

Daresbury Laboratory

JOINT CCP4 AND ESF-EACBM

NEWSLETTER ON

PROTEIN CRYSTALLOGRAPHY



An informal Newsletter associated with the SERC Collaborative Computational Project No.4 on Protein Crystallography and the ESF Network of the European Association of the Crystallography of Biological Macromolecules.

Number 27

November 1992

Contents

CCP4 Suite News and Administrivia from Daresbury D. Love	... 1
MIR with one derivative M. Noble & R. Pauptit	... 3
Haloalkane dehalogenase K.H.G. Verschueren, H.J. Rozeboom, K.H. Kalk & B.W. Dijkstra	... 7
Crystal Structure of Avidin-Biotin Complex at 2.7Å resolution L. Pugliese & M. Bolognesi	... 11
SRS Specialist User Group for Laue Diffraction J. Campbell & B. Kariuki	... 14
Notes on FEBS/EACBM Advanced Course "Techniques in Protein Crystallography"	... 17
Status Report from ESRF C.-I. Brändén	... 22
The Crystallography Workstation for X-Ray Structural Investigations of Single Crystals at the TNK Storage Ring A.N. Popov, D.M. Kheiker & E.H. Harutyunyan	... 25
Availability of the MOSFLM program suite for processing image plate and film data A.G.W. Leslie	... 30
On the Reliability of Atomic Models in the PDB: Stereochemical Point of View A.G. Urzhumtsev	... 31

Editors: Wojciech M. Wolf

Keith S. Wilson

SERC Daresbury Laboratory
Daresbury, Warrington WA4 4AD UK

EMBL c/o DESY
Notkestrasse 85, D-2000 Hamburg 52
Federal Republic of Germany

PROCHECK: A Program to Check the Stereochemical Quality of Protein Structure	
R.A. Laskowski, M.W. MacArthur, D. Moss & J.M. Thornton	... 34
FAN Scalar Field Analysis	
E. Vernoslova & V. Lunin	... 35
FFT. Discrete Three-Dimensional Fourier Transform	
E. Vernoslova	... 36
CAN. Comparative Analysis of Three-Dimensional Functions	
E. Vernoslova & V. Lunin	... 37
3rd European Workshop on Crystallography of Biological Macromolecules, Como, May 1993	... 38
The CCP4 Overview	... 39
Main, Supported CCP4 Programs	... 44
Miscellaneous and Unsupported CCP4 programs	... 48
List of the CCP4 programs which are very likely to be removed from the distribution kit in the next release of the CCP4 package	... 50
Positions available	... 52

The Joint CCP4 and ESF-EACBM Newsletter is not a formal publication and permission to refer to or quote from articles reproduced here must be referred to the authors.

CCP4 Suite News and Administrivia from Daresbury

Dave Love
SERC Daresbury Laboratory,
Warrington WA4 4AD, UK
d.love@dl.ac.uk

Since joining CCP4 earlier this year I have put a lot of work into attempting to improve portability and introducing some software engineering practices to make it, I hope, better maintainable and more reliable. The results of this work may not be immediately apparent, but I hope they will result in wins in the future. This has been a learning process and your indulgence is appreciated where things may have gone wrong; constructive comments are very welcome.

Distribution policy

The distribution of the suite is no longer subject to continual update. Distinct releases will now be made periodically, in line with normal practice for distributions. This, along with the revision control information now kept on the software, should allow us to provide a better service by being able to identify the versions in which problems occur. An effort will be made to test at least the compilation of each release on a reasonable number of systems, too. Major bug fixes will be announced on the ccp4bb e-mail list and patches for them will be made available from the ftp area and the info-server between releases. The first distinct release (2.1) has now been made; this contains the usual bug fixes and enhancements to the programs plus support for AIX, HPUX and ESV Unix systems and a new installation procedure (see below).

Installation/configuration

Unix

There is a new installation system for Unix which, we hope, will make installation in different Unix versions easier and more reliable. It is no longer necessary to edit any files other than the site-dependent ccp4.setup script—the Makefiles and system-dependent code are configured under control of a script which only needs to be told the system type and, optionally, locations of sources and binaries.

VMS

The distribution now contains the necessary files to build the CCP4 library without a C compiler.

X-windows graphics

There is now a tool available to scan the .log files of those programs which have been appropriately instrumented so far and plot the tabulated data under X-windows. This is not part of CCP4 and brought to you courtesy of Jan Zelinka (York/Glaxo). A VMS port of this xloggraph system is not yet available, but we'll be happy to see one if anyone wants to undertake it.

Electronic mail

CCP4BB

There have been problems recently with the CCP4BB e-mail list for CCP4 news etc., for which, apologies. One problem is that various people seem to have become unsubscribed unintentionally for some reason. If you haven't been receiving messages from CCP4BB and think you should, please re-subscribe by sending the message `sub ccp4bb` to address `ccp4bb-requests@daresbury.ac.uk`; (UK people may need to reverse the domain order, of course). Subscriptions are dealt with by an automatic system.

CCP4-DEV

There's an extra list for people who consider themselves CCP4 developers i.e., hack the code significantly. This is for discussing problems and policy, advertising work in progress etc. Subscription and unsubscription requests should be sent to `ccp4-dev-request@daresbury.ac.uk`; they're processed by hand.

List-server

Please note that the crystal request to the list-server (`list-server@daresbury.ac.uk`) is now obsolete—all the VMS source is now kept in the same directories as the Unix code, so use `request: sources` instead.

Ftp distribution

The internet anonymous ftp distribution area (`anonymous@gserv1.dl.ac.uk:pub/ccp4`) has been somewhat re-vamped. VMS users can get tools to deal with the compressed tar (.tar.Z) files and make and apply Unix-style patches from the directory `VMSutils` as C source and executables. The 'aggregated' software is now also available from the ftp area. VMS binaries of the CCP4 programs are not currently available via ftp.

