(5)	-1(4)	8(4)	-5(4)
C32	47(5)	63(6)	52(5)	-4(4)	7(4)	-21(4)
C33	55(6)	79(7)	61(6)	15(5)	-1(4)	-21(6)
C34	84(7)	52(7)	67(7)	17(5)	-21(5)	-14(5)
C35	106(7)	40(6)	61(6)	-11(5)	-7(5)	7(5)
C36	77(6)	53(5)	51(5)	-15(4)	10(4)	8(5)
Pa	53(1)	44(1)	40(1)	0(1)	21(1)	0(1)
O1a	46(3)	65(4)	37(3)	-14(3)	9(3)	-8(3)
O2a	60(4)	42(3)	49(4)	-1(3)	32(3)	3(3)
O3a	59(4)	48(4)	40(3)	8(3)	25(3)	7(3)

[a] $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[b] Estimated standard deviations are given in parentheses for the least significant digits.

Table B.7. (Continued)

atom	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
O4a	70(4)	100(4)	47(4)	-13(4)	28(3)	-10(4)
C1a	51(5)	32(5)	35(4)	8(4)	16(4)	6(4)
C2a	63(6)	44(6)	45(5)	-1(5)	36(5)	2(4)
C3a	64(6)	43(5)	48(5)	1(5)	29(5)	-3(4)
C4a	40(5)	51(6)	60(6)	-13(5)	20(5)	-5(5)
C5a	70(6)	97(8)	48(5)	-12(6)	32(5)	-5(6)
C6a	68(6)	86(7)	62(6)	2(5)	31(5)	-8(5)
C11a	76(7)	31(5)	43(6)	1(5)	25(5)	-7(5)
C12a	87(7)	45(7)	55(6)	18(6)	35(6)	17(5)
C13a	118(9)	80(7)	62(7)	38(7)	49(7)	19(6)
C14a	198(14)	44(8)	76(9)	17(9)	78(9)	8(7)
C15a	164(12)	53(8)	64(7)	-27(7)	61(8)	-13(6)
C16a	114(9)	53(7)	50(6)	-17(7)	46(6)	-12(6)
C21a	44(5)	32(5)	51(6)	-10(4)	20(5)	-16(4)
C22a	68(7)	60(6)	49(6)	-6(5)	35(5)	-3(5)
C23a	62(7)	69(7)	69(7)	-7(6)	41(6)	-1(6)
C24a	59(7)	58(7)	89(7)	-9(6)	42(6)	-21(6)
C25a	45(6)	57(6)	72(7)	-8(5)	24(5)	-18(5)
C26a	59(6)	47(6)	48(6)	-12(5)	25(5)	-8(5)
C31a	44(6)	41(6)	41(5)	-11(5)	12(4)	-1(5)
C32a	59(6)	44(5)	57(6)	-16(5)	22(5)	-12(5)
C33a	75(7)	57(7)	54(6)	-9(6)	26(6)	-13(5)
C34a	54(6)	64(7)	61(6)	-18(6)	14(5)	-10(5)
C35a	49(7)	66(6)	59(6)	9(5)	18(5)	2(5)
C36a	42(6)	48(6)	52(5)	-8(5)	11(5)	-6(5)

Table B.8. Positional[a] ($x10^4$) and Thermal ($Å^2$, $x10^3$) Parameters for Hydrogen Atoms in $PO_4C_{24}H_{25}$

atom	x	y	z	U[b]
H1	7887	1402	8142	76.0
H2	7553	793	8500	76.0
H3	7740	3286	8566	76.0
H4	7303	3572	9083	76.0
H5	5883	2493	8024	76.0
H6	6013	1578	7773	76.0
H71	6340	2041	9293	76.0
H81	7282	1601	9815	76.0
H91	6580	1000	9122	76.0
H72	7129	1054	9528	76.0
H82	6887	2095	9698	76.0
H92	6186	1492	9004	76.0
H10	6263	2459	6847	76.0
H12	10802	522	11388	76.0
H13	10260	-683	12055	76.0
H14	8822	-507	12127	76.0
H15	7899	882	11526	76.0
H16	8392	2113	10837	76.0
H22	10216	924	9703	76.0
H23	11665	267	9795	76.0
H24	13086	854	10831	76.0
H25	13027	2062	11769	76.0
H26	11600	2694	11679	76.0
H32	9649	2697	11963	76.0
H33	9779	4328	12535	76.0
H34	10238	5680	11931	76.0
H35	10626	5425	10806	76.0
H36	10495	3807	10254	76.0
H1a	9250	633	6528	76.0
H2a	9831	1579	6807	76.0
H3a	8181	3580	5902	76.0
H4a	9262	3354	6465	76.0
H5a	8375	1563	7251	76.0
H6a	8033	2645	7039	76.0
H7a1	7160	2139	5722	76.0
H8a1	7523	1081	5893	76.0
H9a1	7583	1663	5200	76.0
H7a2	7684	1117	5487	76.0
H8a2	7261	1591	6010	76.0

[a]Hydrogen positions were calculated but not refined.

[b]$T = \exp[-8\pi^2U(\sin\theta/\lambda)^2]$.

Table B.8. (Continued)

atom	x	y	z	U^b
H9a2	7321	2175	5318	76.0
H10a	9391	2333	8145	76.0
H12a	7924	2863	3091	76.0
H13a	7665	4561	2626	76.0
H14a	8827	5793	3266	76.0
H15a	10261	5323	4338	76.0
H16a	10543	3658	4808	76.0
H22a	10011	2516	3228	76.0
H23a	11323	1839	3098	76.0
H24a	12321	709	4072	76.0
H25a	11985	242	5199	76.0
H26a	10662	910	5346	76.0
H32a	9439	471	3598	76.0
H33a	8191	-656	2968	76.0
H34a	6692	-282	2932	76.0
H35a	6472	1202	3554	76.0
H36a	7731	2328	4202	76.0

determined from a difference electron density map and the rest of the hydrogen positions were calculated. None of the hydrogen positions were refined. The final positional and averaged isotropic thermal parameters for the non-hydrogen atoms are given in Table B.6, and the anisotropic thermal parameters for these atoms are given in Table B.7. Table B.8 contains the calculated positions and isotropic thermal parameters for the hydrogen atoms.

Structure Determination of $PO_2C_6H_{11}$

Experimental

Crystal Data $PO_2C_6H_{11}$, M = 146.13, orthorhombic, a = 10.2107(12), b = 10.5541(23), c = 6.4639(6) Å, V = 697.54 $Å^3$, space group $P2_12_12_1$ from systematic absences h00:h=2n, 0k0:k=2n and 00l:l=2n, D_c = 1.391 g cm^{-3}, Z = 4, μ(Mo) = 3.077 cm^{-1}.

Crystallographic studies Crystals of this compound were prepared in Dr. Verkade's group (Department of Chemistry, Iowa State University). For data collection, a cylindrical crystal of approximate dimensions 0.35 mm in length and 0.15 mm in diameter was mounted in a 0.2 mm diameter thin-walled capillary tube, sealed with wax, and attached to a standard goniometer head. The crystal was aligned on a four-circle SYNTEX X-ray diffractometer. A preliminary rotation photograph was taken. The approximate positions of 12

reflections were selected from the photograph, refined and used as input to an automatic indexing program[42]. The resulting reduced cell and reduced cell scalars indicated a primitive orthorhombic lattice. The predicted layer spacings for this cell were observed, within experimental error, on three axial ω-oscillation photographs.

Intensity data were collected at room temperature (20°C) using Mo K_α (λ = 0.70169 Å) radiation on the above mentioned diffractometer equipped with a graphite monochromating crystal and interfaced to a NOVA computer in a real-time mode. A total of 1867 intensities were measured, corresponding to reflections in octants hkl and $\bar{h}$kl, using an ω-step scan technique with a scan half-width of 0.5 degree. The data were collected to an maximum $\sin\theta/\lambda$ of 0.6. As a check on the stability of the crystal, the intensity of one standard reflection was measured every 75 reflections during data collection. No significant decay was observed.

