DARESBURY LABORATORY

INFORMATION QUARTERLY for PROTEIN CRYSTALLOGRAPHY

An Informal Newsletter associated with Collaborative Computational Project No. 4 on Protein Crystallography

Number 4 August 1980

Contents

Editor:	Pella Machin	Science Research Council, Daresbury Laboratory, Daresbury, Warrington WA4 4AD, England
Deputy Editor for Imperial College:	Dr. Alan Wonacott	Imperial College of Science & Technology, The Blackett Laboratory, Prince Consort Road, London SW7 2BZ
Deputy Editor for Birkbeck College:	Dr. David Moss	Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX

EDITORIAL

PELLA MACHIN

This issue of the newsletter contains reports about work going on in some of the university groups and at Daresbury.

Photographic film scanning is of interest to most protein crystallography groups. Contributions describing relevant work in microdensitometry are therefore requested for future editions of this newsletter. Mike Elder has helped begin this by describing the SRC microdensitometer service.

Plans are going ahead for the study weekend on the "Refinement of Protein Structures" to be held 14-16 November, 1980. Information and application forms can be obtained from Mrs. S. Lowndes, SRC, Daresbury, Warrington. If anyone knows of recent developments which should be discussed at the meeting please let me know.

REPORT ON COMPUTING BY THE SHEFFIELD PROTEIN GROUP

PHIL BOURNE

Currently Rob Stansfield and myself, under the direction of Prof. Pauline M. Harrison, are improving the structure of the iron storage protein horse spleen apoferritin. The small size of our group has meant that collaboration on computer software has been, and will continue to be, a great asset. Although not writing any major programs to contribute to the Daresbury system,I have updated some existing programs and fixed bugs in others. (Apoferritin crystallizes in space group F432, but the most readily available subgroup for the space group general programs is P222 which appears not to have been extensively tested.).

CLISTS and documentation (typed sheets) which reflect user experience are available for all programs used in Sheffield providing easier access to future users of the system at Sheffield and possibly some other CCP-4 users.

Details of modifications to programs at Daresbury which may be of use to users are set out below:-

FCALC - space group independent structure factor calculation. I have provided a CLIST and improved write-up. Furthermore I have added a job step which calculates a conventional residual for regions of $1/d^2$ and also applies linear rescale factors.

NORMAL - absolute scale and temperature factor calculation. A program modified from the MULTAN suite.
Write-up and CLIST available.

FFT - space group independent Fourier and difference Fourier calculation. The P222 version of this program does not work correctly.
I experimented with applying the sub-group symmetry outside the program and using an IBM sort to order the data (a necessity for the first transform). This bought no significant saving in cpu time for space group F432.
The program should be modified to use a direct access scratch file which would significantly reduce the I/O between the first and second transforms.
Updated write--up available.

PLUTO - plotting of molecules and contour maps. The quality of the contour maps plotted on the CALCOMP have been satisfactory, although the need for more than one pen colour should be stressed. The rate determining step has been the time for plots to travel from Daresbury to Sheffield. USG have been most helpful in organising postage.

When contouring maps it is advisable to submit individual plots containing stacks of 10 sections or less. This minimises loss during any plotter or machine failure and does not tie up the plotter with one job for long periods (my last map took over 20 hours on the Calcomp).

MAIN - generation of a poly-alanine structure given alpha carbons and carbonyl oxygens and includes a molecular geometry calculation. CLIST and documentation available.

DISTAN - polypeptide geometry calculation. CLIST and updated documentation available.

MODFT - polypeptide co-ordinate refinement. CLIST and updated documentation available.

TORSION- torsion angle calculation including a standard deviation calculation for better resolved structures. CLIST and write-up available.

DIRDIF - application of direct methods to partially known structures (van den Hark et al, (1976), Acta Cryst. A32, 816). CLIST and documentation available.

Future developments will include the use of phase combination programs and the implementation of the Konnert-Hendrickson restrained least-squares refinement program for F432.

A COMPARISON OF STRUCTURE FACTORS CALCULATED USING FAST FOURIER TRANSFORM AND CLASSICAL METHODS

JOHN W. CAMPBELL (DARESBURY LABORATORY)

1. Introduction

At the March, 1980, meeting of Working Group II, Phil Evans reported differences in structure factors calculated by Eleanor Dodson's version of the Isaacs Agarwal fast fourier transform refinement program and those calculated by another program. It was proposed that a study should be made at Daresbury to compare the structure factors calculated by Eleanor Dodson's program with those calculated using a classical structure factor calculation. This note presents the results of such a study and should be of particular interest to those using fast fourier transform methods for structure factor calculation.

2. Calculating Structure Factors Using the Fast Fourier Transform (FFT)

The method for calculating structure factors using FFT techniques has been described and discussed by L.F. Ten Eyck[1]. In outline, the method involves the calculation of a fast fourier transform of a modelled electron density map.