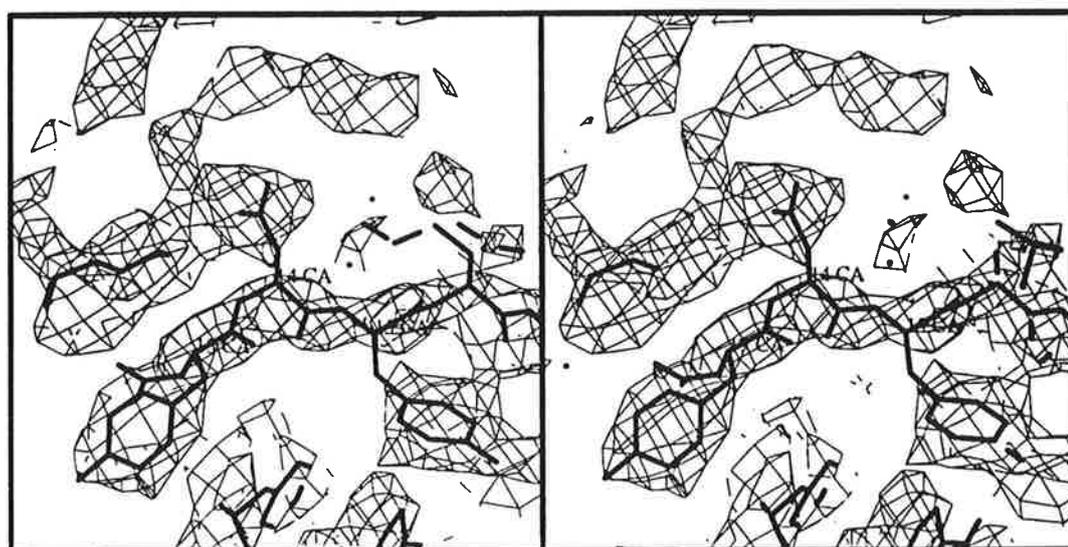
MIR with one derivative

Martin Noble (EMBL Heidelberg, Germany) in%”noble@EMBL-Heidelberg.DE”
& Richard Paupitit(ICI Pharmaceuticals, UK) in%”paupitit@iciukph.uucp”

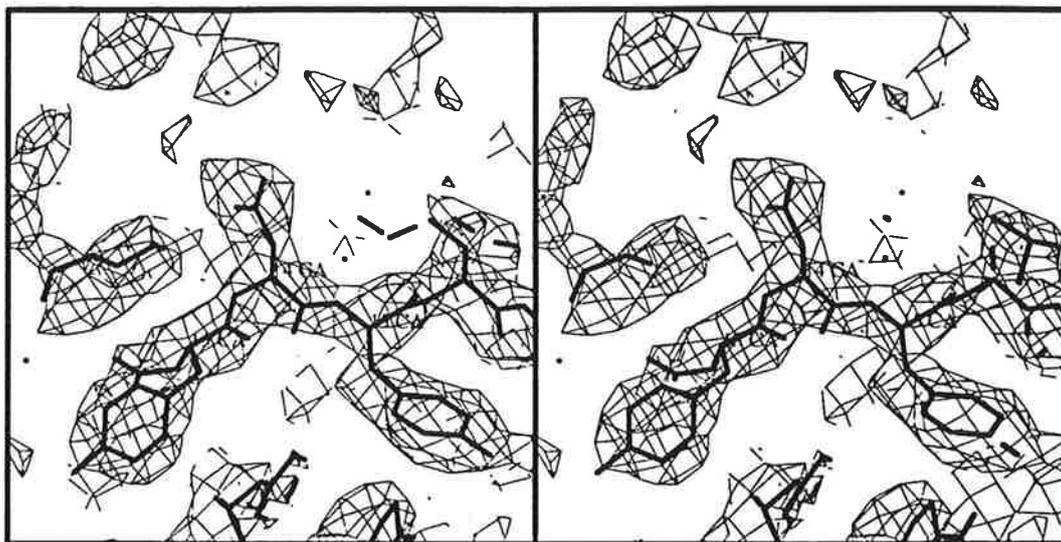
In the structure determination of the SH3 domain from chicken brain spectrin (Nature, in press), only one Pt derivative was found after extensive searching. SIRAS, using data collected on a FAST detector with CuK α radiation, did not lead to interpretable density, although it allowed the correct heavy atom enantiomer to be established. The wrong enantiomer did not allow solvent regions to be discerned. It seemed that the best strategy would be to optimize the anomalous signal using synchrotron radiation. This was attempted twice, but both times the heavy atom occupancy was unfortunately low. Data were recollected in the home laboratory in an attempt to (1) establish conditions for high Pt occupancy and (2) improve the completeness of the Friedel pairs. The second home data set had the same Pt site, different occupancy, and also gave rise to an uninterpretable electron density map. When the two data sets were treated as separate derivatives, the overall figure of merit increased substantially and the map became interpretable, allowing structure solution. The improvement in the map interpretability is illustrated below for one region of the protein.

Figure 1

a) SIRAS electron density using Pt1:



b)"MIRAS" electron density using Pt1 & Pt2:



It is believed that this improvement was due to a statistical improvement in the quality of the isomorphous difference and especially of the anomalous data. A similar effect was observed in the structure determination of p2 myelin (Jones et al, 1988. EMBO J. 7, 1597-1604), where the structure was solved using the anomalous signal of two different derivatives that occupied the same site. The implication is that it may be worthwhile using a second data set as a second derivative whenever an SIRAS map is uninterpretable, or, in fact, always. Table I presents the phasing statistics for the two SIRAS and the "MIRAS" cases.

TABLE I. (space group P212121)

A) Data:

	Native	PT1	PT2
Cell axes	34.1 42.4 50.0	34.1 42.5 50.5	34.0 42.4 50.2
Merging R to 2.6 ang	6.6	3.9	4.5

B) Heavy atom parameters:

	Riso (on F)	X	Y	Z	Occ	B	rms phasing power (acentrics/centrics)
PT1	23.7	-0.130	-0.076	-0.108	1.78	29.4	2.0/1.4
PT2	15.8	-0.131	-0.076	-0.107	1.01	24.1	1.8/1.1

Riso (Pt1 vs Pt2) = 12.9 %

C) Phasing statistics: (fom's in parentheses are without the anomalous contribution, i.e. sir)

	Pt1	Pt2	Pt1 and Pt2
Mean figure of merit (acentrics)	0.44 (.38)	0.47 (.31)	0.71 (.51)
Mean figure of merit (centrics)	0.61 (.61)	0.54 (.53)	0.76 (.77)

The improvement in the quality of the electron density as a result of treating the two data sets as different derivatives could be a result of:

- Improved completeness of anomalously phased reflections
- Differences in cell dimensions of the two derivative crystals (nonisomorphism), so that each provides effectively independent information
- Averaging of the isomorphous difference and anomalous information to improve the signal to noise ratio in phasing.

