02-Methods for Structure Determination and Analysis,
Computing and Graphics

02.08 - Crystallographic Computing

OCM-02.08.01 HYPERTEXT FOR CRYSTALLOGRAPHY
by R.Diamond*, Medical Research Council, Laboratory for
Molecular Biology, Hills Road, Cambridge CB2 2QJ, England
and A Munnynck, Oxford University Press, Walton Street,
Oxford, OX2 6DP, UK

"Hypertext" designates a system of text and related material,
prepared in machine-readable form so that its presentation on
a computer screen includes within it "links" which enable the
reader to invoke supporting material, references figures etc. as
he wishes, using a mouse or similar device to control the
process. One such system, Molecular Structures in Biology,
will be described both from the user's point of view and in terms
of its internal construction, which is based on a "web" which is a
file defining many "nodes" which may themselves invoke
software to provide a pop-up reference, or another text, or
graphics software or a stored bit map image etc. The web itself
is traversed by software known as the "spider" which puts into
execution actions specified by nodes of the web. Such techniques,
coupled with the high storage capacity of CD-ROM (600 Megabytes),
provide an efficient means of publishing high-volume
numerically-intensive data, especially data with a high degree of
permanence, as in databases, where in crystallographic
application seems most likely to be fruitful MSB itself is both
a textbook and a reference work, containing over 500 coordinate
sets from the Protein Data Bank, plus twelve chapters of text
explaining hypertext techniques, plus many references, over 1000
illustrations (bit maps), and the capability for the user to create
an unlimited number of others.

OCM-02.08.02 IF WE CAN DO IT, YOU CAN DO IT:
WRITING A GUI INTERFACE FOR A CRYSTALLOGRAPHIC
PROGRAM*. By Paul N. Sweaton and Beverly R. Vincent,
Molecular Structure Corporation, 2220 Research Forest Drive
Woodlands, TX 77381 USA

Many crystallographic programs are still based on card-image
input. The user is required to use a text editor prepare an
instruction file for program execution. While this can be a
simple effective means of program control, it fails to take
advantage of modern software developments.
Examples will be given of graphical user interfaces (GUI) that
have been developed for crystallographic programs using the
MOTIF tool kit. Actual examples of source code will be shown
in order to demonstrate how it is to develop a modern graphical
interface.

OCM-02.08.03 PC: A CRYSTALLOGRAPHIC COMPUTING
TOOL FOR EXPERTS AND NOVICES. By V.K. Pecharsky,
Dept of Inorganic Chemistry, L'Iviv State University, L'Iviv, Ukraine.

The choice of type of computer(s) for daily use in the laboratory
for the tasks of diffraction-data processing, crystal-structure
solution, refinement and final representation faces every
crystallographer. For work on inorganic crystal structures and on
small organic molecular structures (up to 100-150 independent
atoms in the unit cell, with ca. 1000000 least-squares parameters),
personal computers (PCs) can be used very effectively. PCs have
many advantages, viz. they are open and friendly systems, they can
easily be used by students, by regular scientific staff and by
experts, and their capital cost is low. With proper software and
suitable user interface, success is assured for everyone. PCs are
excellent tools for crystallographic computing today and probably
in the future.

We will illustrate our experience with PCs by describing the
software package CSD (Crystal Structure Determination) for
crystallographic computing on single-crystal and powder, X-ray and
neutron diffraction data developed over the past few years. It will
be shown that the average time spent by a crystallographer in the
solution and refinement of a crystal structure is significantly
less than the time necessary to collect the diffraction data. It does not
usually exceed 1-4 hours, depending on the number of free
least-squares parameters. Over two thirds of this time is spent on the
final stages of refinement.

The author wishes to thank L. Akemi, Yu. Grin and P.
Zavilj in collaboration with whom the crystallographic system
CSD has been successfully developed and used, and without whose
kind help this work would not have been possible.

OCM-02.06.04 THE PROFILE FITTING IN SINGLE-CRYSTAL X-RAY
DIFFRACTION. By Dariusz Galbecka, Institute of Low
Temperature and Structure Research Polish Academy of
Sciences, ul. Ukielna 2, 50-950 Wroclaw, Poland.
The whole-profile fitting that plays a fundamental role in
crystal-structure determination from powder samples (the
Powell method) is rather occasionally used for processing the single-crystal intensities
traditional background-peak-background procedure is considered to be sufficient in the latter case. The
purpose of the paper is to work out a suitable method for approximation of the single-crystal diffraction
profiles and to test the effect of the careful data
processing on the precision and reliability of the
crystal-structure determination. The subjects of
detailed considerations are such problems as criteria of the goodness of fit, the choice of the basic
approximating function (shape function), the proper
number of independent adjustable parameters, the
dependence of parameters of the profiles on the Bragg
gle and direction junctions of the diffraction vector, and – recently discussed by Schwarzenbach &
Flack (Acta Cryst. (1991), A47, 134-137) and Lenstra,
Geise & van Beusekom (Acta Cryst. (1991), A47,
597-598) – treatment of the background and 'negative'
reflections. The basis for the considerations is the
papers mentioned above and two papers by the authors
116-121). Results of the present work are currently
being incorporated into a computer procedure which
approximates the diffraction profile and calculates
the integrated net intensities. The newest results of
the crystal-structure determination that includes the
profile fitting, as compared with those obtained using
the common approach, will be presented.