Accurate unit cell parameters were obtained by carefully centering fifteen independent high-angle reflections on the diffractometer noted above and inputing the $\pm 2\theta$ values of these reflections into a least-squares fitting program. Intensity data were corrected for Lorentz-polarization effects, but not for absorption due to the low value of μ. Of the 1867 reflections measured, the 918 that had $I \geq 3\sigma(I)$

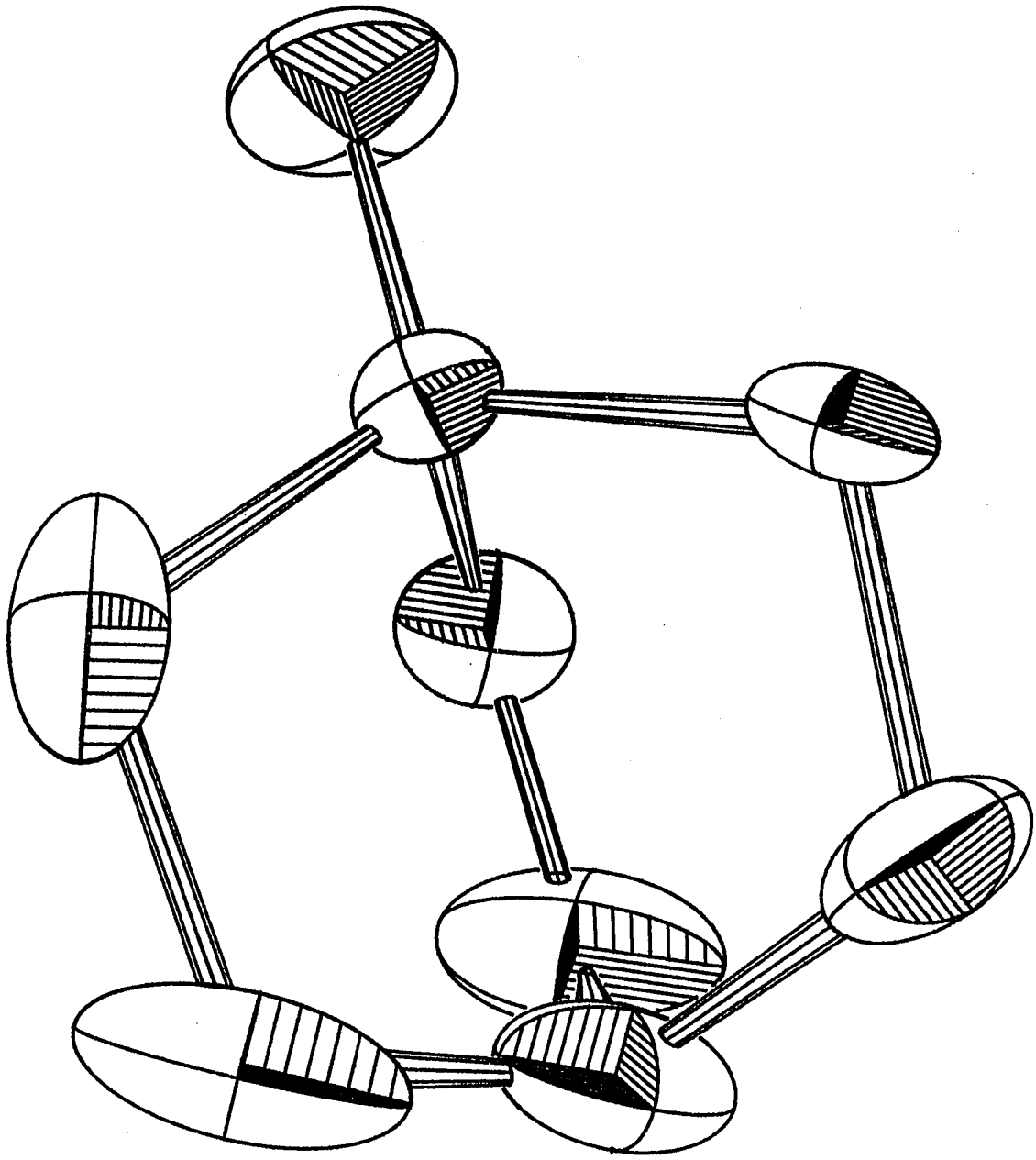


Figure B.4. ORTEP drawing for $PO_2C_6H_{11}$

Table B.9. Positional[a] ($x10^4$) and Averaged Isotropic Thermal ($Å^2$, $x10^3$) Parameters for $PO_2C_6H_{11}$

atom	x	y	z	U_{av}[b]
P1	5755(3)[c]	3377(3)	7840(4)	42
O1	6687(10)	3054(11)	9516(16)	88
O2	5555(9)	2227(6)	6243(14)	57
C1	6176(12)	4655(12)	6206(24)	67
C2	4116(15)	3795(11)	8546(16)	52
C3	4607(19)	2491(12)	4545(22)	92
C4	5091(16)	4745(12)	4469(20)	81
C5	3338(12)	3948(16)	6569(20)	79
C6	4088(16)	3752(11)	4669(17)	55

[a]Positional parameters are given as fractions of the unit cell.

[b]$U_{av} = (U_{11} + U_{22} + U_{33})/3$. $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[c]Estimated standard deviations are given in parentheses for the least significant digit.

Table B.10. Anisotropic Thermal Parameters[a] ($Å^2$, $x10^3$) for the Non-Hydrogen Atoms in $PO_2C_6H_{11}$

atom	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
P1	35(1)[b]	57(2)	35(1)	-2(2)	-8(2)	3(1)
O1	70(7)	130(8)	65(6)	16(6)	-35(6)	12(6)
O2	68(6)	48(4)	56(5)	14(4)	-3(5)	0(4)
C1	49(8)	65(8)	86(10)	-19(6)	17(8)	-2(8)
C2	48(7)	90(9)	16(5)	1(7)	-1(6)	-3(5)
C3	164(23)	68(8)	46(9)	0(10)	-35(11)	-21(7)
C4	130(14)	83(9)	30(7)	23(10)	-15(9)	13(7)
C5	33(5)	164(14)	39(7)	32(8)	-13(6)	-9(9)
C6	76(10)	64(9)	26(6)	24(8)	-3(7)	2(5)

[a] $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[b] Estimated standard deviations are given in parentheses for the least significant digits.

Table B.11. Positional[a] ($x10^4$) and Thermal ($Å^2$, $x10^3$) Parameters for Hydrogen Atoms in $PO_2C_6H_{11}$

atom	x	y	z	U[b]
H1	7127	4486	5469	38.0
H2	6201	5537	7018	38.0
H3	4101	4671	9411	38.0
H4	3679	3046	9492	38.0
H5	3882	1783	4538	38.0
H6	5202	2386	3067	38.0
H7	5609	4632	2974	38.0
H8	4669	5660	4542	38.0
H9	2914	4891	6542	38.0
H10	2533	3255	6581	38.0
H11	3236	3943	3392	38.0

[a]Hydrogen positions were calculated but not refined.

[b]$T = \exp[-8\pi^2 U(\sin\theta/\lambda)^2]$.

were considered to be observed. The symmetry equivalent reflections were averaged (R_{av} = 0.099), yielding 542 reflections which were used in the refinement of the structure.

Solution and Refinement

The position of the phosphorus atom was determined by examination of a Patterson map. The positions of the remaining non-hydrogen atoms were determined from successive structure factor[43] and electron density map calculations[51]. The initial positional and isotropic thermal parameters were refined using a block-matrix least-squares procedure. The positional and anisotropic thermal parameters were refined to their final values using a full-matrix least-squares procedure, minimizing the function $\Sigma w(|F_o|-|F_c|)^2$ to a conventional residual index of $R = \Sigma(|(|F_o|-|F_c|)|)/\Sigma|F_o| = 0.088$ and a weighted residual index of $R_w = [\Sigma w(|F_o|-|F_c|)^2/\Sigma w|F_o|^2]^{1/2} = 0.095$, where $w = 1/\sigma^2(|F_o|)$. The positions of the hydrogen atoms were calculated and not refined. The final positional and averaged isotropic thermal parameters for the non-hydrogen atoms are given in Table B.9, and the anisotropic thermal parameters for these atoms are given in Table B.10. Table B.11 contains the calculated positions and isotropic thermal parameters for the hydrogen atoms.

Structure Determination of $(NHC_5H_5)_2I_{10}$

Experimental

Crystal Data $(NHC_5H_5)_2I_{10}$, M = 1429.26, orthorhombic, a = 17.3305(37), b = 17.6146(26), c = 4.5336(7) Å, V = 1383.996 $Å^3$, space group $Pm2_1b$, systematic absences hk0:k=2n and 0k0:k=2n, D_c = 3.429 g cm^{-3}, Z = 2, μ(Mo) = 110.8 cm^{-1}.

Crystallographic studies For data collection, a crystal of approximate dimensions 0.35 x 0.15 x 0.15 mm was mounted in a 0.2 mm diameter thin-walled capillary tube, sealed with wax, and attached to a standard goniometer head. The crystal was aligned on a four-circle SYNTEX X-ray diffractometer. A preliminary rotation photograph was taken. The approximate positions of 13 reflections were selected from the photograph, refined and used as input to an automatic indexing program[42]. The resulting reduced cell and reduced cell scalars indicated a primitive orthorhombic lattice. The predicted layer spacings for this cell were observed, within experimental error, on three axial ω-oscillation photographs.

Intensity data were collected at room temperature (20°C) using Mo K_α (λ = 0.70169 Å) radiation on the above mentioned diffractometer equipped with a graphite monochromating crystal and interfaced to a NOVA computer in a real-time mode. A total of 1827 intensities were measured, corresponding to reflections in the hkl octant, using an ω-

step scan technique with a scan half-width of 0.5 degree. The data were collected to an maximum $\sin\theta/\lambda$ of 0.6. As a check on the stability of the crystal, the intensity of one standard reflection was measured every 75 reflections during data collection. No significant decay was observed.

Accurate unit cell parameters were obtained by carefully centering fifteen independent high-angle reflections on the diffractometer noted above and inputing the $\pm 2\theta$ values of these reflections into a least-squares fitting program. Intensity data were corrected for absorption (T_{min}/T_{max} = 0.180/0.219), as well as for Lorentz-polarization effects. Of the 1827 reflections measured, the 555 that had $I \geq 3\sigma(I)$ were considered to be observed.