The Riso values are consistent with the relative occupancies. In particular, the agreement between the two derivative data sets (Riso=12.9%) is greater than the agreement between either derivative data set and the native data, consistent with the smaller difference in relative occupancy (0.77). Thus, we find no indication of nonisomorphism. A breakdown of the observations used in producing the final interpretable MIRAS map at 2.8 angstroms is presented in table II

TABLE II

	Column:	Number of reflections:	Percent of possible
A)	Fo	1807	91.4
B)	Fo and F(Pt1)	1713	86.6
C)	Fo and Dano(Pt1)	1360	68.8
D)	Fo and F(Pt2)	1727	87.3
E)	Fo and Dano(Pt2)	1325	65.7
F)	(B) or (D)	1797	90.8
G)	(B) and (D)	1643	83.1
H)	(C) or (E)	1405	71.0
I)	(C) and (E)	1254	63.4

Either of the two derivative data sets gave rise to SIR phases for about 87 % of the possible reflections, and SIRAS phases for about 67 %. Use of the two data sets together increased these quantities to 91 % (row F), and 71 % (row I) respectively; a relatively minor increase. Most noticeable, is that of the 1797 reflections with calculated phases, 1643 were observed in both derivative data sets, and of the 1405 reflections for which an anomalous signal was available, 1254 had anomalous signals observed in both data sets. Thus the improvement in the map interpretability seems more likely to result from improved quality of phases due to

the use of two derivative data sets, than from complementation of the phased reflections from one data set by the other. This is supported by the statistics in table I, which show a considerably higher figure of merit for the MIRAS phases than for either of the two SIRAS phase sets. The greater improvement when anomalous data are included suggests that the quality of the anomalous data stood more to gain, perhaps due to the weak nature of the anomalous signal.

The use of an additional data collection is encouraged, not as a substitute for careful data collection, but for cases where it becomes apparent that the data are not quite adequate for good phasing. In the example presented here, the improvement was sufficient to solve the structure. We would be delighted to hear of any similar success stories.

Haloalkane dehalogenase

K.H.G. Verschuieren, H.J. Rozeboom, K.H. Kalk, B.W. Dijkstra

Lab. of Biophysical Chemistry and BIOSON-Research Institute

University of Groningen

Nijenborgh 4, 9747 AG Groningen, the Netherlands

Introduction.

Several years ago the soil bacterium *Xanthobacter autotrophicus* GJ10 was found to be capable of growth on a medium containing 1,2-dichloroethane or 2-chloroethanol as a sole carbon- and energy source. In order to utilize these compounds, the organism produces at least two different enzymes that release chloride from halogenated chemicals : one specific for halogenated alkanes (haloalkane dehalogenase), whereas the other acts on halogenated carboxylic acids (haloalkanoate dehalogenase). The haloalkane dehalogenase has been cloned, expressed, purified and characterized by the group of Dick Janssen, Lab. of Biochemistry, Groningen. The enzyme converts haloalkanes to their corresponding alcohols without any requirement for oxygen or co-factors. A number of different halogenated haloalkanes are hydrolyzed by the enzyme, among which are the environmentally important chlorinated hydrocarbons methylchloride, ethylchloride and 1,2-dichloroethane. The three-dimensional structure was initially solved by Sybille Franken (1991) and has now been refined at 1.9 Å resolution.

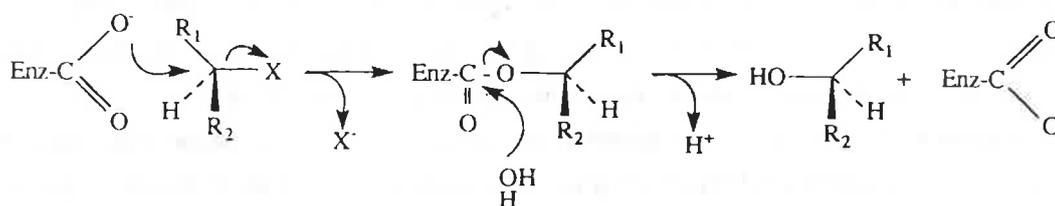
The putative active site is in an internal, buried cavity with catalytic residues Asp124, Asp260 and His289. The other residues lining this cavity are hydrophobic. A tunnel is stretched out from the solvent to this cavity, but it seems to be blocked by the side-chain of Leu262. Asp260 and His289 form a pair as has been observed before in other enzymes (e.g. the serine proteases, phospholipase A2). The cavity has a size perfectly fit to accommodate a substrate molecule like 1,2-dichloroethane, as we could show by docking experiments.

Dehalogenases offer an intriguing perspective to be applied as an " anti-pollutant ", as well as to be used in the synthesis of chiral compounds. Our research focuses on the elucidation of the reaction mechanism, and the factors that determine the substrate specificity. Long time goals (in collaboration with the group of Dick Janssen) are to broaden the enzyme's substrate specificity towards longer chain haloalkanes (both by site directed mutagenesis, and by selection of " in vivo " mutants), to improve the enzyme's catalytic activity and to make the enzyme applicable for the synthesis of chiral compounds.

The catalytic mechanism.

The activity of the enzyme is optimal at pH 8, at pH 6 only some residual activity is found. The initial structure was determined at pH 6.2. Afterwards, crystals at pH 8.2 were obtained by soaking the native crystals in mother liquor of pH 8.2. This procedure did not have any dramatic effect on the crystals. The three-dimensional structure of the dehalogenase at this new pH remained the same as the native structure, except for the H-bond between Asp124 and His289 which has been weakened at pH 8.2 (at pH 6.2 : Oδ1 Asp124 - Ne2 His289 = 2.71 Å ; at pH 8.2, this distance becomes 3.3 Å). Thus, at the pH of optimal activity, His289 is mostly deprotonated. From this observation, the architecture of the active site and the

position of a substrate molecule docked into the cavity, we could propose a reaction mechanism: the carboxylic side chain of Asp124 performs a nucleophilic attack on the carbon atom of the substrate to which the halogen is bound (the O δ 1 atom is about 3.1 Å away from the attacked carbon atom of the modelled substrate). Two tryptophan residues (Trp125 and Trp175) point with their ring nitrogens towards the cavity in such a way that they could stabilize the resulting halide ion as well as a negatively charged five-coordinated carbon in the transition state (as occurring in a binuclear nucleophilic substitution reaction). A covalently bound intermediate would be formed which is an ester (formed between the Asp124 side-chain and the substrate). This ester can subsequently be cleaved by a water molecule in a nucleophilic attack on the C γ atom of the Asp124 side-chain, and the alcohol is released.