In the Isaacs Agarwal method[2], the electron density is built up using the contributions from each atom within or near (within the atom radius) to the part of the electron density map being calculated. The atomic form factors are approximated by gaussian functions with normally 2 or 5 term gaussians being used depending on the accuracy required. The radii of the atoms used in the calculation are chosen to balance the requirements of speed (small radius) and accuracy (large radius). A three dimensional FFT of the modelled electron density map is used to calculate the structure factors. If the sampling grid used is too coarse, significant errors may occur in the calculation of the structure factors. These errors may be reduced, either by using a finer grid for the calculation, or by adding an artificial extra temperature factor to all the atoms when modelling the density. This factor is also taken into account when the calculated structure factors are compared with the observed structure factors.

3. The Test

The program, chosen for the classical structure factor calculation, was 'FCALC' from Birkbeck College, London, as made available by Phil Bourne from Sheffield. The calculations were carried out in the space group P1 for both programs using a test structure consisting of the C, N and O atoms of the Co-enzyme B12 (actual space group $P2_12_12_1$). The program FCALC was used with an overall temperature factor of 5.0 and a B value of 5.0 was assigned to each atom for the Isaacs Agarwal calculation. Four parameters of the Isaacs Agarwal calculation were investigated:-

(i) The effect of varying the grid size.

(ii) The use of an artificial additional temperature factor to compensate for a coarse grid.

(iii) The use of 2 or 5 term gaussian approximations to the atomic form factors.

(iv) The choice of atomic radius used when modelling the electron density.

The program FCALC was used to calculate a set of structure factors which were then used as the Fobs measurements for input to the Isaacs Agarwal program.

The results of the various calculations are presented in graphical form in the sections below and a table giving the numerical values used in plotting the graphs is given in an Appendix. The R factors shown on the graphs are those calculated after re-scaling the observed to the calculated data. The total number of atoms in the unit cell was 428 and the number of reflections used was 15,331.

4. Definition of Terms

The following terms are used in presenting the results below:-

Hmax - This is the maximum index along any axis. Thus a grid of 3 x Hmax has 3 x hmax points along x, 3 x kmax points along y and 3 x ℓmax points along z. Where exact values could not be used in the calculation the nearest appropriate values were taken.

BADD - The artificial additional temperature factor used to compensate for errors when using a coarse grid for the calculation.

DLIM - The radius squared of the atoms used in the electron density calculation.

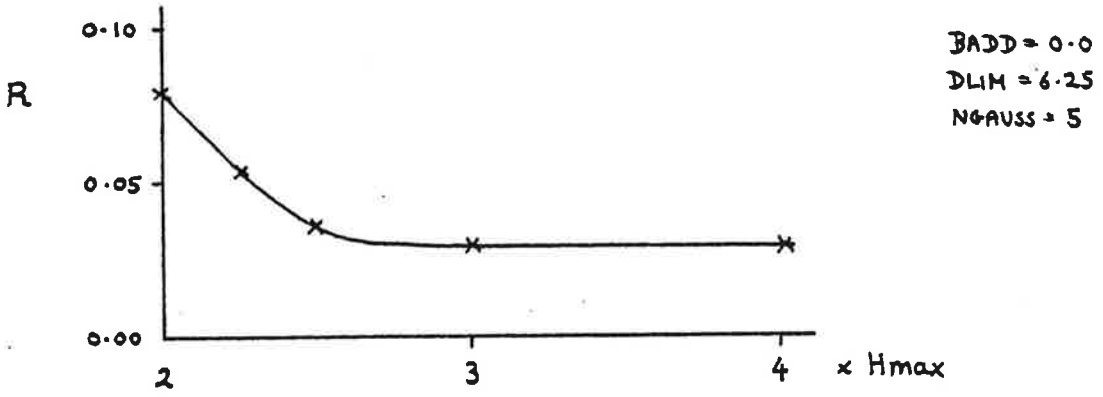
R - The standard reliability index = $\frac{\Sigma|Fo - Fc|}{\Sigma Fo}$

NGAUSS - The number (2 or 5) of terms used in the gaussian approximations to the form factors.

