Materials Science Division of the US Department of Energy under contract No. DE-AC03-76SF00098.

References


IMPAS - A Simple Structure Determination Procedure Based on Intermolecular Patterson Vectors

BY PETER LUGER AND JOACHIM FUCHS
Fachbereich Chemie der Freien Universität Berlin, Takustrasse 6, D-1000 Berlin 33, Federal Republic of Germany

(Received 15 February 1985; accepted 22 April 1986)

Abstract
A computer method for interpreting Patterson functions of crystals belonging to space groups of higher symmetry than P1 is described. A table of all Patterson peaks in the entire cell is initially searched for vectors between symmetry-related atoms. Such 'generalized Harker vectors' are identified not only by known symmetry-fixed components, but also by statistical criteria based on a systematic search for pairs of vectors which relate additional atoms to the pair involved in the generalized Harker vector. Subsequent searches, using the same basic principle, are carried out to find additional atoms. Pairs of vectors, whose components add up to the previously established Harker vector, indicate newly found atom positions after appropriate transformations into crystal space. The correctness of these atom positions can be further tested by applying all symmetry operations of the space group in question and looking for a complete set of related vectors. In contrast to many other Patterson search methods no information about known molecular fragments is requested. Also described are four structures which were solved with the IMPAS procedure following unsuccessful attempts to solve them by direct methods.

Introduction
50 years ago, Patterson (1934) introduced the Fourier transform of the squared structure amplitudes as one of the most important crystallographic functions into the literature. For more than 30 years the deconvolution of the Patterson function was the major procedure for solving the crystallographic phase problem. Several techniques were developed [superposition method (Jacobson, 1966), image-seeking functions (Buerger, 1959), Faltsmolekülmethode (Huber & Hoppe, 1965), etc.] mainly aimed at searching for suitable intramolecular Patterson vectors.

In the last decade 'direct methods' have superseded Patterson methods as the major tool in solving crystal structures. Among the organic and organometallic structures published in 1982 the phase problem was solved with direct methods in 72% of the cases, and with Patterson methods in 28% (de Ranter, 1984). Nevertheless, a number of structural problems remain, where direct methods fail or are successful only after overcoming considerable difficulties. These problems occur in a few unsuitable space groups such as P1 or C2, in cases with high internal molecular symmetry (mainly observed in inorganic
compounds), in cases of superstructure lattices and also if moderate-intensity data sets were measured.

For these cases we are now observing a 'renaissance' of Patterson methods. A number of authors have developed new efficient Patterson search programs, e.g. the new vector search programs of Nordman (1980) and Strumpel (1983), a modified superposition method, Dietrich (1983), or the integrated Patterson and direct methods of Egert & Sheldrick (1985). Generally, these methods need a known input fragment to allow for a search for intramolecular vectors, at least in the initial stages. Methods which do not use molecular-fragment structures have been described by Jacobson & Beckman (1979) and by Simonov (1982).

As a further development of the latter type of application we have established a procedure which is based on a systematic search for intermolecular Patterson vectors related by one of the space-group symmetry elements.

Description of the method

For convenience we use two spaces. One is the crystal space $C$; vectors $r \in C$, $r = (x, y, z)$ generally represent position vectors of atoms relative to one origin of the unit cell. The second is the Patterson space $P$; vectors $u \in P$, $u = (u, v, w)$ represent Patterson vectors, having the property $u = r_1 - r_2$, where $r_1, r_2 \in C$. Generalizing the usual Harker vector notation, we regard each vector $k \in P$ as a Harker vector if $k$ is the difference vector of two symmetry-related atoms (in that sense $k = 2r$, where $r$ is a position vector in a centric structure, is taken as a Harker vector).

If the Patterson function is regarded as an $n$-fold image of an $n$-atom structure, it can be said that a structure is solved (a) if one connected image can be identified and (b) if the corresponding viewing atom is brought to its correct position relative to a valid unit-cell origin.

For problem (b) Harker vectors play an important role. If one Harker vector is known, the unit-cell origin can be fixed, at least in one or two dimensions. Generally, in cases where many atoms of equal weight are present, a Harker vector is difficult to recognize. A procedure which searches for appropriate Patterson vectors and supplies two-dimensional or one-dimensional statistics (for Harker planes and lines) has been written for identifying Harker vectors. Once this problem has been overcome, further searches are provided to solve problem (a).

To handle searches of the type described above an interactive computer program IMPAS (= intermolecular Patterson search) has been developed, which is based on systematic searches for pairs of vectors satisfying a given symmetry condition. Subsequent searches for additional symmetry elements may be executed.