Solution and Refinement

The correct space group and the locations of two iodines were determined as described in Chapter 5. The positions of the remaining non-hydrogen atoms were determined from successive structure factor[43] and electron density map calculations[51]. The initial positional and isotropic thermal parameters were refined using a block-matrix least-squares procedure. The positional and anisotropic thermal parameters were refined to their final values using a full-matrix least-squares procedure, minimizing the function $\Sigma w(|F_o|-|F_c|)^2$ to a conventional residual index of $R = \Sigma(|(|F_o|-|F_c|)|)/\Sigma|F_o| = 0.048$ and a weighted residual index

Table B.12. Positional[a] ($x10^4$) and Averaged Isotropic Thermal ($Å^2$, $x10^3$) Parameters for the Anion in $(NHC_5H_5)I_{10}$

atom	x	y	z	U_{av}[b]
I1	5000	5000	-635(18)[c]	77
I2	5000	1786(8)	4244(11)	43
I3	3385(2)	1834(5)	7402(8)	54
I4	1887(2)	1854(5)	9636(8)	65
I5	5000	3525(5)	1407(15)	60
I6	0	-83(6)	906(12)	48
I7	0	3313(5)	6146(15)	54
I8	0	1685(5)	3371(13)	45

[a]Positional parameters are given as fractions of the unit cell,

[b]$U_{av} = (U_{11} + U_{22} + U_{33})/3$. $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[c]Estimated standard deviations are given in parentheses for the least significant digit.

Table B.13. Anisotropic Thermal Parameters[a] ($Å^2$, $x10^3$) for the Anion in $(NHC_5H_5)_2I_{10}$

atom	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
I1	69(5)	60(5)	102(6)	0	0	-23(6)
I2	37(2)	51(3)	41(3)	0	0	0(5)
I3	47(2)	59(2)	57(2)	4(3)	-16(2)	2(3)
I4	44(2)	81(3)	70(3)	6(3)	-14(2)	-4(4)
I5	46(4)	72(5)	61(5)	0	0	-24(4)
I6	50(3)	41(3)	52(4)	0	0	-2(4)
I7	46(3)	60(4)	56(5)	0	0	13(4)
I8	35(2)	49(5)	52(4)	0	0	-2(4)

[a] $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

[b] Estimated standard deviations are given in parentheses for the least significant digits.

Table B.14. Positional[a] ($x10^4$) and Isotropic Thermal ($Å^2$, $x10^3$) Parameters for the Cation in $(NHC_5H_5)I_{10}$

atom	x	y	z	U[b]
N	1922(29)[c]	3915(31)	3014(121)	83(17)
C1	2751(33)	4883(44)	6197(120)	60(16)
C2	3028(31)	4232(43)	5210(209)	90(20)
C3	2623(37)	3627(36)	3437(135)	71(22)
C4	2082(43)	5109(44)	5145(198)	111(26)
C5	1645(27)	4601(30)	3635(123)	55(17)

[a]Positional parameters are given as fractions of a unit cell.

[b]$T = \exp[-8\pi^2U(\sin\theta/\lambda)^2]$.

[c]Estimated standard deviations are given in parentheses for the least significant digits.

Table B.15. Positional[a] ($x10^4$) and Thermal ($Å^2$, $x10^3$) Parameters for Hydrogen Atoms in $(NHC_5H_5)I_{10}$

atom	x	y	z	U[b]
H1	1589	3500	1724	50.7
H2	1063	4788	2991	50.7
H3	1864	5670	5806	50.7
H4	3085	5298	7487	50.7
H5	3610	4045	5854	50.7
H6	2841	3066	2776	50.7

[a]Hydrogen positions were calculated but not refined. Positional parameters are given as fractions of a unit cell.

[b]$T = \exp[-8\pi^2U(\sin\theta/\lambda)^2]$.

of $R_w = [\Sigma w(|F_o|-|F_c|)^2/\Sigma w|F_o|^2]^{1/2} = 0.048$, where $w = 1/\sigma^2(|F_o|)$. The positions of the hydrogen atoms were calculated and not refined. The final positional and averaged isotropic thermal parameters for the anion are given in Table B.12, and the anisotropic thermal parameters for these atoms are given in Table B.13. The positional and isotropic thermal parameters for the cation are given in Table B.14. Table B.15 contains the calculated positions and isotropic thermal parameters for the hydrogen atoms.

Twinning Effects in $YBa_2Cu_2O_{7-x}$

Crystallographic Studies

Crystals of the superconducting material were prepared by Dr. Gschneidner's group (Ames Laboratory, Iowa State University). Three apparently single-orientation crystals were mounted on glass fibers using Duco cement and attached to standard goniometer heads. The crystals were aligned on a four-circle DATEX diffractometer. Four preliminary ω-oscillation photographs were taken at various ϕ settings. The approximate positions of 10 reflections were selected from the photographs and used as input to an automatic indexing program[51]. The resulting reduced cell and reduced cell scalars indicated a primitive orthorhombic lattice. Additional reflections having $2\theta > 25°$ were used to determine a more accurate orientation matrix.

A computer program was written to scan a region of reciprocal space using a "step and count" technique. In this technique, the detector is moved through reciprocal space using a small stepsize, and the intensity is measured after each step for several seconds. For each crystal, regions of reciprocal space around the (220), the $(2\bar{2}0)$, the (400) and the (040) reflections were scanned using a stepsize of 0.01 reciprocal lattice units and a counting time of two seconds. These regions were scanned a minimum of four times and the results of the scans were averaged for each crystal. The plots of the averaged scans for a given reflection were similar for each crystal. Representative plots for the (220), $(2\bar{2}0)$, (400), and (040) reflections are given in Figures B.5, B.6, B.7, and B.8, respectively. It is quite obvious from the appearance of these plots that the crystalline samples of the material were not single-orientation and that a twinning process was occurring.

Twin Modelling

A computer program was written to generate the expected shape of a reflection for a given mode of twinning. The effect of instrument broadening was treated by expressing the peak shape as Gaussian,

$$I_h = I_{h_o} \exp(-\alpha(h - h_o)^2) \quad , \qquad (B.1)$$

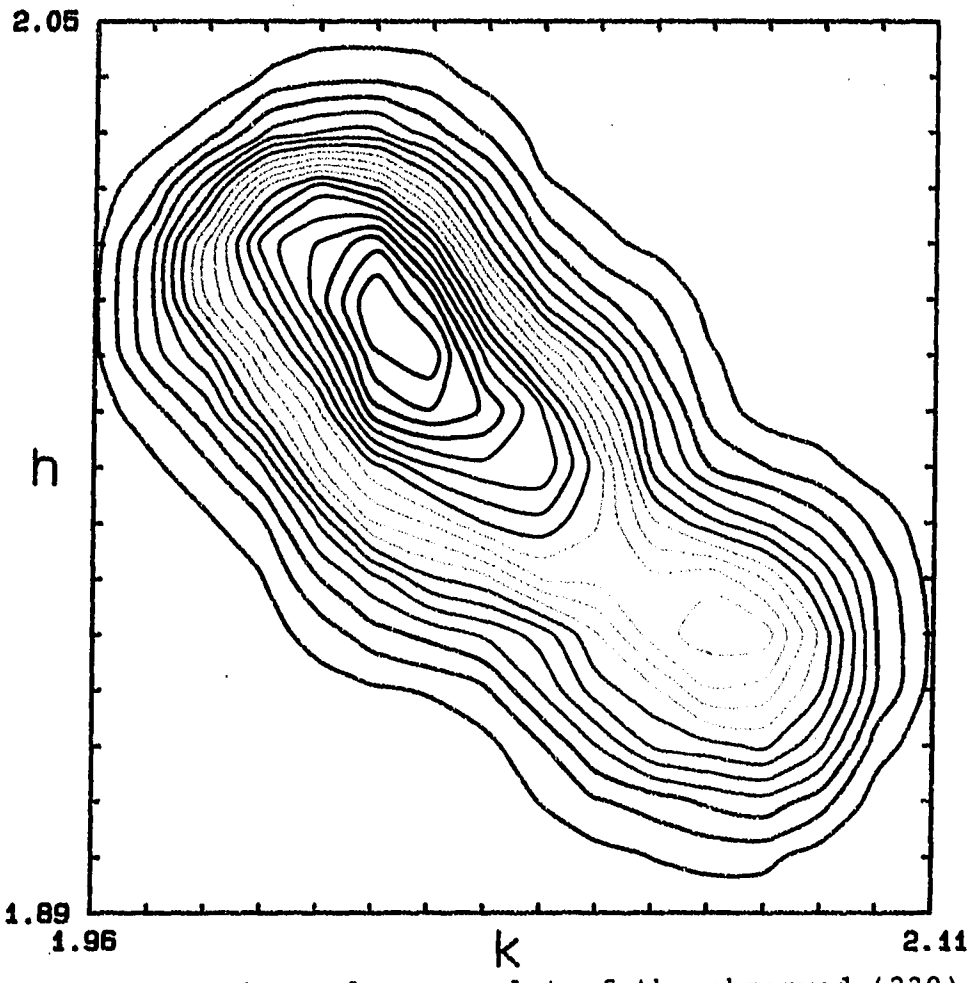


Figure B.5. Reciprocal space plot of the observed (220) reflection for $YBa_2Cu_3O_{7-x}$