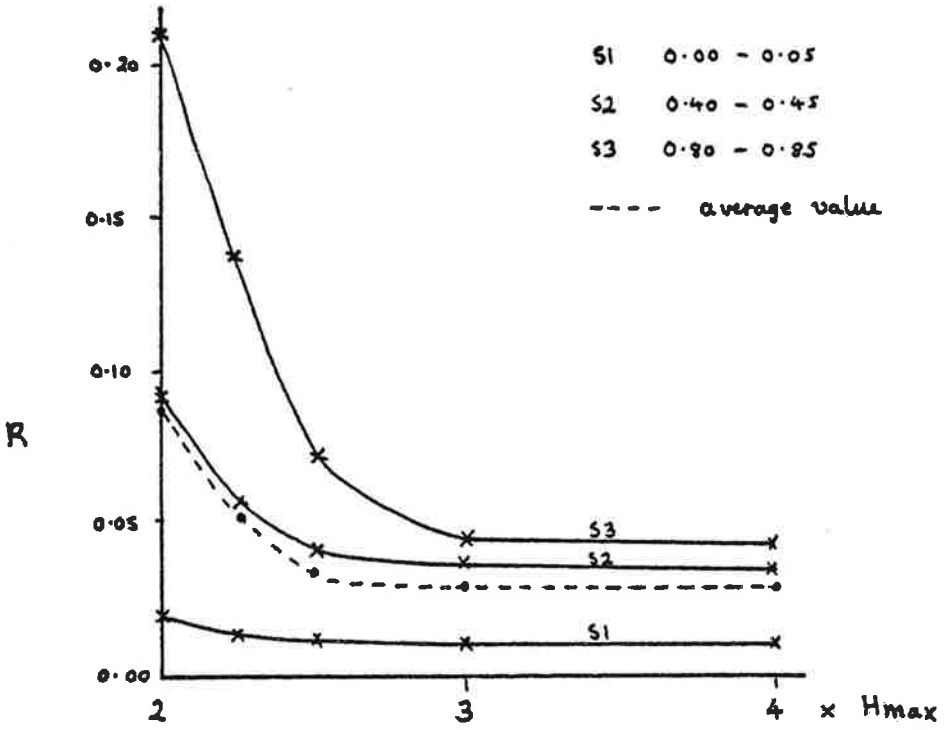
S - $4\sin^2\theta/\lambda^2$

5. The Effect of Grid Size

In this test the grid size was varied between 2 and 4 times the maximum index along each axis. No artificial additional temperature factor was used. The graph below shows the variation of R factor with grid size.



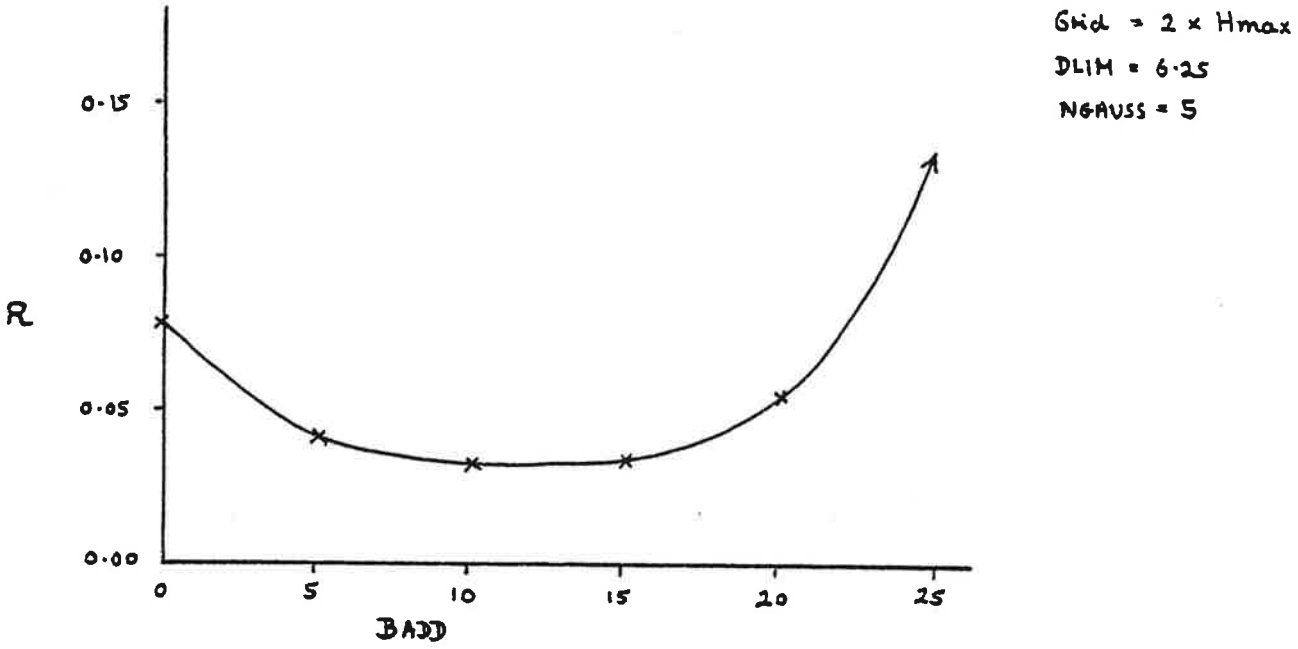
It may be noted that the minimum R value of 0.030 obtained in this test represents the best agreement reached between the results of the two programs. Possible factors contributing to this error are discussed below in paragraph 9. The graph below gives a more detailed analysis of the results by showing the variation of the R factor with S as well as with grid size.



It can be seen that, not only is there an increase in R factor with decreasing grid size but also that the effect is much more pronounced for the higher values of S.