In total, IMPAS offers three types of searches:

1. Routines allowing searches for Harker vectors on a statistical basis and thus helping implicitly in defining the origin of $C$.

2. If routines of type (1) have successfully been applied, or if the user himself has identified one Harker vector (e.g. in the case of a heavy-atom structure), searches are made for Patterson vector pairs $u, u' \in P$ which satisfy the symmetry condition associated with the given Harker vector. If a pair $u, u'$ has been found, it is transformed into the corresponding pair $r, r' \in C$ as far as is possible via the Harker vector.

3. If a supposed set of connected position vectors in $C$ has been determined, searches in $C$ may be executed to sort out ghosts.

The application of IMPAS to the structure determination of the heteropolytungstophosphate $K_{13}[KP_2W_{20}O_{72}] \cdot xH_2O$ (Fuchs & Palm, 1984) may illustrate the three types of search described above. Although this compound crystallizes in the space group $P2_1/c$, which is usually favorable for the application of direct methods, no solution with MULTAN (Main, Lessinger, Woolfson, Germain & Declercq, 1977) could be obtained. Two $[KP_2W_{20}O_{72}]$ units are in the cell, hence ten independent W atoms have to be located.

With respect to the twofold screw axis, a search for Harker vectors using one of the type-(1) IMPAS routines can be applied as follows. The Patterson vectors of two symmetry-related atoms $r$ and $r'$ relative to a reference atom $r_0$ are

$$u = r - r_0 \quad \text{and} \quad u' = r' - r_0.$$

In components,

$$u = (u, v, w) = (x - x_0, y - y_0, z - z_0)$$
$$u' = (u', v', w') = (-x - x_0, \frac{1}{2} + y - y_0, \frac{1}{2} - z - z_0).$$

It follows that

$$v' - v = \frac{1}{2}$$
$$u' + u = -2x_0$$
$$w' + w = \frac{1}{2} - 2z_0,$$

which are components of the Harker vector $k$ associated with $r_0$. To obtain a solution for $k$ the program searches for pairs of Patterson vectors which differ in $v$ by $\frac{1}{2}$. If such a pair is found, the vector sum is registered in a two-dimensional $u, w$ grid. After the search is completed, a $u, w$ grid point having a large number of such sums registered is accepted as representing the $u, w$ components of a Harker vector.

A search of this type was executed with the 457 highest Patterson vectors of the tungstophosphate structure. With a subdivision of the $uw$ plane into $50 \times 50$ grid points one accumulation of $u, w$ vectors was registered at $u = 4.5/50$, $w = 4.5/50$ with a total...
of 52 counts around this point (Fig. 1) and regarded as one of the tungsten Harker vectors.

With one Harker vector \( k \) of the structure known, a type-(2) search can be executed, which utilizes the fact that from a knowledge of \( k \) the displacement between spaces \( C \) and \( P \) can be calculated in those dimensions where \( k \) is operative.

The situation is illustrated in Fig. 2 in projection on the \( xz \) plane. Let \( s \) be the vector representing the displacement of the symmetry element from the unit-cell origin [in the case of the 2\( _1 \) axis in \( P2_1/c \), we have \( s = (0, 0, \frac{1}{2}) \)]. Let us view the Patterson space from the atom \( r_0 \) for which the Harker vector \( k \) is known. If \( r_{0y} \) and \( k_y \) are the projections of \( r_0 \) and \( k \) on to the \( xz \) plane, we have

\[
k_y = 2(s - r_{0y}).
\]

Let \( r, r' \) be the position vectors of a further atom pair of the structure, symmetry-related via the 2\( _1 \) axis, in \( C \):

\[
r = (x, y, z), \quad r' = (-x, \frac{1}{2} + y, \frac{1}{2} - z).
\]

The corresponding Patterson vectors relative to the Patterson origin at \( r_0 \) are \( u = r - r_0 \) and \( u' = r' - r_0 \) which have the property that their \( v \) components differ by \( \frac{1}{2} \) and that their projections on the \( xz \) plane satisfy

\[
u_y + u'_y = k_y, \quad k_y = (U_k, 0, W_k).
\]

Now IMPAS searches for pairs of Patterson vectors which differ in \( v \) by \( \frac{1}{2} \) and whose vector sum has \( u, W \) components equal to the Harker vector components \( U_k \) and \( W_k \). When such a pair is found, it is assumed to represent a pair of atoms relative to the Patterson origin atom \( r_0 \). From the knowledge of \( k \) [equation (1)] the displacement vector \( \frac{1}{2}k_y \) between \( r_{0y} \) and the symmetry element at \( s \) is known.