Proposed reaction mechanism.

To prove this proposed catalytic mechanism, we performed several experiments :

1. *Soaking with halide compounds : halide binding site.* From the literature it is known that iodoacetamide inhibits the dehalogenase. Previously, this finding was taken to indicate that the enzyme should have a cysteine residue in the active site. Because there is no cysteine present in our proposed active site, we were curious to find out how iodoacetamide inhibits the enzyme. Therefore we soaked native dehalogenase crystals in a iodoacetamide solution and determined the 3D-structure (2.2 Å resolution). A large spherical blob of electron density, consistent with a iodide ion, appeared between the ring nitrogens of Trp125 and Trp175. No density could be seen for the acetamide part of the reagent. A similar result was obtained by soaking crystals in chloroacetamide, both at pH 6 and pH 8 : no evidence for the acetamide moiety in the active site, but a chloride ion near the side chains of Trp125 and Trp175. This result could be further corroborated by soaking the crystals in a 5 mM iodide solution at pH 8.2. Again, a spherical density for a iodide appeared between the Trp125 and Trp175. It is not clear to us how iodoacetamide could inhibit the enzyme. The easiest explanation would be that halide ions (either present as a contamination, or produced by the enzyme from the halo-acetamide) cause product inhibition. Indeed, our chloroacetamide preparations do contain chloride ions, as we could determine by a colorimetric method as described by Bergmann and Sanik (1957).

2. *Soaking with 1,2-dichloroethane.* Our proposed reaction mechanism suggests, that, in principle, it should be possible to trap the putative covalent intermediate at low pH. If His289 is protonated, it can no longer take up a proton from the nucleophilic water molecule, and therefore hydrolysis of the acyl-enzyme may be slow enough to accumulate the covalent intermediate.

Therefore, we performed soaking experiments of native crystals at pH 6.2 and 5.7, in the presence of the substrate 1,2-dichloroethane. After solving the structures, it again appeared that a spherical electron density was present between the ring nitrogen of Trp125 and Trp175, indicating that at these pH's the enzyme still had some residual activity, and had produced a chloride ion. Interestingly, in contrast with the experiments with

the halo-acetamide, in this structure the nucleophilic water molecule has disappeared. This suggests that the first step of the reaction has taken place, resulting in the acyl-enzyme, and that, subsequently, the acyl-enzyme has been hydrolyzed by the nucleophilic water molecule. The alcohol has already left the cavity, but the chloride ion is still present. This means that first the alcohol product leaves the active site, and afterwards the halide ion. Probably, after the chloride has diffused out of the active site, the nucleophilic water is replenished. We see in the cavity extra density for an incoming water molecule. This result should be further confirmed by kinetic measurements.

Because at pH 5.7 the enzyme had still some residual activity, we decided to repeat the same experiment at pH 5.0. The difference in electron density between the native, pH 6.2 structure and the structure with substrate at pH 5.0, now indicates that the reaction proceeds very slowly. We see a mixture of probably four states and after occupancy refinement of this structure (2.0 Å resolution, R-factor = 18.75 %), we could assign :

- 10 % of the molecules in the crystal have no substrate bound,
- 25 % are in the acyl-enzyme state (O δ 1 Asp124 has attacked the substrate) with the nucleophilic water molecule still at its position,
- 60 % of the enzyme just after cleavage of the ester bond with the hydroxyl of the nucleophilic water molecule attached to the side chain of the Asp124 ; the alcohol product has already left the cavity but the chloride ion is still present between the two Trp's.
- 5 % of the enzyme is in the state as we see it in the soaking experiments at pH 6.2 and pH 5.7.

So far, this result is in agreement with the hypothesis for the reaction mechanism.

Evolutionary aspects of the structure.

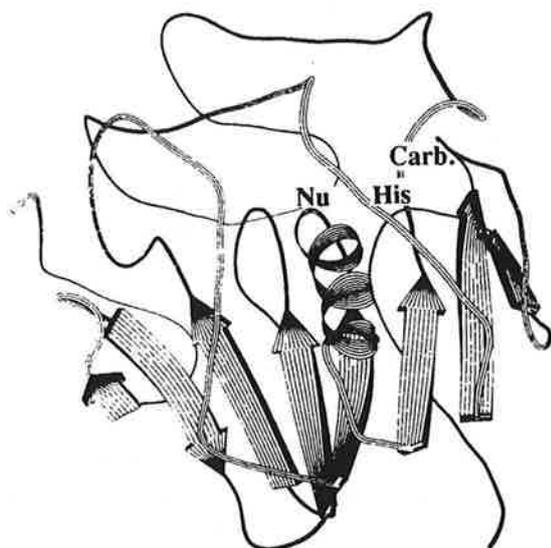
The topology of the secondary structure elements in the dehalogenase is similar to that of a number of other enzymes like dienelactone hydrolase (Pathak *et al.*, 1988), a lipase from *Geotrichum candidum* (Schrag *et al.*, 1991), a Ser-carboxypeptidase from wheat (Liao *et al.*, 1990) and acetylcholinesterase from *Torpedo californica* (Sussman *et al.*, 1991). A comparison of the five structures has appeared recently (Ollis *et al.*, 1992).

All five enzymes have a central catalytic domain of unique topology and three-dimensional structure ("the α/β hydrolase fold"). In four of the five proteins the topology of the central eight strands is identical, while in carboxypeptidase a β -hairpin loop is inserted between the seventh and eighth strand. All five proteins possess catalytic triads (e.g. Ser-His-Asp), and, in each case, the triad residues occur at the same topological location, and at similar relative positions in the primary structure (see table).

protein	nucleophile	carboxylate	His
dienelactone hydrolase	Cys123	Asp171	His202
haloalkane dehalogenase	Asp124	Asp260	His289
Ser-carboxypeptidase	Ser146	Asp338	His397
acetylcholinesterase	Ser200	Glu327	His440
lipase (<i>Geotrichum candidum</i>)	Ser217	Glu354	His463

The α/β hydrolase fold enzymes.