6. The Effect of an Artificial Temperature Factor

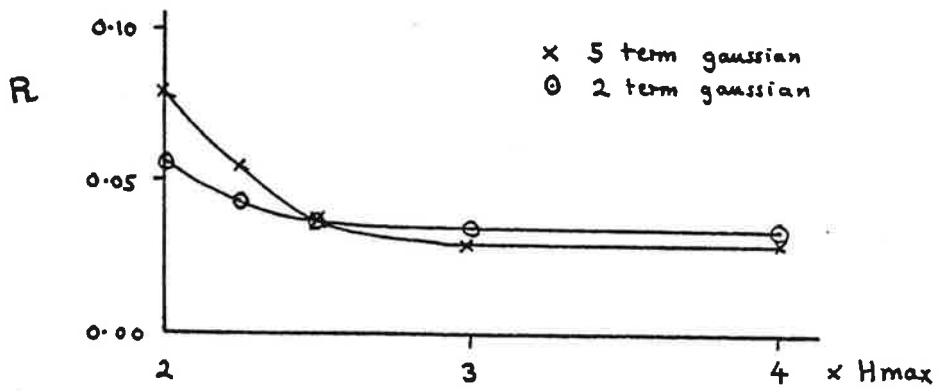
The effect of an artificial additional temperature factor to compensate for too coarse a grid size was investigated using the worst possible grid size of 2 x Hmax allowable for the method. The graph below illustrates the variation of R factors for different values of BADD. The best R factor obtained of 0.034 corresponds to the results obtained with a grid size of 2½ x Hmax.



The results demonstrate that, though the method can be effective, care must be taken in selecting the value of BADD to be used. Too small a value will give insufficient improvement and too large a value can lead to disastrous results (see also the R values for BADD = 30.0 and 50.0 in the table in the Appendix). The use of too large a value of BADD is also indicated by a significant change in the observed/calculated scale factor. The change of scale factor does not however act as a very sensitive indicator that something is amiss. The variation of scale factor with BADD can be seen in the values given in the Appendix.

7. The use of 2 or 5 Term Gaussians

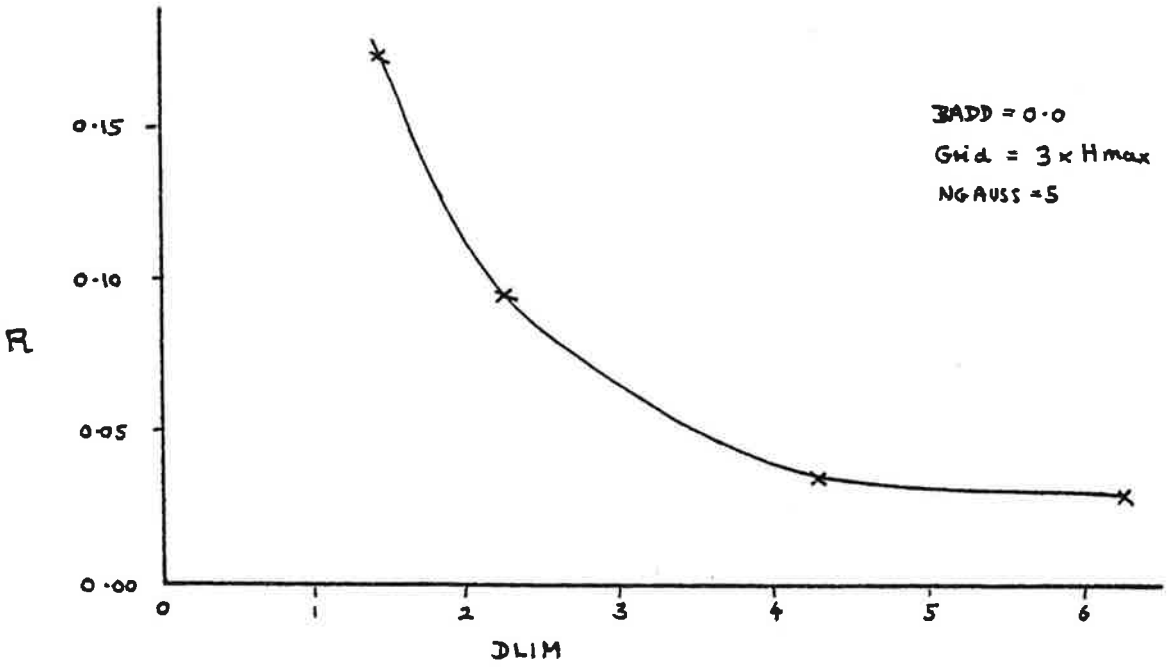
The first test (variation of R factor with grid size) was repeated using 2 term gaussian approximations for the form factors. The graph below compares the results from the two tests.



This test produced the only unexpected result of the study. It indicated that, if a calculation is being done with a reduced grid size to save computing time, there is no advantage in using a 5 term gaussian over a 2 term gaussian approximation. Hence in such cases a double advantage can be gained.