This makes it possible to assign the crystal-space \( x, z \) coordinates to the Patterson origin atom \( r_0 \), as well as the atom coordinates of \( r \) and \( r' \), with an indeterminate \( y \) coordinate:

\[
r_\Delta = (x, y + \Delta, z), \quad r'_\Delta = (-x, \frac{1}{2} + y + \Delta, \frac{1}{2} - z),
\]

where \( \Delta \) is unknown, i.e. the origin in the \( y \) direction is not yet fixed. Nevertheless, from a knowledge of \( k \) this search has allowed us to obtain a number of
atomic positions $r_{ai}$, $r'_{ai}$ ($i = 1, 2, \ldots$) based on the proper origin of $C$, except for their $y$ components.

The displacement $\Delta$ can be determined with a one-dimensional type-(1) search, where only the $r_{ai}$- and $r'_{ai}$-type vectors are used as input. Since the space group $P2_1/c$ also has the inversion center and the $c$-glide plane as symmetry elements, the search described above operates on vectors of the following types:

\begin{align*}
A: & \quad r = (x, y + \Delta, z) \\
B: & \quad r' = (-x, \frac{1}{2} + y + \Delta, \frac{1}{2} - z) \\
C: & \quad r'' = (x, \frac{1}{2} - y + \Delta, \frac{1}{2} + z) \\
D: & \quad r''' = (-x, -y + \Delta, -z).
\end{align*}

Among these vectors a search is made for pairs whose $x$ and $z$ sums equal zero (modulo 1). For these pairs the frequency of the sum of $y$ coordinates is plotted versus $y$. The maximum of this one-dimensional distribution is probably $2\Delta$, coming from the pairs $A/D$ and $B/C$:

\begin{align*}
A/D: & \quad r_{ai} + r''_{ai} = (0, 2\Delta, 0) \\
B/C: & \quad r'_{ai} + r''_{ai} = (0, 1 + 2\Delta, 1).
\end{align*}

Subtraction of $\Delta$ relates the $y$ coordinates to the proper crystal-space origin, so that after this search a set of position vectors $r_1, r_2, \ldots \in C$ is obtained, giving, in the case of a successful solution, all atom positions in the unit cell (not only in the asymmetric unit!).

However, in an actual case, some atom positions may be missing or vectors representing false atom positions (ghosts) may be in the list. While the first problem is not easy to overcome, an additional type-(3) search may solve the latter one.

Ghosts may be eliminated, if in the set of crystal-space vectors a search is made for a further space-group symmetry operation, in that for each vector $r \in C$ a check is made whether its symmetry-related vector $r'$ is in the set. For the tungstophosphate structure this was done by a search with respect to the $c$-glide plane. A final result with seven of the ten W atoms of the asymmetric unit was obtained; this was sufficient to phase the complete structure (see Fig. 3).

In addition to one- and two-dimensional searches for Harker vector statistics a 'zero'-dimensional search may be selected for the identification of Harker vectors in the space group $P1\bar{1}$. This search is based on the experience that among the highest Patterson vectors one Harker vector $k = (2x, 2y, 2z)$ should be present. In order to identify a vector of this type, the Patterson vectors are taken from the peak list in order of decreasing height, and assumed to be Harker vectors $k$ of an inversion center. Then a search for vector pairs $u, u'$ with

$$u + u' = k$$

is executed. The number of such vector pairs found, and their 'intensity sum',

$$\sum_{\text{all pairs}} I(u)I(u'),$$

where $I(u), I(u')$ are the peak heights of the Patterson vectors $u$ and $u'$, may serve as a criterion, whether the input vector is a Harker vector or not.

An example of the successful application of the zero-dimensional search option is the structure of $2,2'$-bipyridylium oxonium bis(diiodobromide) (Parlow & Hartl, 1985) space group $P1\bar{1}$. The second-highest Patterson vector was recognized as the Harker vector $(2x, 2y, 2z)$ of an I atom. With this vector as input to a further search the correct solution of the structure was obtained.

The program structure

**IMPAS** was written in Fortran IV and was developed as an interactive program on the CDC Cyber 835 computer. Since only the maxima of the Patterson syntheses are stored, the memory requirement is less than 32K 60-bit words, for up to 1000 Patterson vectors. The program can easily be adapted to Fortran 77 so that **IMPAS** can be run on computers of different type. An illustration of the program structure is given in the flowchart of Fig. 4.*

* The source program listing, a program description (in English) and a user's manual (in German) have been deposited with the British Library Lending Division as Supplementary Publication No. 42888 (48pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.
Except for the Patterson vectors, which must be provided on a separate input file (TAPE4) all input is entered interactively where the user is guided by the program to respond to self-explanatory prompts. Since it is the basic principle of IMPAS to search for symmetry-related vectors, it has to be ensured that not only an asymmetric set but Patterson vectors of the entire cell are supplied.

