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

The extensive use of Fourier transform methods and the large sizes of unit cell data sets in macromolecular crystallography make it highly desirable to use efficient methods of computation. Fast Fourier transform (FFT) methods are very useful for solution of these problems. With a few exceptions, the libraries of FFT subroutines that have been available in the past have been restricted to sample sizes (numbers of grid points per period) that were products of small, prime numbers, which results in rather large gaps between numbers that may be large for large unit cells and moderately high resolution. Efficient FFT routines for different types of numbers – prime numbers, powers of prime numbers, or products of distinct prime numbers and their powers – follow many different procedures, so it is not possible to write general routines that can handle all cases. Also, when many transforms are required for sample sequences of the same size, a routine that is optimized for that particular size is desirable.

The crystallographic problem allows exploitation of certain special conditions to make computation more efficient. Electron density is a real function, so its Fourier transform has Hermitian symmetry. Three-dimensional transforms may be performed as sequences of one-dimensional transforms, and the data in the intermediate steps may have a symmetric structure. For example, a 2D space axis gives a data structure in which values separated by half of a lattice translation are complex conjugates, and centered lattices produce rows that are alternately symmetric and antisymmetric. Techniques for exploiting space group symmetry for space groups that have no rotation axes of order higher than two have been described by Tan Eyck (Acta Cryst., 1971, A29, 183-181), but space group symmetry with three-, four-, and six-fold axes may also be used to reduce the computational load.

We have developed a package of FORTRAN routines for one-dimensional FFTs for all odd number sample sizes from 11 to 99 and all even number sample sizes from 20 to 280 that have no prime factors larger than 11. The package also includes transforms from real sequences to Hermitian sequences, and from Hermitian sequences to real sequences, as well as between sequences that have translational conjugate symmetry and sequences that have alternating Hermitian symmetry and Hermitian antisymmetry. These routines are utilized by transform routines for a set of space groups that contain more than 80% of reported protein crystals.

1,2-dihydroxybenzene (catechol) and 1,2-dimethoxybenzene molecules were exploited to elucidate the procedure. Structural information was retrieved from the Cambridge Structural Database and two-dimensional potential energy maps were calculated using the molecular mechanics method.

**PS-02.08.10 SHELX-92 A NEW LEAST-SQUARES REFINEMENT PROGRAM FOR USE WITH SINGLE CRYSTAL DIFFRACTION DATA**

George M. Sheldrick, Institut fur Anorganische Chemie, der Universit鋞, G鰚tingen, Germany, and Ward T. Robinson, Department of Chemistry, University of Canterbury, Christchurch, New Zealand.

SHELX-92 is a FORTRAN-77 program for the refinement of crystal structures from X-ray or neutron diffraction data, and is primarily designed for single crystal data from small structures (1000 atom) at atomic resolution. It is intended to be easy to install and use on a wide variety of computers, and replaces the structure-refining part of SHELX-76.

SHELX-92 is general and efficient for all space groups in all settings and there are no arbitrary limits to the size of problems which can be handled, except for the total memory available to the program. All instructions are in machine-independent free format, with extensive use of default settings to minimize the amount of input required from the user. Instructions and data are taken from two standard (ASCII) text files, so that input files can easily be transferred between different computers. SHELX-92 is a PUBLIC DOMAIN program; it is provided in source form, also as a precompiled version which has been optimized for PCs. There are no restrictions on its use or distribution anywhere in the world for non-commercial purposes.

The program produces all the information required for efficient development of a complete structural model from a small, correctly oriented data set. The program can be run on any computer that has a standard compiler and Fortran-77 libraries, and the program is distributed with a comprehensive manual. The program is designed to be easy to use, and the instructions are concise but thorough. The program is intended to be used by crystallographers who are familiar with the basic principles of crystallography.

A hypothesis concerning possible minimum energy conformations and conformational interconversions can be supported by statistical analysis of similar molecular fragment potentials, potential energy, calculation, topology and other methods (Klauk D., and Dimitri D., Acc.Chem.Res., 1985, 16, 153-161; Munge P.G., Potential Energy Hypersurfaces, Amsterdam: Elsevier, 1987). Molecular geometry of a fragment under study gives additional limiting conditions (Frei H., Papier A., and Gmehn H.U., Top.Complex., 1979, 81, 1-97). It is believed that properly chosen model of static and/or dynamic behaviour of the molecular fragment leads to a good mutual agreement among results reached by the methods mentioned above. Potential energy minima correspond to chemical entities which represent the most frequently occurring conformations (molecular templates). On the other hand, sparsely scattered conformations bridging stable conformations span the probable reaction pathways during the conformational interconversions.

**PS-02.08.09 STATISTICAL ANALYSIS AND REACTION PATHWAYS OF MOLECULES**

by K. Humbert, P. Soballe, and W. Hummel, Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Hradeckého nám. 2, 182 06 Praha 6, Czech Republic.

A hypothesis concerning possible minimum energy conformations and conformational interconversions can be supported by statistical analysis of similar molecular fragment potentials, potential energy, calculation, topology and other methods (Klauk D. and Dimitri D., Acc.Chem.Res., 1985, 16, 153-161; Munge P.G., Potential Energy Hypersurfaces, Amsterdam: Elsevier, 1987). Molecular geometry of a fragment under study gives additional limiting conditions (Frei H., Papier A., and Gmehn H.U., Top.Complex., 1979, 81, 1-97). It is believed that properly chosen model of static and/or dynamic behaviour of the molecular fragment leads to a good mutual agreement among results reached by the methods mentioned above. Potential energy minima correspond to chemical entities which represent the most frequently occurring conformations (molecular templates). On the other hand, sparsely scattered conformations bridging stable conformations span the probable reaction pathways during the conformational interconversions.
Traditionally, crystallographic programs are run from a few standard data files for routine tasks, edited with the values for the current compound and run as batch jobs. For non-routine tasks the user is left searching the manual for elusive information.