02-Methods for Structure Determination and Analysis, Computing and Graphics

Many scientific fields deal with the calculation or observation of 3-dimensional data, for example electron density, Van der Waals surface and electronic potential calculations. Conventionally, this data is stored in grid form where a large number of data points are needed to fully describe it, which leads to problems with both storage and analysis.

An alternative approach is to approximate the grid by fitting a set of orthogonal functions to it. Once this has been carried out one only the function coefficients need to be stored - i.e. a few hundred numbers instead of several thousand. The 3-d information can then be regained when required by reversing the calculation.

The functions we have chosen are very similar to those used to describe atomic orbitals, they differ only in that they have been scaled to represent a whole molecule and not just the volume associated with a single atom. This technique has been applied previously for electron density display and fitting in the field of Protein Crystallography but has not been used extensively for smaller molecules.

We are particularly interested in the following applications of this technique:

Crystallography:

The representation of molecular shape and charge for use in crystal packing studies.

The method can be used to represent voids within crystals, something that is difficult to achieve by other means.

Drug Design:

Molecular shapes are important for explaining and/or predicting the interaction of a drug with a putative binding site.

QSAR studies on the coefficients can be easily carried out. The number of parameters involved is relatively small and the nature of the functions allow the use of fast rotation algorithms to aid molecular comparisons.

The applications of the technique are limited only by the types of information that such functions can be expected to represent. If required another set of functions can be used for other cases as the choice is only limited by the need for orthogonality. Also, as the functions are able to represent a complex 3-D grid, two independent sets of data for the same molecule can be dealt with simultaneously. One is entered as though it was the real part of the grid and the other as the imaginary part; the two sets of numbers are kept completely separate during the calculations but may be displayed and analyzed together if required.

As part of our project in Molecular Scene Analysis (Fortier, S., Cashdole, I., Ghogaw, J. I., Conklin, D., Waddley, C., Lehar, L. & Allen, F. H. (1997); ORCritel, The Oak Ridge Critical Point Network Program, Chemistry Division, Oak Ridge National Laboratory, USA) we have been investigating methods to assist in the spatial and visual analysis of electron density maps at varying resolution. In particular, we have assessed the usefulness of the topological approach for the segmentation of medium (2 A) resolution maps of proteins and their interpretation in terms of structural motifs. We have followed the approach implemented by Johnson (Johnson, C. K. (1977); ORCritel, which provides a global representation of the electron density distribution through the location, identification and linkage of its critical points (points where the density gradient vanishes). In the first part of our study, the topological approach was applied to (calculated) maps of three proteins of small to medium size so as to develop a methodology - rules, heuristics or templates - that could then be used for analyzing maps of medium resolution. The methodology was then applied to both calculated and experimental maps of penicillopepsin at 3 A resolution. The study shows that the networks of critical points provide a useful segmentation of the maps, tracing the protein main chains and capturing their conformation. In addition, these networks can be parsed in terms of secondary structure motifs, through a geometrical analysis of the critical points. The procedure adopted for secondary structure recognition was phrased in terms of geometry-based rules. It provides a basis for an automated implementation through the use of artificial intelligence techniques.

PS-02.08.07 THE USE OF A MODIFIED DOUBLE PATTERSON FUNCTION IN DIRECT PHASE DETERMINATION.

By Yuan-Feng Wang*, Ross Peachar and Henk Schenk, Laboratory for Crystallography, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands.

Nowadays in X-ray crystallography most single crystal structures are solved by direct methods, which are based on probability theory. From this theory the valence of structure invariant and unitary are estimated, one of them being the triplet phase sum $\psi$. If interatomic triangles ($r_1$, $r_2$, $r_3$) can be obtained, $\langle \cos(\psi) \rangle$ and $\langle \sin(\psi) \rangle$ can be expressed in more accurate formulas. (Koenen, Thesis, Univ. of Amsterdam, 1992, 67-75).

It is commonly known that the Patterson function consists of interatomic vectors. Two arbitrary Patterson vectors, however, do not necessarily share a common root, so the construction of triangles ($r_1$, $r_2$, $r_3$) from Patterson vectors is a complicated task. In this respect the double Patterson function (Vaughan, Acta Cryst., 1958, 11, 111-115):

$$P_{(u,v)}(1) = \sum \langle o(r) \rangle \cos(u(v) \cdot r)$$

is more interesting for two related reasons:

1) $P_{(u,v)}(1)$ is the Fourier transform of the triplet phase sum.

2) non-zero $P_{(u,v)}(1)$ are possible only if $u$ and $v$ form an interatomic triangle.

A modification of the double Patterson function is proposed. Calculations for both model and real structures are presented which show that interatomic triangles can be constructed more safely with the double Patterson function than with the normal Patterson function.

PS-02.08.08 A PACKAGE OF FAST FOURIER TRANSFORM ROUTINES FOR MACROMOLECULAR CRYSTALLOGRAPHY.

E. Prince1, M. An2, C. Le3 and R. Tonkyn1.

1. Center for Molecular Biology, National Institute of Standards and Technology, Gaithersburg, MD 20899, U. S. A.

2. Aware, Inc., 1 Memorial Drive, 4th Floor, Cambridge, MA 02142, U. S. A.

3. Department of Computer and Information Sciences, Towson State University, Towson, MD 21204, U. S. A.