As is typical of the active sites in α/β proteins, the nucleophile is located at the strand crossover point of the parallel β -sheet. The five enzymes constitute four different structural families of α/β hydrolase fold enzymes, the acetylcholinesterase and lipase have very similar sequences (24 % identity), and have rather similar structures. The four groups have no sequence homology and different amounts of structural similarity.



Nevertheless, all members have a similar arrangement of the central eight β -strands : their β -strands are superhelically twisted so that the surface of the sheet covers about half a cylinder, and the first and the last strand cross each other at an angle of about 90° . Although the nature of the nucleophile (Nu.) varies, it is always the central residue of an extremely sharp γ -like turn between β -strand 5 and the subsequent helix. This strand-nucleophile-helix feature, also called the "nucleophile elbow", is the most conserved structure in the α/β hydrolase fold. Also the carboxylate (Carb.) and the His residue are in locally conserved regions at the end of β -strand 7 and 8, respectively. The similar position in sequence and three-dimensional structure of the catalytic

triad residues suggests that also the catalytic mechanisms will resemble each other. All five enzymes catalyse a hydrolytic reaction, in which an oxy-anion is an intermediate. It appears that in all the enzymes the oxy-anion hole is a small cavity between the end of the nucleophile elbow strand and the nucleophile. The negative charge of the tetrahedral intermediate in the hydrolysis of the acyl-enzyme is then stabilized by the peptide NH of the residue following the nucleophile (in haloalkane dehalogenase this is the NH of residue Trp125).

Thus, this family of hydrolytic enzymes provides an example of divergent evolution, in which the arrangement of the catalytic residues has been conserved, but the substrate binding site has been adapted to the specific requirements of the substrate.

Literature.

- Bergmann, J.G. and Sanik, J. (1957) *Anal. Chem.*, **29**, 241-243.
 Franken, S.M.,Rozeboom, H.J., Kalk, K.H. and Dijkstra, B.W. (1991) *EMBO*, **10**, 1297- 1302.
 Liao, D. and Remington, S.J. (1990) *J. Biol. Chem.*, **265**, 6528.
 Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G. and Goldman, A. (1992) *Protein Engineering*, **5**, 197-211.
 Pathak, D., Ngai, K.L. and Ollis, D. (1988) *J. Mol. Biol.*, **204**, 435-445.
 Schrag, J.D., Li, Y., Wu, S., Cygler, M. (1991) *Nature*, **351**, 761-764.
 Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. and Silman, I. (1991) *Science*, **253**, 872-879.

CRYSTAL STRUCTURE OF AVIDIN-BIOTIN COMPLEX AT 2.7 Å RESOLUTION.

L. Pugliese* and M. Bolognesi*#.

*Dipartimento di Genetica e Microbiologia. Università di Pavia - Italy.

IST - Genova - Italy.

Avidin is a tetrameric protein found in hen-egg white. Each monomer binds one molecule of biotin and the K_d of the complex is 10^{-15} M. The strong interaction between avidin and biotin is the basis of all the many biotechnological application of the complex. Some examples of its uses are: sensibility enhancement of immuno-assay, protein purification and analysis of membrane sites.

The avidin-biotin complex was crystallized from 2.0 M ammonium sulfate solutions with phosphate buffer at pH 5.7. The crystals obtained belong to space group $P4_22_12$ with unit cell edges $a=80.15$ Å $c=85.28$ Å, two molecules per asymmetric unit (1).

A first data set was collected on a Siemens area detector installed on a Rigaku rotating anode at the Weizmann Institute of Science - Rehovot - Israel.

Since after structure solution and refinement on that data set it was found that only 60% of the binding sites were occupied, a second data was collected

at the NCI - Frederick (MD) - USA, on one crystal previously soaked in a solution saturated with biotin. Both data set were processed using the Xengen program package.

The first data set was used to solve the structure by molecular replacement using an avidin coordinates set model built on the basis of an avidin aminoacid sequence, of the streptavidin three dimensional structure, and of low resolution crystallographic data of an orthorombic egg-white avidin crystal form (2), as a search model. The program suite Merlot was used both for the rotation and the translation searches. The first and the third peak found in the rotation search were shown to be related by the local two fold axis relating the two monomers in the asymmetric unit. The first peak was then used for the translation search in which seven consistent vectors were found.

After release of some tight contacts between two loops of symmetry

related molecules the refinement was started by alternating cycles of conventional restrained minimization, using the TNT program package, with model building with FRODO. The R factor decreased from an initial 0.471 to 0.168 in 16 cycles with r.m.s. on bond length of 0.020 Å and of 3.00 degrees on bond angles. The structure of the complex with full occupancy (second data set) was then refined in the same way to a final R factor of 0.168 (10 - 2.7 Å resolution), with r.m.s. on bond length of 0.020 Å and of 2.9 degrees on bond angles.

As a general fold, avidin is an eight stranded β barrel and the biotin binding site is deeply buried inside the barrel. In the internal part of the barrel there are several polar groups able to form hydrogen bonds with biotin ureidic ring and among themselves. The biotin binding sites are placed on two opposite sides of the tetramer and show several water molecules. In Figure 1 the $C\alpha$ trace of the tetramer is represented while in Figure 2 there is the $C\alpha$ trace of one avidin monomer with its biotin bonded.

Figure 1: STEREO VIEW OF THE TETRAMER.

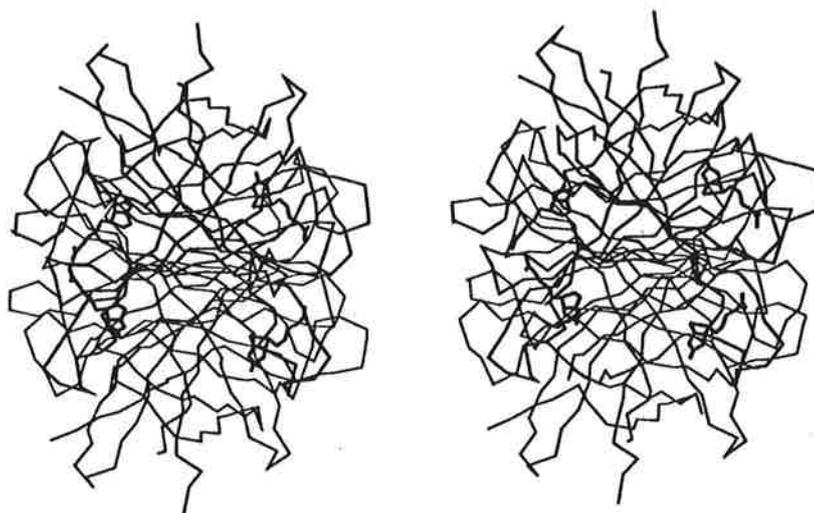


FIGURE 2: STEREO VIEW OF THE MONOMER IN TWO DIFFERENT ORIENTATIONS.