At the present stage the symmetry elements \(1, 2, 2_1, 4, 4\), certain types of sixfold axes, and all types of mirror and glide planes can be processed. Hence the program is applicable to all space groups except \(P1\), with the further exception that searches for certain types of three- or sixfold axes cannot be done.

In contrast to many other Patterson search programs no input of molecular fragments is required, so IMPAS may be used for structures of totally unknown geometry. Even if the space group is ambiguous, the success or failure of a search for a proposed symmetry element may help in establishing the presence or absence of the symmetry element in question.

**Examples**

A few further examples illustrate the application and usefulness of this program. Each of the examples given below refers to structures where direct methods failed or were successful only after overcoming some difficulties.

1. **Structure of the phosphorus sugar 1,2,3-tri-O-acetyl-4,5-dideoxy-4-C-[(R)-phenylphosphinyl]-\(\alpha\)-L-lyxofuranose** (Luger, Yamamoto & Inokawa, 1982).
   The space group, \(C2\), is known to present difficulties for 'direct methods', and a number of attempts with MULTAN failed. For IMPAS this structure was straightforward. The Harker vector of the phosphorus atom is \(k = P - P' = (u_p, 0, w_p)\), where \(P, P'\) are the P-atom positions relative to the twofold axis. This vector was easily located at \(k = (0.460, 0, 0.430)\). Then a type-(2) search for the twofold axis (\(INTER = 3\)) was executed. Fig. 5 shows the twofold axis at \(k/2\) in the Patterson space. Each atom vector pair \(r = (x, y, z)\) and \(r' = (-x, y, -z)\) produces a Patterson vector pair

\[
\begin{align*}
u &= k/2 + r \\
u' &= k/2 + r',
\end{align*}
\]

with

\[
\begin{align*}
u &= (u_p/2 + x, y, w_p/2 + z) \\
u' &= (u_p/2 - x, y, w_p/2 - z).
\end{align*}
\]

\(u\) and \(u'\) have equal \(v\) and their \(u, w\) components obey the condition \(u + u' = k\). IMPAS searches systematically among the set of Patterson vectors for vector pairs having this relation to a given \(k\).

If \(u_p/2\) and \(w_p/2\) are subtracted from a found vector pair, a transformation into crystal space is accom-
plished. Since the \( y \) origin can be chosen arbitrarily in this space group, no further search is necessary. The result shows the structure and its mirror image with respect to the \( xz \) plane. A discrimination between the two images must then be made either graphically or by inspection of bond lengths and angles.

The result for the phosphorus sugar is shown in Fig. 6. Among the 88 positions found in the cell eight were recognized as ghosts, so that 80, i.e., 20 unique positions were left, corresponding to 20 of the 25 atoms of the molecule. The computation time for this run was 2.8 s.

2. Structure of tetraethylammonium hexakis(diiodocuprate) \([(\text{Et}_4\text{N})_6\text{Cu}_6\text{I}_{11}]\text{I} \) (Mahdjour-Hassan-Abadi, Hartl & Fuchs, 1984), hexagonal space group \( P6_3/m \). No solution of the phase problem was obtained by direct methods. The structure determination was done with IMPAS as follows. First a two-dimensional Harker vector search was executed with respect to a twofold screw axis. One solution of this search was then input as a Harker vector to a search giving vector pairs related via the screw axis. Addition of a one-dimensional statistical search gave the relative position of the inversion center, i.e., the origin in the \( z \) direction.

Since the origin in the space group \( P6_3/m \) has to be fixed on the \( 6_3 \) axis, the procedure described above works correctly only if the previously chosen \( 2_1 \) axis coincides with the \( 6_3 \) axis. This cannot be ensured because twofold screw axes also exist in this space group, at \((0.5, 0, z),(0, 0.5, z)\) and \((0.5, 0.5, z)\), and are not identical to a \( 6_3 \) axis. Therefore an additional hexagonal search had to be carried out, which allowed a check whether the vectors found also satisfied a threefold-axis symmetry. When a wrong screw axis was input, this search gave no solution, but with a valid screw axis 32 vectors were found. This set was reduced to 24 vectors by further mirror plane and inversion center searches. These corresponded to the correct positions of all 24 I atoms in the unit cell (Fig. 7).