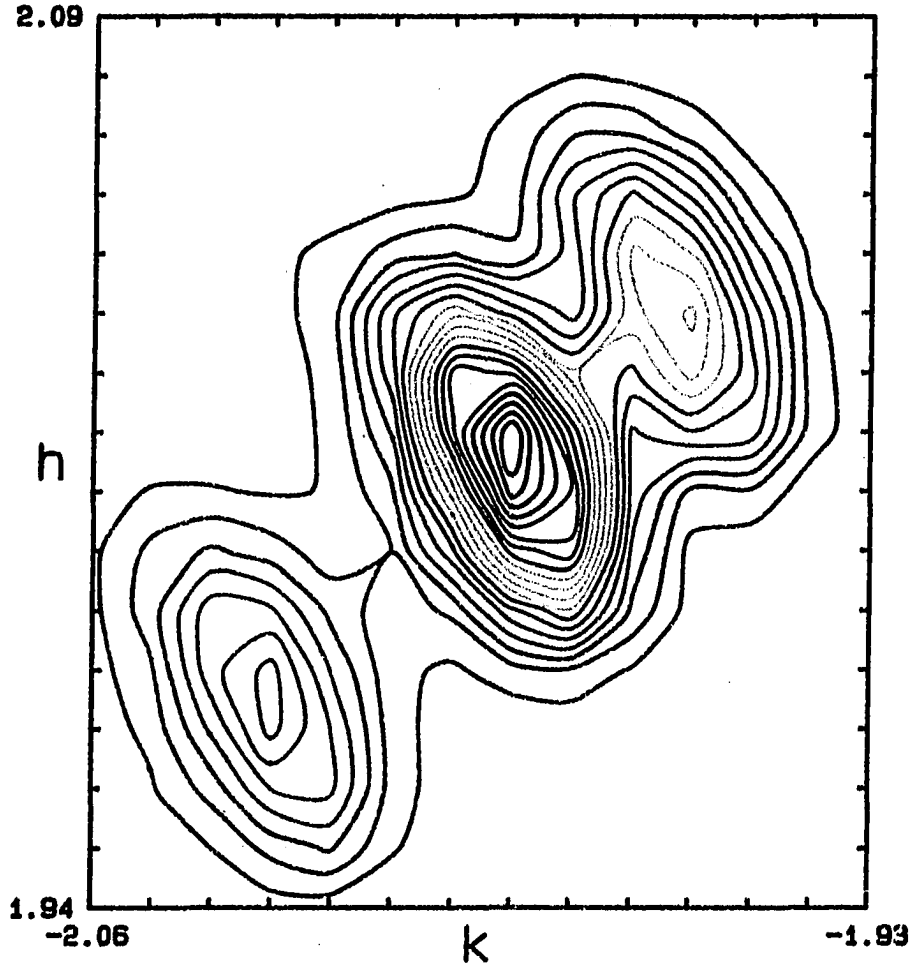


Figure B.6. Reciprocal space plot of the observed $(2\bar{2}0)$ reflection for $YBa_2Cu_3O_{7-x}$

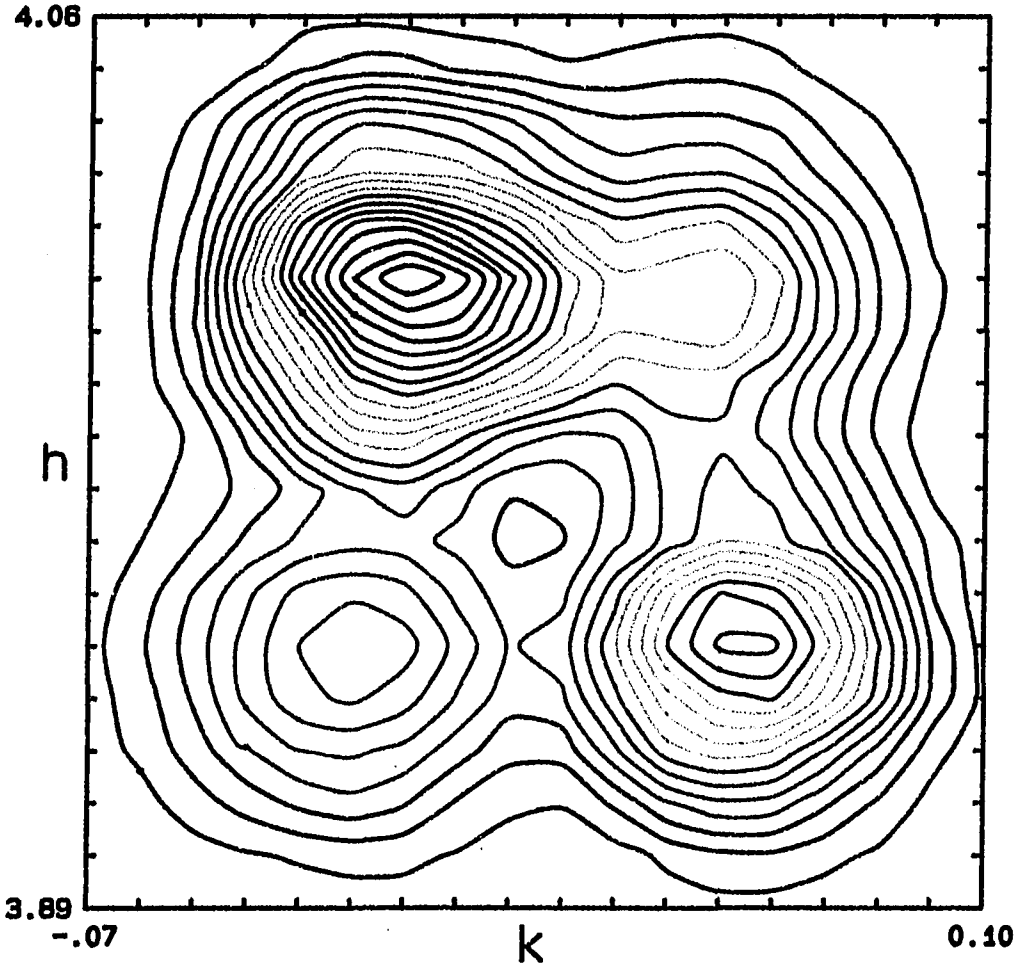


Figure B.7. Reciprocal space plot of the observed (400) reflection for $YBa_2Cu_3O_{7-x}$

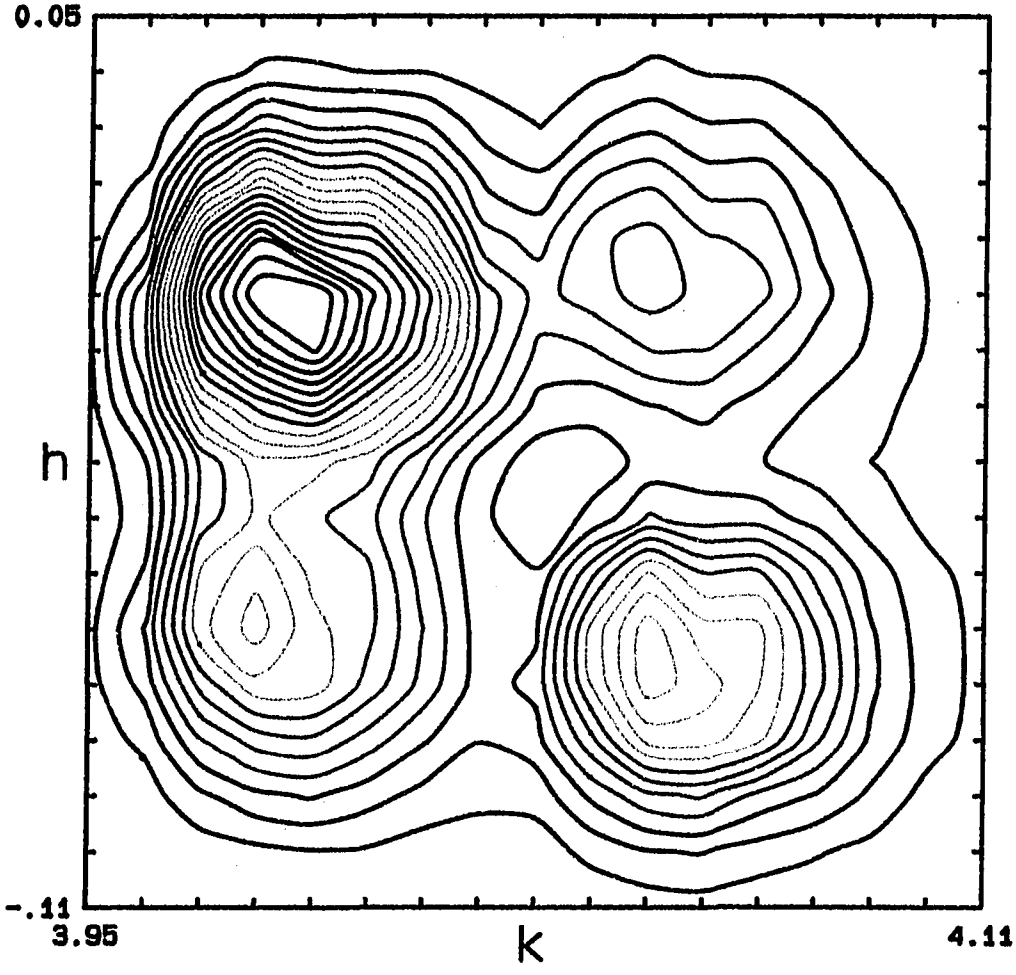


Figure B.8. Reciprocal space plot of the observed (040) reflection for $YBa_2Cu_3O_{7-x}$

where α is an adjustable parameter which determines the width of the peak. Three modes of twinning were modelled: (1) a mirror plane along $h\bar{h}0$, (2) mirror planes along hh0 and $h\bar{h}0$, and (3) a main crystal with a mirror along $h\bar{h}0$ with a smaller satellite, also having a mirror along $h\bar{h}0$, rotated 90° relative to the orientation of the main crystal. Plots of the (400), the (220), and the $(2\bar{2}0)$ reflections for each of the twinning modes are shown in Figures B.9 through B.17.