SRS SPECIALIST USER GROUP FOR LAUE DIFFRACTION

John Campbell (Daresbury Laboratory) & Benson Kariuki (University of Liverpool)

The first meeting of a new SRS Specialist User Group for Laue Diffraction was held at The Daresbury Laboratory on July 9th, 1992. It was chaired by Dr Marjorie Harding from the University of Liverpool. There was an encouraging response with more than 35 participants from both Biological Science and Material Science fields.

Experimental Facilities

Station 9.7

Simon Clark (Daresbury) summarised the facilities on station 9.7 including the water cooled slits, the experimental bench, the 30 millisecond fast shutter and the single and six film carousels. He said that a camera was currently being modified at Keele University to incorporate a diode array photometer.

He reported that they were working on the development of small furnaces, high pressure cells and 4-circle software to drive the camera. Two main areas of development were planned, setting up facilities for combined EDPD/single-crystal experiments and the use of capillary focussing optics.

A number of issues were discussed including the possible use of image plates on the station.

Station 9.5

Sean McSweeney (Daresbury) summarised the facilities on station 9.5 which was designed for monochromatic and MAD data collection as well as for focussed beam Laue experiments. A toroidal focussing mirror was used and the wavelength range was 0.45 to 2.6 Angstroms.

In addition to a film camera, a MAR image plate system was installed with an 18 cm diameter radially scanned plate with 1200x1200 pixels each 150 micrometres square. There was approximately 2 minutes required between exposures to allow for scanning and erasure times.

A new Arndt-Wonacott camera without a mu circle was being built; this would enable minimum crystal to image plate distances to be reduced from about 120mm to 65mm.

There was a programmable fast shutter (down to 30 milliseconds) and an ultra-fast shutter with a minimum opening time of 80 microseconds, programmable up to hundreds of milliseconds. Work was underway to enable the ultra-fast shutter to be used for longer opening times so that it could be used for all purposes.

Future developments planned were a 30cm MAR image plate system with a programmable radius of scan (this was ordered and should be available during the next beam allocation period) and Helium cones for the image plate system. Other possible areas for development included better ways

to characterise the lambda-min and lambda-max values using semi-transparent mirrors, new slits to collimate for very small crystals and a flash lamp facility.

Software Suite

John Campbell (Daresbury) described the current state of the Daresbury Laboratory Laue Software Suite. The following points were covered:-

- a) A new release of the Laue software was available; many of the programs had been made more portable and the suite no longer required the use of the NAG and GHOST libraries.
- b) The suite was now available to Academic users via an Anonymous FTP account at Daresbury subject to accepting the licence conditions. Details are given below.
- c) Some minor programs which had been added to the suite were described. Particular interest was shown in the program GNOMCONV which converted a Laue image to a corresponding gnomonic projection image.
- d) Test data sets were now available and distributed with the suite; a revised version of the Laue Demo, based on these data sets, was being prepared.
- e) It was agreed that the replacement of LCF files with MTZ files should be done as soon as possible.
- f) The set of programs, developed by Paul Carr, for using Gnomonic projection methods, was now distributed with the Laue software suite.
- g) Ian Clifton (Oxford) had specified new files to replace .gen and .gel/.ge2 files. These would be examined in more detail before incorporating them in the Daresbury version of the suite.
- h) The current state of development of the X-windows based program LAUEGEN was described. This incorporated the facilities of NEWLAUE, GENLAUE and SPOTIN and had image display facilities. A pre-release version would soon be made available to a few interested parties.

Anonymous FTP Use

The Laue suite is now available to Academic users who have signed the Daresbury Laboratory Laue Software Suite Licence agreement (or the CCP4 Program Suite Licence agreement) via 'ftp' from Daresbury. If you are eligible, you should first get the LAUEINSTALL file via the 'ftp' route as follows and then follow the further instructions in that file.

The logon 'id' is ftp and you are asked to give the name of your computer system as the password (this is not checked).

The ftp address is 148.79.80.10 and the session would be as follows (password 'dummy' used for machine name):

```
ftp 148.79.80.10
```

```
Name: ftp
Password: dummy
ftp> cd pub/ccp4
ftp> get LAUEINSTALL
ftp> quit
```

Scientific Section

Liz Duke (University of Oxford) described time dependent Laue diffraction studies on Glycogen Phosphorylase. A method of reaction initiation, using the photolabile 3,5-dinitrophenyl phosphate, had been developed. The need for the deconvolution of multiplets to obtain more complete data was emphasised in her talk.

It was noted that a new approach for unscrambling multiplets which involves the use of direct methods is being developed by Quan Hao (Liverpool/Manchester/Daresbury).

Trevor Greenhough (University of Keele) talked about his visits to the USA and France. A conclusive comparison between the Daresbury and CHESS Laue suites is yet to be carried out but a general satisfaction with the Daresbury software has been expressed. In the USA, there are indications that film will be commonly used until the image plate is proved on large molecules. Fast single bunch and undulator experiments were also described.

Axel Scheidig (Max-Planck Institute, Heidelberg) described a time resolved study on the Ha-ras product 'p21' currently in progress. Conformational changes, coupled to the intrinsic GTP hydrolysis, were being investigated. Acceptable Laue data sets at six different time steps had been processed.

Marjorie Harding (University of Liverpool) briefly compared stations 9.5 and 9.7 for small molecule data collection. Substantially larger numbers of useful reflections can be measured on station 9.7 because of the availability of wavelengths in the range 0.25 to 0.5 Angstroms (one third more reflections with $F > 5 \cdot \sigma(F)$ in the example given).