8. The Choice of Atomic Radius

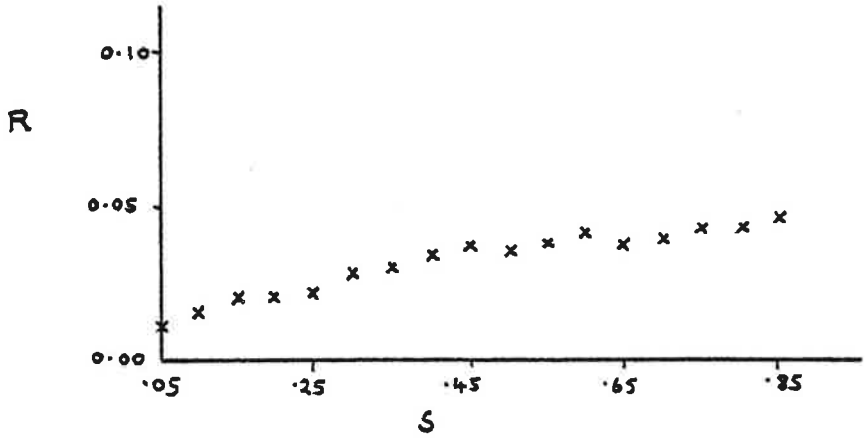
The graph below illustrates the variation of R factor with atomic radius squared.



The results indicate that DLIM may be reduced to about 4.25 before there is a rapid deterioration in the calculated results. Again, in this test, there is a change of scale factor as DLIM is reduced (see results in the Appendix).

9. The Remaining 3%

As mentioned earlier, the best agreement reached between the results from the two methods was an R factor of 0.030. When this figure is analysed further, it can be seen that there is a systematic variation of R factor with S as shown in the graph below.



This may indicate that the main cause of the differences between the programs lies in their approach to the use of form factors.

There is, however, evidence that the results from FCALC are not entirely accurate. First, the F values calculated for the $P2_12_12_1$ systematic absences have significant values (values of 2 to 5 where the average F value overall is ~ 30) and secondly there are differences between the calculated values for equivalent reflections (differences of 1 or 2 in many reflections). For the systematic absences and equivalent reflections looked at, the Isaacs Agarwal results were free of these defects.

There is also a source of error in the F values from FCALC used as the observed F values for the Isaacs Agarwal program as the values are rounded to the nearest integer. These errors would probably only account for a maximum discrepancy of about 1% and that only in the higher S ranges.

10. Conclusion

It would be reasonable to conclude from the study that there are no serious errors in the calculation of structure factors using Eleanor Dodson's version of the Isaacs Agarwal refinement program. More important however, the results should act as a reminder that the calculation of structure factors using the FFT method should be approached with care. It can be used in a safe mode with a grid of > 3 x Hmax, with a DLIM value of 6.25 and using 5 term gaussian approximations for the form factors. Considerable gains in speed can however be made by a careful choice of parameters depending on the accuracy required.

References

{1} Ten Eyck L.F., Acta Cryst., (1977), A33, 486.

{2} Agarwal R.C., Acta Cryst., (1978), A34, 791.

Appendix: Table of results from the structure factor calculation study

Grid	BADD	NGAUSS	DLIM	R	Rsmin	Rsmax	Runsc	Scale
4	0.0	5	6.25	0.030	0.012	0.046	0.030	1.009
3	0.0	5	6.25	0.030	0.012	0.046	0.030	1.009
2.5	0.0	5	6.25	0.035	0.013	0.072	0.036	1.007
2.25	0.0	5	6.25	0.055	0.013	0.138	0.056	1.008
2	0.0	5	6.25	0.079	0.020	0.212	0.079	0.998
2	5.0	5	6.25	0.041	0.012	0.117	0.041	1.007
2	10.0	5	6.25	0.034	0.011	0.083	0.035	1.008
2	15.0	5	6.25	0.035	0.011	0.081	0.036	1.009
2	20.0	5	6.25	0.054	0.011	0.172	0.055	1.013
2	30.0	5	6.25	0.308	0.159	1.603	0.281	0.854
2	50.0	5	6.25	0.915	0.997	0.750	***	0.003
4	0.0	2	6.25	0.035	0.011	0.095	0.037	1.012
3	0.0	2	6.25	0.035	0.011	0.095	0.037	1.012
2.5	0.0	2	6.25	0.036	0.011	0.095	0.038	1.012
2.25	0.0	2	6.25	0.044	0.011	0.128	0.045	1.012
2	0.0	2	6.25	0.056	0.012	0.171	0.057	1.008
3	0.0	5	7.00	0.032	0.023	0.050	0.033	1.006
3	0.0	5	4.25	0.034	0.031	0.047	0.041	0.982
3	0.0	5	2.25	0.095	0.141	0.112	0.242	0.846
3	0.0	5	1.25	0.231	0.327	0.410	0.560	0.712

Key: Grid - the number of points as a multiple of the maximum index along each axis
Rsmin - the R factor for the range S=0.00 to 0.05
Rsmax - the R factor for the range S=0.80 to 0.85
Runsc - the R factor before re-scaling the data

IMPLEMENTATION OF OFF-LINE PROGRAMS FOR ROTATION CAMERA DATA PROCESSING

TREVOR GREENHOUGH (KEELE)

OSCAR

AUTOOSCAR has been tailored to the particular needs of the Daresbury hardware and a full working version is now available. The refinement of misorientation angles and reciprocal cell parameters and c/f distance (seg 1) has been detached into a separate program through personal preference, and will now take mis-setting data from films measured on various measuring devices.