Discussion

At high temperatures, this compound exists in a tetragonal form, having mirrors along hh0 and $h\bar{h}0$ and a four-fold rotation axis. As it cools, it undergos a transformation to an orthorhombic phase. The accepted space group for this phase is Pmmm. This space group contains mirror planes along each axis, so that a twinning mode containing a mirror $h\bar{h}0$ must, by symmetry, have a mirror along hh0. In such a twinning mode, the (220) and the $(2\bar{2}0)$ reflections should have the same shape. However, examination of the observed reflections clearly indicates that this is not the case. The (400) reflection should have three major components, as shown in Figure B.10, but again, this is not the case.

The twinning model that gives the best agreement to the observed peak shapes has a main crystal with a mirror along

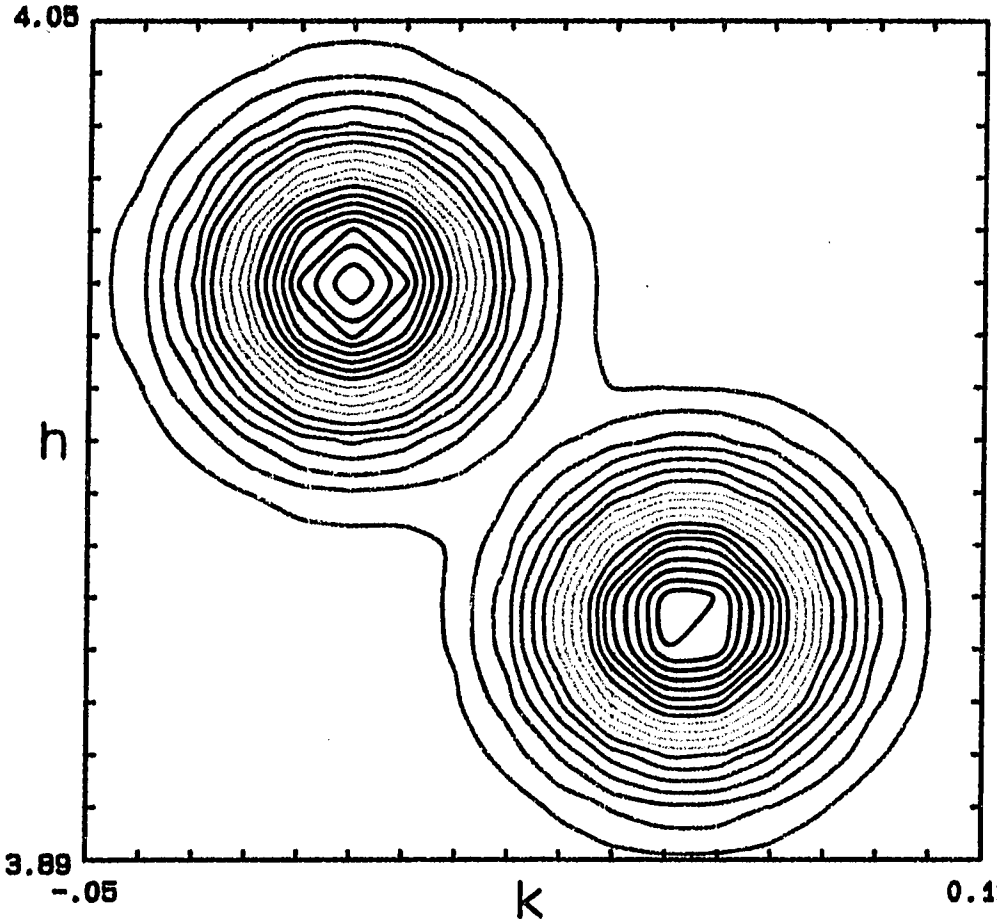


Figure B.9. Reciprocal space plot of the calculated (400) reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (1)

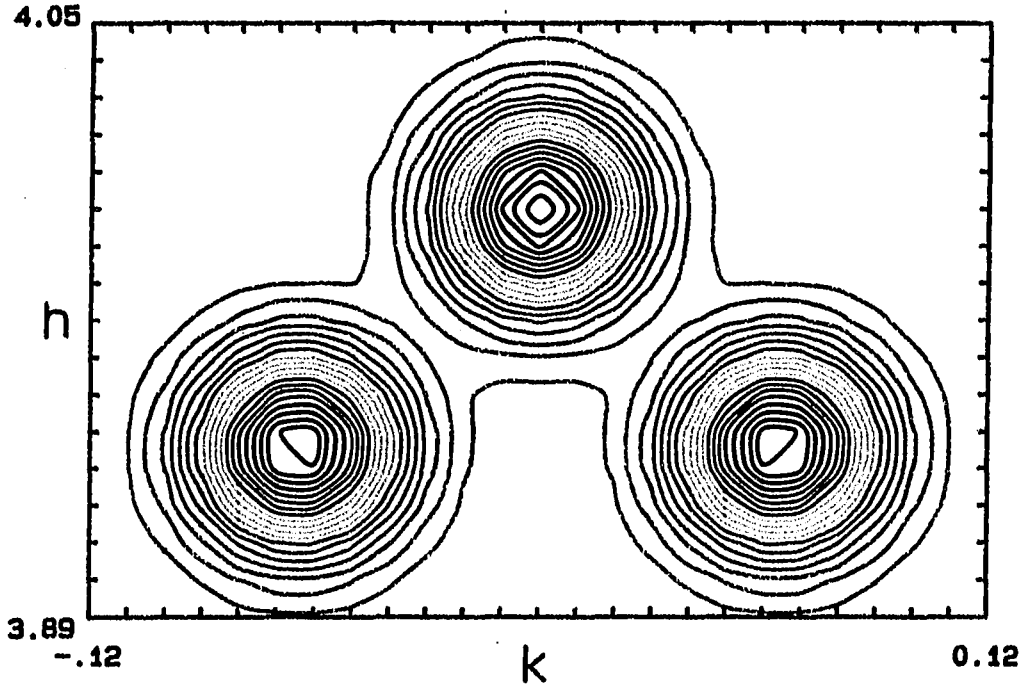


Figure B.10. Reciprocal space plot of the calculated (400) reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (2)

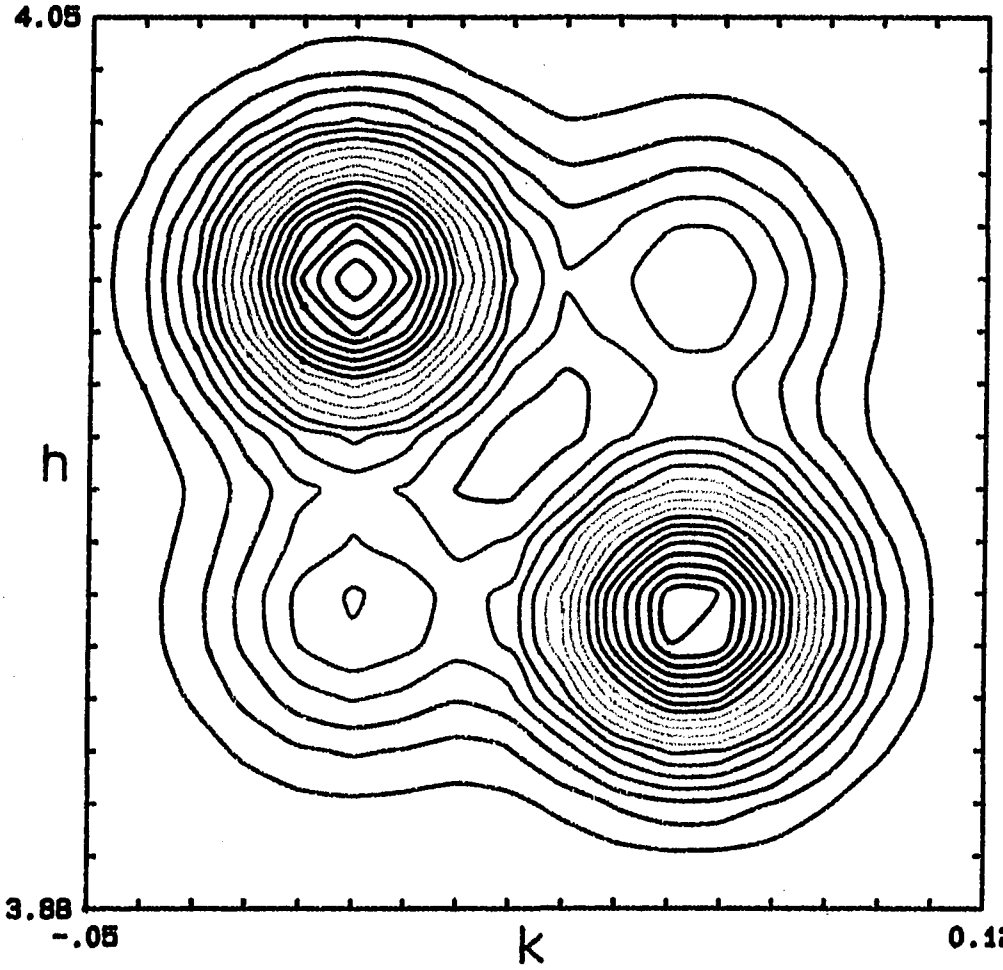


Figure B.11. Reciprocal space plot of the calculated (400) reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (3)