Durward Cruickshank informed the meeting that the proceedings of the January meeting on Kinetics and Macromolecular Crystallography at the Royal Society were to be published in book form by the Oxford University Press and should be available in September.

Next Meeting

The next meeting would be planned for April or May 1993.

THEORY OF THE CRYSTALLIZATION OF BIOLOGICAL MACROMOLECULES

Richard GIEGE, Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique, 15 rue René Descartes, F-67084 Strasbourg Cedex, France.

The historical background of crystallization methods in biochemistry and structural biology, as well as the general principles governing crystal growth of macromolecules, will be presented. The multiparametric nature of crystallization experiments will be discussed. Prenucleation, nucleation, growth and cessation of growth steps will be defined and methods used to monitor these steps will be presented. The use of phase diagrams (knowledge of solubilities) will be emphasized. Finally, the concept of purity and examples that illustrate how small amounts of contaminants can affect nucleation and growth will be discussed.

Selected readings :

Generals : (1) A. McPherson (1982) Preparation and analysis of protein crystals. Wiley, New York ; (2) C.W. Carter (ed) (1990) Methods, A Companion to Methods in Enzymology, 1, 1. (3) P.C. Weber (1991) Physical principles of protein crystallization, *Adv. Protein Chem.* **41**, 1-36 ; (4) A. Ducruix & R. Giegé (eds) (1992) Crystallization of Nucleic Acids and Proteins, A practical approach" IRL Oxford University Press.

See also the four special issues of J. Crystal Growth (**78**, 1986, 529 ; **90**, 1988, 1 ; **110**, 1991, 1 and 1992 in press).

History : eg Chapter 1 of ref 4.

Theory : eg A.A. Chernov (1984) Modern Crystallography III, Springer Series in Solid-State Sciences 36, Springer-Verlag, Berlin ; A.S. Myerson & K. Toyokura (1990) Crystallization as a Separation Process, ACS Symposium Series 438, Washington DC.

Physical-chemistry : eg M. Skouri, M. Delsanti, J.P. Munch, B. Lorber & R. Giegé (1991) *FEBS Lett.*, **295**, 84-88 ; Chapters 9 & 10 of ref. 4.

Purity : eg C. Abergel, M.P. Nesa & J.C. Fontecilla-Camps (1991) *J. Crystal Growth*, **110**, 11-19 ; Chapter 2 of ref 4.

FEBS/ESF-EACMB Advanced Course "Techniques in Protein Crystallography"

9-18 June 1992, University of Aarhus

The above mentioned course was held at Department of Chemistry, University of Aarhus. It was mainly supported by the Federation of European Biochemical Societies (FEBS) through a substantial grant to the course itself and through Youth Travels Funds (YTF) covering travel and accomodation for 20 students. The course was also supported by the European Science Foundation (ESF) through the European Association of Crystallographers in Molecular Biology (EACMB). The course was finally supported by companies supplying equipment for use at the course. These companies were: Evans & Sutherland Computer Ltd. (Three ESV3+ Graphics workstations), Silicon Graphics A/S (Three INDIGO workstations), Leica A/S (Two stereo microscopes).

APPLICANTS

Shortly after the deadline of March 1st 1992 there were exactly 100 applicants to the course. Of these 10 were applicants from outside the FEBS countries and 21 were over the age limit or judged to have none or very little background of basic crystallography. The remaining 69 applicants were from the FEBS countries as follows: Belgium (1), Denmark (1), England (11), France (8), Germany (11), Greece (2), Ireland (3), Israel (3), Italy (3), Norway (2), Poland (1), Portugal (1), Russia (4), Scotland (1), Spain (2), Sweden (10), Switzerland (1), The Netherlands (3). The 22 participants were selected according to the rules: a) as many countries as possible should be covered, b) knowledge of crystallography should be documented by career or by recommendation and c) experiences achieved should be readily applied to projects after the course.

The selection process was really hard and had to go through several rounds. The number of applicants points to the fact that there is a real and unfulfilled need for courses of this type.

ORGANISATION

The course was held at a small conference centre in the campus of University of Aarhus. There was given three one-hour course lectures in the morning at this centre. After lunch practicals were held at Department of Chemistry in two three-hour sessions with different subjects. After dinner normally two one-hour evening lectures were given at Department of Chemistry. The content of these were the present scientific projects of the invited lecturers. Free discussions were possible late in the evening.

The course was thus a very intensive one, with indeed very little free time for the students. This left only little time for the students to have a completely free discussion with lecturers or each other. The last day of the course was left for such free work in the 'Open Shop' sessions, but unfortunately many of the lecturers had then left the course.

CONTENT

The course covered most of the techniques relevant today for a complete structure determination of macromolecules. The lectures (and partly the practicals) were organised in a fairly logical way starting with crystallization and ending with refining the model for the structure. Most lecturers gave two course lectures.

R. Giegé gave the basic theory of crystallization pointing out the difference between the conditions giving nuclei and the conditions necessary for growth. C. Carter talked about the techniques of crystallization and gave an overview of his practical techniques for searching the multi-dimensional space of conditions. W. Welte gave an overview of the use of detergents in crystallizing membrane bound proteins and stressed the fact that many different detergents should be tried for a given protein. T.A. Jones talked on modelling protein structures on the basis of electron density maps using computer graphics. The importance of good maps, of use of fragments of known structures and of experience in model building was indicated. P. Fitzgerald gave an introduction to the molecular replacement method, which is in increasing use for homologous structures. J. Pflugrath gave an overview of datacollection using areadetectors and mentioned that quality of data is absolutely essential to the success of structure determination at all levels. K. Wilson gave examples of datacollection using synchrotron radiation very strongly stressing the importance of collection of complete high quality datasets. H.A.S. Hansen discussed techniques for collecting data using cryo-techniques. He showed that flash-cooling is the technique of choice in projects where radiation damage can be a problem. E.J. Dodson gave an introduction to the CCP4 programme package strongly advocating the spread of this to Europe, the importance of common datastructures and the usefulness of the cooperative ideas behind the CCP4 project. A. Liljas gave a good overview of the Patterson methods especially with respect to solving heavy atom difference Patterson maps. G. Bricogne presented a series of five brilliant lectures on Fourier theory and on the phasing problem in crystallography covering most relevant subjects in these very important areas of protein crystallography. A. Brünger lectured on refinements of models of protein structures stressing the fact that many indicators for the quality of the model should be investigated.