The prediction and plotting of reflections for a given rotation range is also now a separate program which unfortunately suffers from the unsuitability of the Daresbury plotters for this progress. Four poor quality plots are currently produced, of all reflections, fully recorded reflections, start partials and end partials.

The remainder of OSCAR has been amended to take the digital image of a film from the Daresbury scanner on tape (produced by Mike Elder) for variable size scans up to 2380 x 2380 pts. The unforeseen charging of tape I/O to CPU time has led to the required file all being read from tape once to a direct access file (taking some 2 mins. for 2080 x 2080 data pts) with all subsequent operations being performed on the DA file. The tape is controlled via Rutherford SERLAB; DEFINU allows the number of records and record length to be input at run time. The use of DA has much simplified the I/O operations.

The program now allows a raster other than 50 micron (we are currently using 100 micron), and the automatic search for orientation spots is the only option available.

I was impressed by the vector plot of the agreement between observed and calculated spot positions in the Imperial on-line data processing package, so a similar feature has been added to OSCAR. Following integration, an analysis of observed and expected spot positions can be produced, sampling data as

required, along with a plot showing vectors (X50) from observed to expected positions.

All data processed so far has been from flat-plate cassettes; options for cylindrical and V-shaped cassettes are included but not fully tested.

The precession film equivalent of OSCAR (PERCY) has been through preliminary tests and should be available shortly.

Data transferred between programs is on private disc as is the final intensity file (H,K,L,S, partial flag, film, flag) which occupies around 6 tracks for 2000 reflections.

Filmpack

A full working version is available for the merging of two-film packs. A resolution cut-off has been provided. Precession film packs may also be processed.

Results so far have been impressive for the processing of films of the K2PT(CN)4 heavy atom derivative of crystalline 6 - phosphogluconate dehydrogenose from sheep liver taken at the European Molecular Biology Laboratory (storage ring DORIS) Hamburg.

Preliminary investigation of the films at Imperial suggested that they would be extremely difficult if not impossible to process, but the two packs which have so far been through filmpack give merging R factors between films of 0.072 and 0.084 (all data) with a total of 3 bad agreements. R factors for the 3 sigma data are around 4.5%.

We hope the success will continue with the weaker, poorer quality films.

We are continuing implementation of the Oxford off-line suite with ANSC and TRUNCATE underway, and SSM starting shortly.

Further details are available to interested parties.

Report on an NRCC Workshop "Portable Crystallographic Coding: MIR Phasing" which was held at the Lawrence Berkeley Laboratory, California, 31st May - 8th June, 1980

Pella Machin
9th June, 1980

Contents:

Introduction

The first stage of this workshop was held during November, 1979, and it resulted in the cooperative generation of a modular, portable, MIR phasing program. Methods of phasing, details of the XTAL80 system and the use of a preprocessor (RATMAC) were discussed at the first meeting, before participants devoted the remaining time to coding routines in RATMAC for an MIR phasing program for the XTAL system.

The aim of the second workshop was to test the MIR phasing program and to document it.

The participants were:

Steven Freer	(La Jolla)
Robert Munn	(Maryland)
Arthur Olson	(NRCC)
Steven Sheriff	(Los Angeles)
James Stewart	(Maryland)
Jurgen Sygusch	(Quebec)
Keith Watenpaugh	(Seattle)
Gerrard Bricogne	(Cambridge)
Wayne Hendrickson	(Washington)
Pella Machin	(SRC Daresbury)

(The last 3 had not attended the first workshop).

The Workshop

The workshop was held at the NRCC, Berkeley, California. The computer used for most of the work was the NRCC VAX 11/780 and attempts were made in addition to transfer the code to the local CDC 7600. Due to the enthusiasm and dedication of the participants the working day was usually 8.30 (after a stiff climb up the hill) to midnight, with a brief stop for lunch and a longer stop for dinner. In general participants worked away at their particular task and 2 brief meetings were held each day to discuss progress, problems and to generally aid communication.

The main division of tasks was such that the experts on the phasing problem (Sygusch, Hendrickson, Bricogne) when possible debugged existing code with help (particularly from Watenpaugh and Sheriff); the statistics were covered by Sheriff who had written the routines with Ten Eyck and matrix building by Watenpaugh; the remaining participants were kept busy maintaining the system, providing test data, checking RATMAC and CDC activities and debugging XTAL routines. Documentation was an additional topic for everyone.

Although the November, 1979, workshop produced a lot of code, it was not put together and in particular it was not tested within XTAL. Richard Alden (La Jolla) was given the task of assembling the necessary XTAL routines and additional XTAL programs, during the time between the 2 workshops. Due to ill health he was not able to complete this task. Jim Stewart, Steve Freer, Keith Watenpaugh and Robert Munn, spent most of the 2 weeks before the workshop, working day and night to get the program into a state such that the MIR phasing routines could be tested. They almost succeeded, but the first 2 or 3 days of the workshop uncovered some more XTAL-type bugs.