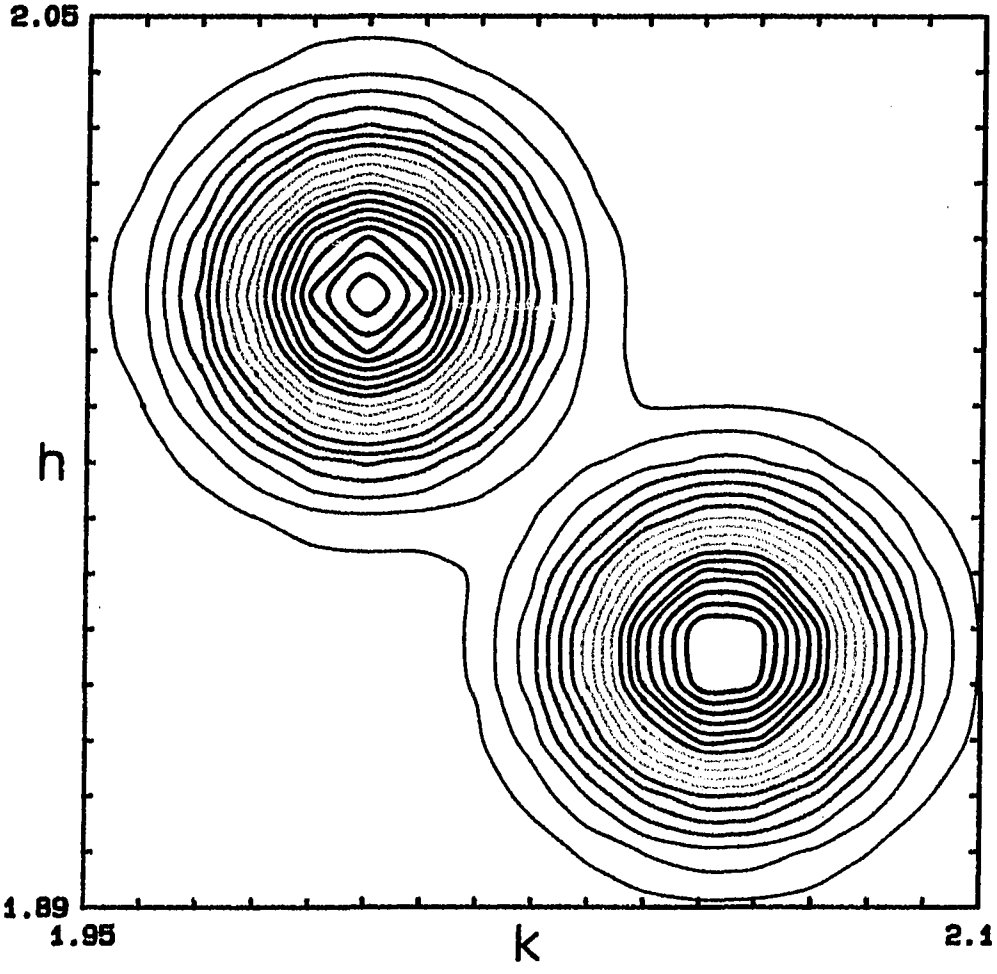


Figure B.12. Reciprocal space plot of the calculated (220) reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (1)

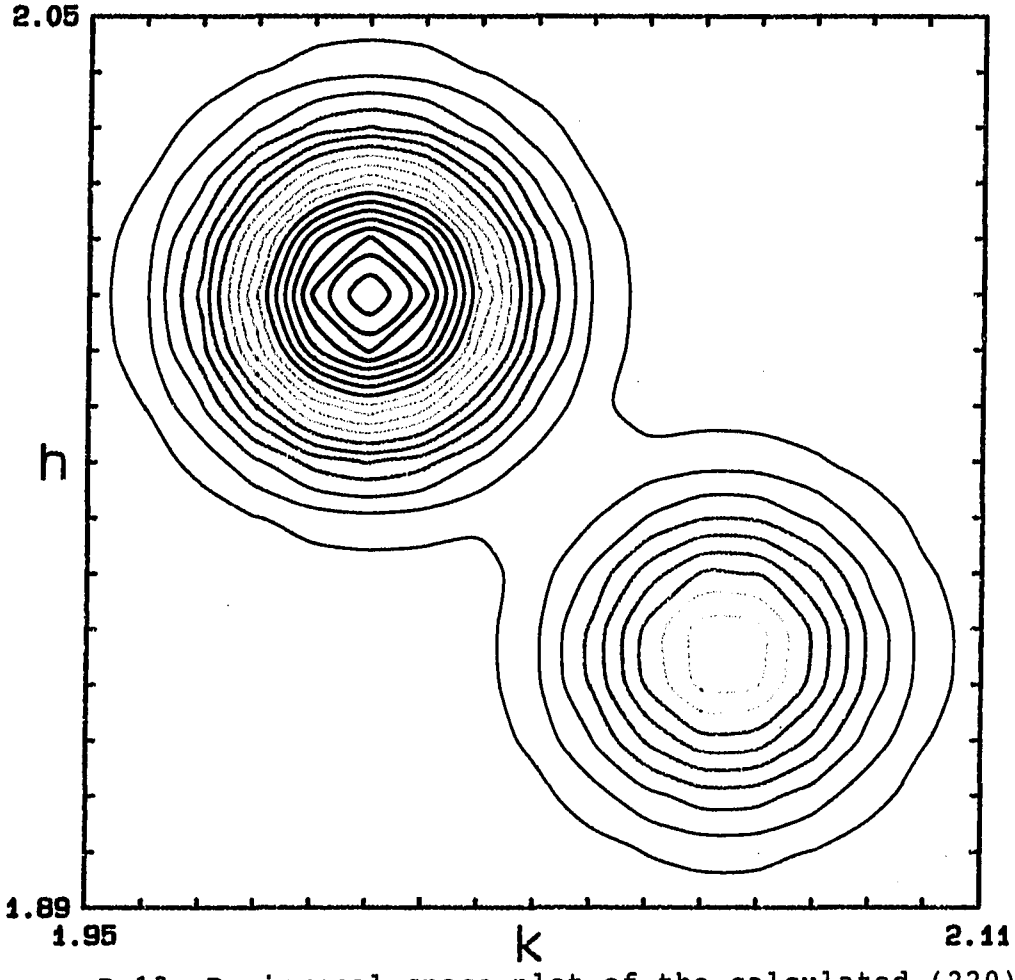


Figure B.13. Reciprocal space plot of the calculated (220) reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (2)