3. Structure of tetr phenylphosphonium acetonitrile triiodozincate(II) (Hartl, Bridgman & Salhab, 1986), \( \text{P(}C_6\text{H}_5\text{)}_4\text{ZnI}_3\cdot \text{CH}_3\text{CN}, \) space group \( C2cm \). No solution was obtained with direct methods. This structure was solved by IMPAS utilizing consecutively the relatively large number of symmetry elements in this space group.

Inspection of the distribution of Patterson maxima allowed a safe assignment of the space group. The space group \( C2cm \), rather than \( Cmc2 \), was indicated by the presence of strong Patterson vectors of the type \((0, 0, w)\) but not of the type \((u, 0, 0)\).

**Fig. 6.** Crystal structure (unit cell projected down \( y \)) of 1,2,3-tri-O-acetyl-4,5-dideoxy-4-C-[(\( R \)-phenylphosphinyl)-\( \alpha \)-L-lyxo-furanose as found by IMPAS. Circles designate atom positions not found by IMPAS.

**Fig. 7.** (a) Unit cell of \([(\text{Et}_4\text{N})_6\text{Cu}_6\text{I}_{11}]\text{I} \) projected on the \( ab \) plane; only the \( \text{Cu} \) and \( \text{I} \) atoms are shown (refinement results). (b) Structure found by IMPAS; all \( \text{I} \)-atom positions were correctly positioned. The asymmetric unit contains four \( \text{I} \) atoms, of which three are on special positions with multiplicities six, four and two.
Table 1. Structures solved with IMPAS

<table>
<thead>
<tr>
<th>Structure</th>
<th>Space group</th>
<th>MULTAN solution</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>The phosphorus sugar C₁₇H₂₁OTP</td>
<td>C₂</td>
<td>No</td>
<td>Luger, Yamamoto &amp; Inokawa (1982)</td>
</tr>
<tr>
<td>Dioxane C₅H₄O₂</td>
<td>P₂₁/n</td>
<td>Yes</td>
<td>Buschmann, Luger &amp; Müller (1983)</td>
</tr>
<tr>
<td>K₃P₂W₁₀O₄₂·xH₂O</td>
<td>P₂₁/c</td>
<td>No</td>
<td>Fuchs &amp; Palm (1984)</td>
</tr>
<tr>
<td>P(C₆H₅)Zn₃(C₂H₅CN)</td>
<td>C₂cm/P₂₁/n</td>
<td>No</td>
<td>Hartl, Brüdgam &amp; Salhab (1986)</td>
</tr>
<tr>
<td>[N(CH₃)₄]₄Na₂·OH₆·HPW₁₀O₄₂·xH₂O</td>
<td>C₂/c</td>
<td>No</td>
<td>Fuchs &amp; Palm (1986)</td>
</tr>
<tr>
<td>The caesium tellurate Cs₂Te₂.O₆·(OH)₆</td>
<td></td>
<td>Yes</td>
<td>Hartl &amp; Mahdjour-Hassan-Abadi (1984)</td>
</tr>
<tr>
<td>The iodocuprate [CH₃(CsHaN)₂]₂Cu₂I₆</td>
<td>P₂₁/c</td>
<td>Yes</td>
<td>Hartl &amp; Roth (1986)</td>
</tr>
<tr>
<td>Ethylthiobullvalene C₁₆H₁₆S</td>
<td>I₄</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

From a two-dimensional type-(1) search a suitable Harker vector with respect to the twofold axis in the a direction was chosen. This Harker vector was then input to a 2₁-axis search. Addition of c-glide (₁b) and mirror-plane (₁c) searches led to the positions of the two I, the Zn and the P atoms. No further peaks were found by IMPAS; however, in this acentric space group the mirror images of the heavy atoms also appeared. Bond-length considerations easily allowed a choice of a connected set.

Concluding remarks

The examples illustrated above and the list of structures already solved with the help of IMPAS (Table 1) demonstrate that a suitable Patterson vector search program can be a valuable alternative to direct-methods programs. We have had the best experiences with this program mainly for higher-symmetrical inorganic structures. For these structures IMPAS has the advantage – in comparison with some other vector search procedures – that no previous information on any molecular geometry is necessary. The present drawback may be that it is not applicable to all space groups (e.g. P₁ is totally excluded) and that the user must have some education in basic crystallography, especially in crystal symmetry.

It should be pointed out that IMPAS was not introduced as a substitute for direct-methods programs. For large organic light-atom structures these methods are still the first choice.

The authors thank Miss E. Müller for her assistance in coding and testing the program.

References