An alternative modern strategy provides the user with a menu of procedures. This can be a useful facility for routine tasks, but offers the user little help with obscure or unusual problems.

In CRYSTALS we provide both of these input modes together with an interactive interface through which the program enters into a dialogue with the user. This dialogue is driven by external 'scripts', and so can be tailored to suit novices (as a teaching aid), to simplify routine tasks, or as an aid for attacking unusual problems.

Since the user can write his own scripts, he can encode in them any special knowledge or experience he has. Since SCRIPTS can read data files, call other SCRIPTS, and can even write new SCRIPTS themselves, the concept has the potential to be developed into an 'Expert System'.

At the meeting we will be showing, as an example, a script written to assist in the anisotropic refinement of large organo-metallic compounds containing very many organo ligands and counter ions. Because each script is just a plain language ASCII data file, users can easily modify them to handle local problems without risk of corrupting the CRYSTALS program itself.

The integration of a wide range of refinement, Fourier, geometry, graphical and analytical functions into a single package which can be used both interactively and in batch, which runs with identical functionality on anything between a PDP-8 and mainframes like the CONVEX makes CRYSTALS suitable for almost any environment or problem.

For users able to work interactively, there are integrated graphics which can be used both for the production of diagrams for publication, and as a powerful tool when dealing with large structures. The integration of Fourier map contouring with principal axis and TLS analysis of molecules is of great value for understanding disorder and thermal motion.

The ability to switch between the various user modes will provide the flexibility needed for efficient structure analysis.

Enquiries about obtaining CRYSTALS should be addressed to David Watkin at the address above.

PS-02.06.12 THE ESTIMATION OF PHASE INVARIANTS FOR LARGE DIRECT METHODS STRUCTURE DETERMINATIONS. By D. A. Langs* and D. Y. Guo, Department of Molecular Biophysics, Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, N.Y., U.S.A.

Simple tangent formula phase determination methods usually reach their limit of applicability when the complexity of a light atom structure exceeds 100 to 150 atoms. A new frequency-based triples phase invariant estimation procedure (Langs, Acta Cryst., A40, 1993, in press), employing traditional quadrupole relationships, has been shown to be successfully applicable to tangent formula structure determinations of greater than 300 atoms complexity. Efforts to extend the reliability of these frequency-based cosine invariant estimates by exploring higher-order relationships among phase invariants is in progress. Our experience using higher than 4th order Karle-Hauptman determinantal constructions has not been overly encouraging. We have however discovered one particularly interesting high order relationship involving 13 E-magnitudes that is most unusual in that it often embeds an unexpectedly high percentage of aberrant triads which have large A-values. We are currently exploring whether the frequency criterion or derived probability distribution can take advantage of these E-magnitude constructs to discern the identity of the embedded negative triples cosine invariants. If this can be achieved we strongly feel that this information will permit the solution of many more structures of greater than 900 atoms complexity by direct phasing methods.

Research support from NIH grant GM-46733 is gratefully acknowledged.

PS-02.08.13 COMPARISON OF F AND F^2 STANDARD REFINEMENTS OF SMALL MOLECULE CRYSTAL STRUCTURES. By S. James, M. Nardelli*, Istituto di Chimica Generale, Università di Parma, Centro di Studio CNR per la Strutturistica Difrattometrica, Viale delle Scienze 78, I-43100 Parma, Italy.

It has been demonstrated that, in least-squares refinement of crystal structures, exclusion of weak (or 'negative') intensities leads to a bias in the remaining experimental reflections towards too high F^2 values and thus to systematic errors in the refined parameters (F.L. Hirshfeld and D. Zahn, Acta Cryst. (1973), A29, 310-313; L. Arnborg, S. Hennemuller and S. Weisman, Acta Cryst. (C1979), A35, 497-499; P. Seiler, W. B. Schweizer and J. D. Dunitt, Acta Cryst. (1984), B40, 319-327). It seems that, omitting F^2 smaller than some predetermined size have some effect in the refinement that should be more important for scaling factors and temperature factors than for positional parameters. Moreover, carrying out two refinements with different weighting schemes might reveal the presence of significant but unobservable defects in the model or systematic errors in the measurements (A.J.C. Wilson, Acta Cryst. (1973), B29, 1488-1490).

Taking the opportunity of checking the newborn SHEXL-92 system for refinement of crystal structures (gamma-test stage) that is based on the use of F^2's, fifteen crystal structures of small organic and metal-organic molecules were refined, using data of standard quality, both on F (by using SHEXL-76) and F^2. The results were compared using the half-normal probability plot technique (A.J.C. Abrahams and T. Keve, Acta Cryst. (1971), A27, 157-165) applied to all the intrinsic distances less than 4.65 Å (W.H. De Camp, Acta Cryst. (1973), A29, 148-150) and to the Ueq values. The main results can be summarized as follows:

- The positional parameters obtained through these refinements are in general not significantly different, as only a limited number of long contacts have ∆ρ > 3, which means that significant differences may be present only in some values of bonds and main torsion angles.
- No indication of systematic errors or some relevance is observed for positional parameter, in the major number of cases. For these parameters in six cases the standard deviations are overestimated by factors ranging from 1.06 to 1.7, while in all the other cases the e.s.d.'s are underestimated by factors ranging from 1.07 to 2.25.