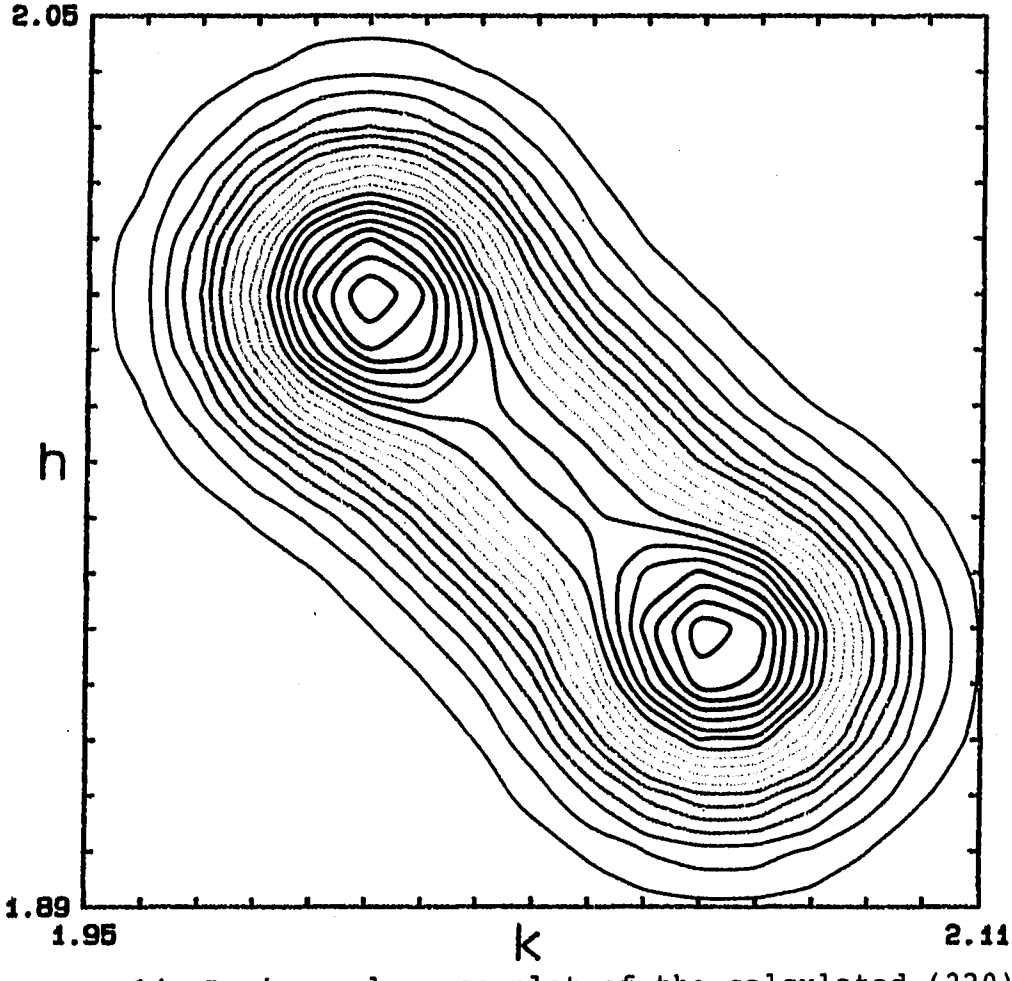


Figure B.14. Reciprocal space plot of the calculated (220) reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (3)

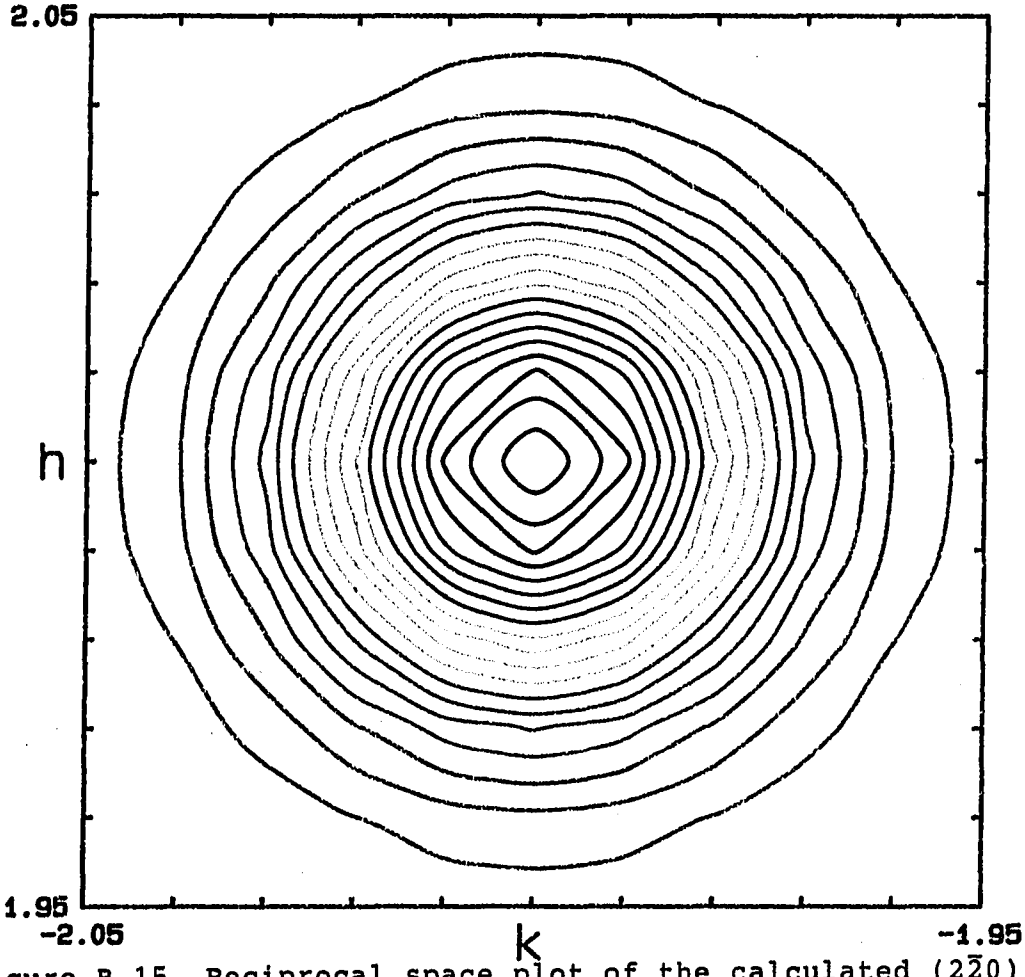


Figure B.15. Reciprocal space plot of the calculated $(2\bar{2}0)$ reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (1)

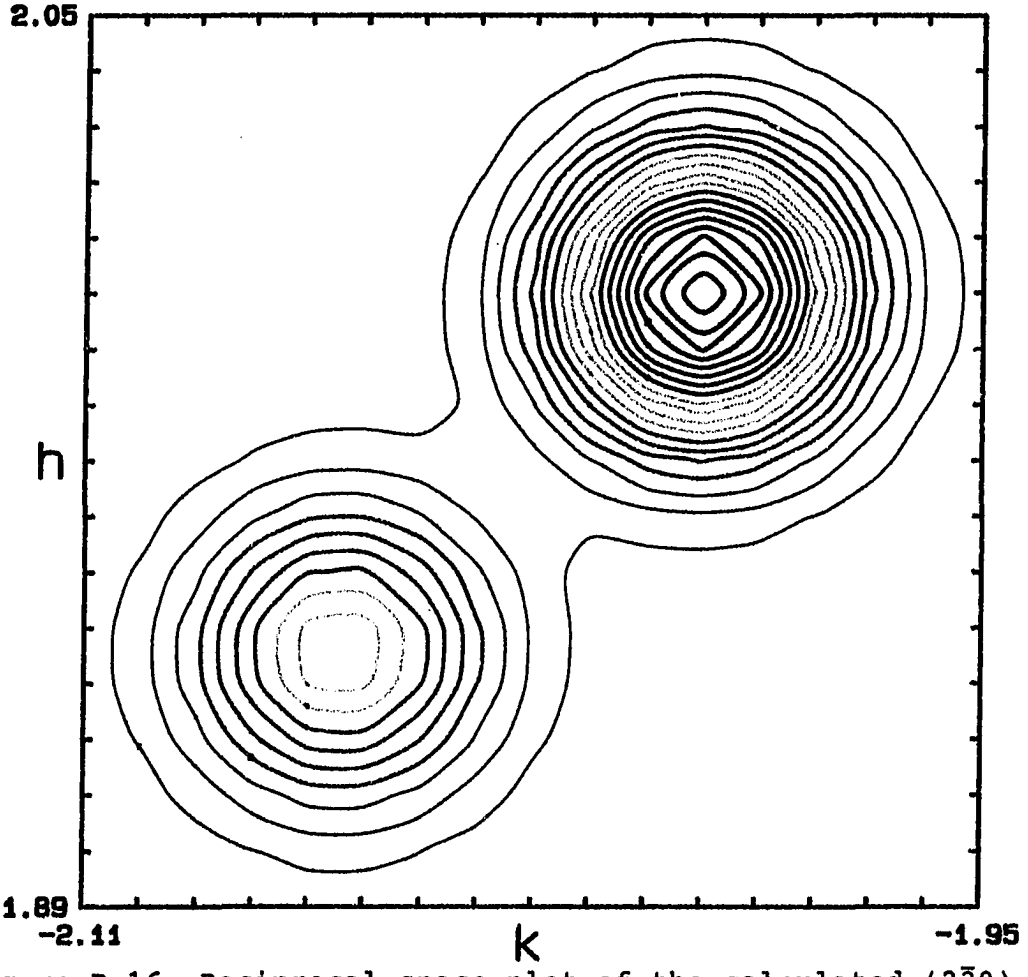


Figure B.16. Reciprocal space plot of the calculated $(2\bar{2}0)$ reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (2)