Final Program State

The state of the program at the end of the workshop was as follows:

1. The MVFC method (of Sygusch) was checked out in detail on an ideal data set of 10 reflections. It was then run on a 5A (ideal) data set of 532 reflections of cytochrome C data and it both phased and refined over 4 cycles, giving figures for the 'goodness of fit' of 24, 12, 8, 6 respectively over the 4 cycles.

2. The Blow and Crick method was also checked (Hendrickson and Sheriff) using the 10 reflection test data set and gave comparable phase values. It was run on the 5A data set and produced satisfactory phases but could not continue to refinement since the statistics package was not in a state to provide accurate statistics on which to base the refinement.

3. The statistics routines were debugged partially but not completely. Various line printer plots were successfully produced for example, but other problems remained as mentioned in (2) above. It should be noted that debugging of these routines was somewhat delayed by the initial problems in producing phases.

4. XTAL routines were debugged to a large extent. (Stewart, Watenpaugh, Freer, Machin). Prior to, and during the workshop the problems which arose concerned XTAL input/output and the additional XTAL routines (loading reflection data and atom parameters to the binary data file).

5. RATMAC on the VAX ran well (Munn) with only very occasional problems. Efficient operation of RATMAC on the CDC presented problems which were mainly overcome by the end of the workshop (Munn and Olson).

6. Bricogne's suggested improvement to the phasing method was well received and was thought possible to implement within the framework of the existing code. Hopefully Bricogne will be able to do this during his 6 weeks at NRCC.

7. Some documentation was produced on paper (details from Sygusch, Hendrickson, Bricogne, Sheriff) as well as some user documentation about card formats (Jim Stewart). Part of this (such as definition of routines and common areas) were input to the computer during the workshop (Munn and Machin). Art Olson acted as a naive user to test out any existing documentation on XTAL.

8. 6-PGDH (phospho gluconate de hydrogenase) data provided by Margaret Adams (Oxford) via Pella Machin was used as a real test data set. Art Olson did most of the work involved in running the initial XTAL routines to set up an adequate binary data file for the MIR phasing program, thereby uncovering various documentation and program problems. On the (last + 1) day of the workshop the data ran through the MIR program for 2 cycles but produced ridiculous 'goodness of fit' values - it was thought that the problem concerned the input of variances.

9. Pella Machin and Gerrard Bricogne spent some time producing a set of IBM specific macros for implementation of this program on an IBM computer.

10. Follow up problems were discussed and in particular the following points were noted:

(i) The MIR program would be used immediately in Seattle and La Jolla. Resulting modifications should be passed on initially (during the next 6 weeks) to Olson and Bricogne (NRCC) and then to Jim Stewart (Maryland).

(ii) If possible major modifications should be made by Gerrard Bricogne during his stay at NRCC.

(iii) Machine specific modifications would be made by Jim Stewart.

(iv) With help available at Maryland, Jim Stewart would complete user input documentation and have it transferred to machine readable form.

(v) After the initial 6 week follow up Jim Stewart will coordinate all updates.

(vi) Keith Watenpaugh agreed to write an article about the workshop which might be submitted to Science. A draft of this would be discussed amongst the workshop participants at the Calgary ACA meeting in August, 1980.

(vii) Great interest was shown in the program. In addition to the VAX implementation IBM versions were required by Pella Machin, Jurgen Sygusch and Gerrard Bricogne. Wayne Hendrickson and Pella Machin wanted documentation on the method and algorithm so that they would be able to rewrite the program to run without the XTAL framework, if necessary.

Conclusions

All participants agreed that this MIR program is an improvement on existing codes with respect to the methods and algorithms employed. Three methods of phasing (MVFC of Sygusch, Blow and Crick, Bricogne's modification) are available in the single program. The program has novel approaches to the treatment of variances and of anomalous scattering data.

The RATMAC preprocessor was used throughout and ran well and efficiently, on the VAX. Participants varied in their degree of conversion to RATMAC, some being sufficiently converted that they began to write jiffy programs in RATMAC by choice, others were still struggling to understand existing RATMAC codes! The XTAL system specifications were sufficiently flexible to handle this kind of problem though further debugging and optimisation have to be done before realistic tests can be made to assess how efficienctly the program runs and what the filestore overheads are.

The program will be further tested and used by the groups in Seattle and La Jolla and XTAL routines will be developed further by Jim Stewart in Maryland.

The facilities and organisation supplied by NRCC were excellent, in particular the provision of more than adequate computer machine time on the VAX, without which this second workshop could not have functioned.

It was generally agreed that the provision of a workshop which aimed to produce some computer codes, rather than just discussing program details, was a welcome innovation and in terms of experimentation it was very worthwhile. It may be argued that the cost of such a workshop could not in general be justified for the production of a single program. The future will show whether resulting use of this MIR program within the XTAL system does in fact justify the costs.

MORE PROTEIN REFINEMENT USING THE KONNERT PROGRAM ON THE CRAY-1 COMPUTER

BILL PULFORD (OXFORD UNIVERSITY)

Preliminary work involving use of the Konnert program on the CRAY-1 computer was described in the last newsletter (May). Since then the program has been enhanced in various ways and several proteins have been refined.

I have been working to produce a version of the program which would enable the refinement of phosphorylase. This is now completed and preliminary timings are indicated below. Dr. Stuart Oatley is working on incorporating new substrate molecules, in particular Thyroxine and like molecules, into the refinement program.

Versions of the program system have been obtained from the U.S.A. which include routines for allowing the exploitation of non-crystallographic symmetry, ideal torsion angles and anisotropic B-factors. These will be included in the programs as soon as possible.

Currently there are two versions of the main program which may be run: PROLSQA for proteins of up to 2500 atoms and PROLSQB (which has additional direct access file overheads) for proteins of up to 7000 atoms. Space groups $P2_1$, C2, $P2_12_12$, $P2_12_12_1$ and $P4_32_12$ are allowed.

Now for details of refinements that have been carried out or are in the course of being completed.

Space Group		Natoms	Nrefls	Time (sec)
$P2_12_12_1$	Arabinose Binding Enzyme	2335	22465	318
$P2_1$	PGK	3067	12200	164
$P2_12_12$	Prealbumen	2000	23000	280
$P2_12_12_1$	Triose Phosphate Isomerase	3750	17500	369
$P2_12_12_1$	Tortoise Lysozyme	1100	19700	120
$P4_32_12$	Phosphorylase	6500	19000	769

THE SRC MICRODENSITOMETER SERVICE

MIKE ELDER (DARESBURY LABORATORY)

Since 1974 the SRC has run a microdensitometer service for small molecule crystallography, first at the Atlas Laboratory, now at Daresbury. When this service began Weissenberg films comprised over 90% of the load, and datasets for about 60 crystal structures a year were produced. Now the figure is more like 40 structures, and this decrease in small molecule structures has been matched by a large increase in the number of films from a wide variety of different sources which are sent for digitization and subsequent image analysis. Recently Daresbury augmented this service by purchasing two new instruments. The increased capacity provided by these instruments means that the laboratory should be able to cope with the film scanning requirements arising from the Synchrotron Radiation Source and may be in a position to cope with any protein crystallographic requirements that exceed local capacity.

Daresbury now has two drum scanners: the original (and ageing) Optronics Photoscan interfaced to a Computer Automation Alpha-16 minicomputer, and a new Joyce-Loebl Scandig 3 instrument. This is the machine with the transparent plastic sleeve upon which film samples of up to 25 x 22 cm^2 may be mounted, identical to the one at Imperial College and the newly installed Leeds machine. The scanning raster can be chosen from 25, 50, 100 or 200 μm with square or circular apertures of the same diameter. The optical density range is switch selectable from O-½OD to O-3 OD with an optional zero offset of up to 1 OD. Where the Optronics used a photomultiplier for intensity measurement the Scandig has a photodiode which should give greater stability and has already been proven to give greater linearity of response. The machine is interfaced to a Data General Nova 3/12 with 32K core, twin 5 Mbyte disks and a 9 track 800 bpi tape unit. At present it does not have the hardware to enable FORTRAN 5 to be run, so it is not compatible with the Imperial College software for precession and oscillation photographs, but an upgrade is planned for 1981. At the moment Weissenberg films are still being processed on the Optronics machine, since the assembler language on-line program will take a while to be moved from the Alpha-16, and the Scandig is being used for digitization to magnetic tape.

The second recent purchase is a Joyce-Loebl flat bed film scanner, a Microdensitometer 6. This instrument has a glass plate upon which flat samples up to 25 x 25 cm^2 may be mounted. Stepping motors move this plate in two mutually perpendicular directions in steps of 2.5 μm, and the plate may be rotated manually between scans. The scanning aperture is chosen from a range of 10 which include square apertures of 5 μm to 250 μm diameter and rectangular apertures of 5 x 50 to 50 x 500. The machine is, of course, slower than a drum scanner, with a maximum data transmission rate of 1000 readings/second. Thus a sample 10 cm x 10 cm scanned at 5 μm resolution would produce 4×10^8 density values and take a few days to scan, so the instrument is designed to be controlled by an interactive program. It is interfaced to another Nova 3/12 with dual diskettes and a 9-track tape unit. Joyce-Loebl provide an interactive controlling language similar to BASIC, and programs can be developed quite quickly for specific image analysis applications.

Anyone interested in either of these machines should get in touch with me at Daresbury.

BROOKHAVEN PROTEIN DATA BANK UPDATES

PELLA MACHIN (DARESBURY)

I received a new copy of the Brookhaven protein data bank soon after the last edition of the newsletter was circulated (containing details of the availability of the data bank at Daresbury) (May 1980)!

In order to inform people of future updates to the data bank I have made a file on the IBM 370/165 at Daresbury which can be inspected by computer users. The file XBE.PROTDIR.DATA details the number of the magnetic tape which contains the most recent version of the data bank and the contents of the tape.