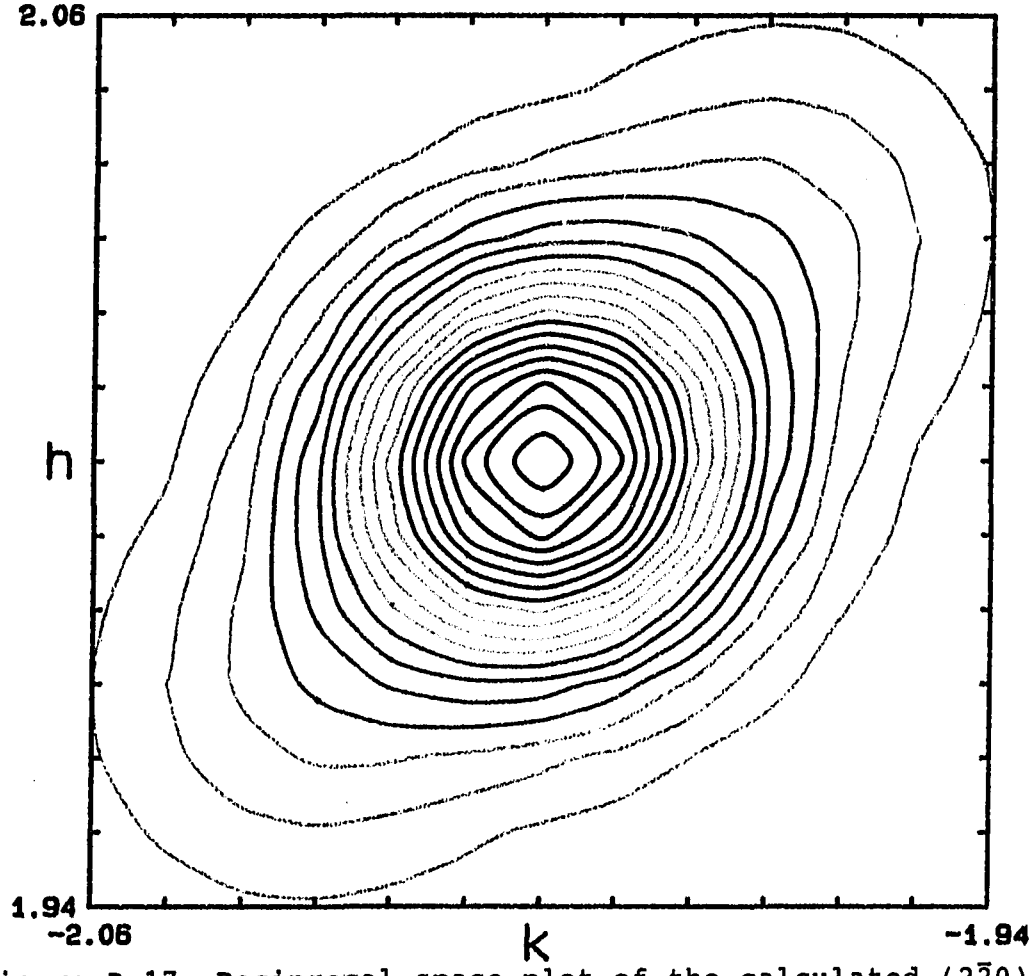


Figure B.17. Reciprocal space plot of the calculated $(2\bar{2}0)$ reflection for $YBa_2Cu_3O_{7-x}$ using twinning mode (3)

$h\bar{h}0$ and a satellite rotated by 90°. It is not unexpected that a rotated satellite is found, as this could easily be the result of the loss of the four-fold rotation axis. It is surprising that the model only contains one mirror. This may indicate that the correct space group is of lower orthorhombic symmetry, or that there is some type of long-range "memory" effect occurring so that once twinning begins to occur in one direction, that direction becomes the preferred twinning direction.

LITERATURE CITED

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

2. D. Harker and J. S. Kasper, Acta Crystallogr., 1, 70 (1948).

3. J. Karle and H. Hauptman, Acta Crystallogr., 3, 181 (1950).

4. D. Sayre, Acta Crystallogr., 5, 60 (1952).

5. M. Cochran, Acta Crystallogr., 5, 65 (1952).

6. W. H. Zachariasen, Acta Crystallogr., 5, 68 (1952).

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

8. W. Cochran, Acta Crystallogr., 8, 473 (1955).

9. J. Karle and H. Hauptman, Acta Crystallogr., 9, 365 (1956).

10. I. L. Karle and J. Karle, Acta Crystallogr., 16, 969 (1964).

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

12. D. White and M. M. Woolfson, Acta Crystallogr., A31, 53 (1975).

13. H. Hauptman, Acta Crystallogr., A31, 671 (1975).

14. H. Hauptman, Acta Crystallogr., A31, 680 (1975).

15. H. Schenk, Acta Crystallogr., A29, 480 (1973).

16. H. Schenk, Acta Crystallogr., A30, 477 (1974).

17. H. Schenk and J. G. H. de Jong, Acta Crystallogr., A29, 31 (1973).

18. Yao Jia-Xing, Acta Crystallogr., A37, 642 (1981).

19. R. Narayan and R. Nityananda, Curr. Sci., 50, 168 (1981).

20. R. Narayan and R. Nityananda, Acta Crystallogr., A38, 122 (1982).

21. G. Bricogne, Acta Crystallogr., A40, 410 (1984).

22. A. L. Patterson, Phys. Rev., 46, 372 (1934).

23. J. Richardson, Jr., Ph.D. dissertation, Iowa State University, Ames, Iowa, 1984.

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

25. A. Mighill, Ph.D. dissertation, Princeton University, Princeton, New Jersey, 1963.

26. J. Clastre and R. Gay, Compt. Rend., 230, 1876 (1950).

27. J. Garrido, Compt. Rend., 230, 1878 (1950).

28. M. Buerger, "Vector Space", Wiley: New York, 1959.

29. R. A. Jacobson, Trans. Am. Crystallogr. Assoc., 2 (1966).

30. D. W. Green, V. N. Ingram and M. F. Perutz, Proc. Roy Soc., A225, 287 (1954).

31. P. Tollin and M. G. Rossman, Acta Crystallogr., 21, 872 (1966).

32. C. E. Nordman, Trans. Am. Crystallogr. Assoc., 2, 29 (1966).

33. G. M. Sheldrick, in "Crystallographic Computing 3", edited by G. M. Sheldrick, C. Kruger and R. Goddard, p. 175, Oxford University Press, New York, 1985.

34. F. Pavelcík, J. Appl. Crystallogr., 19, 488 (1986).

35. P. Luger and J. Fuchs, Acta Crystallogr., A42, 380 (1966).

36. T. C. Terwilliger, S.-H. Kim and D. Eisenberg, Acta Crystallogr., A43, 1 (1987).

37. S.-L. Wang, Ph.D. dissertation, Iowa State University, Ames, Iowa, 1985.

38. Allen, P. W. and Sutton, L. E., Acta Crystallogr., 3, 46 (1950).

39. DeHaven, P. W. and Jacobson, R. A., Cryst. Struct. Commun., 5, 31 (1976).

40. Hackert, M. L., Jacobson, R. A., and Keiderling, T. A., Inorg. Chem., 10, 1075 (1971).

41. Hubbard, C. R. and Jacobson, R. A., Inorg. Chem., 11, 2247 (1972).

42. Lawton, S. L. and Jacobson, R. A., Inorg. Chem., 5, 743 (1966).

43. Lawton, S. L. and Jacobson, R. A., Inorg. Chem., 7, 2124 (1968).

44. Lawton, S. L. and Jacobson, R. A., Inorg. Chem., 10, 709 (1971).

45. Lawton, S. L. and Jacobson, R. A., and Frye, R. S., Inorg. Chem., 10, 701 (1971)

46. Porter, S. K. and Jacobson, R. A., J. Chem. Soc. (A), 1356 (1970).

47. Porter, S. K. and Jacobson, R. A., J. Chem. Soc. (A), 1359 (1970).

48. Porter, S. K. and Jacobson, R. A., Cryst. Struct. Commun., 1, 431 (1971).

49. Schroeder, D. R. and Jacobson, R. A., Inorg. Chem., 12, 515 (1972).

50. Wismer, R. K. and Jacobson, R. A., Inorg. Chem., 13, 1678 (1974).

51. Jacobson, R. A., J. Appl. Crystallogr., 9, 115 (1976).

52. Lapp, R. L. and Jacobson, R. A., "ALLS: A Generalized Crystallographic Least-Squares Program", USDOE Report IS-4708; Iowa State University, Ames IA, 1978.

53. Powell, P. R. and Jacobson, R. A., "FOUR: A Generalized Crystallographic Fourier Program", USDOE Report IS-4737; Iowa State University, Ames IA, 1980.

54. Zachariasen, W. H., J. Less Common Metals, 62, 1 (1978).

ACKNOWLEDGEMENTS

I wish to thank Dr. Robert Jacobson for his guidance and encouragement during the course of this work. I will take with me a very vivid image of him standing at the chalkboard, drawing pictures of unit cells and "solving" the phase problem. Many of the concepts in this approach came from those chalkboards.

I would like to thank our secretary, Brenda Smith, for all of her help and her friendship. I'm sure that she must have thought that I was being very antisocial my first year in the group.

I would also like to thank former and present group members Lance Miller, Raman, Dave Wintergrass, Ying-Zhong Su, and Cathy Day for their help in learning crystallography, their interest in my research and their willingness to be guinea pigs when I modified a program.

During my stay in Ames, I have had the privilege of meeting some very fine people. I think that I shall always remember the Sunday evening barbecues with Ross and Janet Nord, David and Diane Sanders, and Jeff Crain. I will certainly never forget the "turkey wine" in Boone and the infamous trip to Pittsburg for pizza. I am also grateful for the friendships that I have had with fellow graybeard Clark Carleson, Deb Spink, Dick Nagaki and Susan Zvacek.

All of them have made Ames a much more enjoyable place to live. Elen síla lúmenn omentilmo.

And finally, I would like to thank my parents and family for their support. They have stood behind me throughout the course of my graduate studies.

This work was performed under contract No. W-7405-Eng-82 with the U. S. Department of Energy and this thesis has been assigned a report number of IS-T-1408.