Obviously not all subjects relevant for all kinds of structure determinations could be covered in a short course like this one. On the other hand the students had a fairly useful overview of the most commonly used techniques, and will in the future be able to take advantage of such methods as they have heard about in the course.

CONCLUSION

As already mentioned there is a definite and unfulfilled need for courses like this one, where lectures and practicals are mixed, so that students can get an overview of the field and also get many useful hints on how to implement some of the practical details. This statement is thus meant as a suggestion for european funding agencies to support in the future similar courses on a yearly basis.

There is however a potential problem in the rules for support via YTF that FEBS is using. Some of the applications from young european scientists at present employed abroad stated that support for travel could be raised at their present laboratory, but that they would like to have a grant covering accommodation and other costs at the course. Such a type of support could also benefit applicants from the country of the course.

ORGANIZERS

J. Nyborg, M. Kjedaard and S. Thirup, Dept. of Chemistry, Univ. of Aarhus, Denmark.
K. Wilson, EMBL, DESYLAB, Hamburg.

POST SCRIPTUM

Although the course was very intensive, some of the students took the time to publish two final issues of a FEBS Newsletter. With the risk of creating some confusion about the real topics of the course these are enclosed.

FEBS Newslett.

Volume 1, Final Issue

Tuesday, 16 June 1992 -- Price: 2 Beers

Editorial staff: Neil "Spink" Spink, Paul "Big-balls" Hempstead & Gerard "Two-Cheeseburgers-Small-fries-(no-ketchup-please)-and-a-large-vanilla-shake" Kleijwegt

What's on today ?

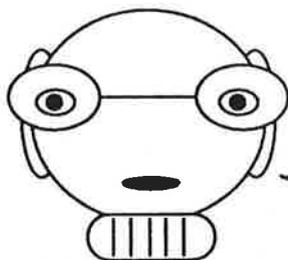
Summer Course programme for Friday, 19 June:

7:30 - 9:00 > Breakfast (so we've been told)
9:00 - 12:00 > Lectures (see above)
9:15 - 11:55 > Alternative: Open shop at the beach
11:55 - 12:00 > Stupid-question time
12:01 - 13:00 > Exquisite banquet (caviar, etc., no wine)
13:00 - 16:00 > Practical session I: Spelling of common precipitants 1 - Precipitants beginning with "Na"
16:00 - 16:23 > Coffee (free), sodas (4 Kr), beer (while stock lasts) & Danish pastry (yummy)
16:22 - 16:30 > Eight-minute break
16:30 - 16:32 > Field trip (visit the historic Aarhus University toilets; Paul says they're great)
16:33 - 18:50 > Practical session II: O17 - Spinking an atom (While we're at it, when did YOU last save your data?)
18:50 - 19:00 > Daily historic walk on the Aarhus University campus
19:00 - 20:30 > Cheeseburgers and shakes at MacDonald's (alternative program: gammon, boiled potatoes, peas, spam, corn, spam, more peas and spam and FREE water (ice cubes extra))
19:39 - 19:45 > World War II (don't mention the war, Gregers !)
20:30 - 22:45 > European "World Series" in football Sweden
22:45 - 7:29 > Social intercourse (except for Mike); alcohol mandatory
7:29 - 7:30 > Free time (sleep not an option)

FEBS QUIZ

A challenging quiz to test if you have actually LEARNED anything in the past ten days !! Can you answer the following questions ???

- 1- Which of the following precipitants does NOT begin with "Na"? Is it NaNO_3 , NaCl , PEG6000 , NaSCN or NaTuRe ?
- 2- This may be a stupid question, but is it really ?
- 3- Can you think of any other questions that test if you have an IQ that exceeds that of an average Brussels sprout ?
- 5- Who spinked question 4 ?
- 4- What's wrong with the numbering of these questions ?
- 6- When did you last save your O-data ?
- 7- What went wrong with the scaling of the two columns of this newsletter ?
- 8- Given the following beer prices, calculate the mean, standard deviation and transform of drunkenness: 6 Kr, 15 Kr, 40 Kr, 34 Kr and 1,200,000 Kr. Hint: remove the four outliers before the data-processing.



2.38 minutes break now.

Alwyn Jones announces new Spink option in O

"Soon there will be a new Spink option in O", Alwyn Jones announced yesterday.

Denmark spinked by Rest of EC

Special reporter: Paula "censored" Fitzgerald

A nerve-wrecking demoralising humiliating suicide-inducing nauseating and thoroughly unpleasant friendly match between Denmark and the Rest of the European Community took place last Saturday.

After an even-matched first half, in which more goals were scored than in the entire European championship, the teams were level at five all. In the second half, the Danes' lack of fantasy, general inability to play football, a slightly larger goal than their opponents, not to mention the "Spink effect", allowed the EC team, in spite of their greater handicap (personified by someone who shall remain anonymous, because Charles "Complementary-hairdo" Carter asked us not to mention him), to secure a whopping 12-6 victory. In the mean time, I took several candid exposures of scarcely dressed, and oh-so sweaty men (life-size reprints and detail blow-ups available at a moderate fee).



FEBS Newslett.

Volume 1, Final Issue

Thursday, 18 June 1992 -- Price: 5 Beers

Editorial staff: Neil "Spink" Spink, Markku "I-like-semen" Saarinen, Paul "Big-balls" Hempstead & Gerard "Two-Cheeseburgers-small-fries-(no-ketchup-or-gammon-please)-and-a-large-vanilla-shake" Kleijwegt

MADNES RELEASE NOTES

by Peter "American-werewolf-in-Uppsala" Dunten

Jim Pflugrath demonstrated the latest improvements to Madnes at Århus, including the coffee command. Madnes will now drive your lab's coffeemaker, handling even the most complicated machines. Coffee can be invoked from any level of Madnes and is compatible with Mr. Coffee, Krupps, Melitta, and all standard coffee makers. Options which can be selected are milk, sugar, espresso, or extra black.

SOLUTIONS TO FEBS QUIZ

The editors have been flooded by requests for the one and only correct solution to the FEBS Ability Test in the previous final issue of this journal. So here it is:

- 1- The answer, of course, is salt.
- 2- Yes.
- 3- Ask Yakov.
- 5- Ask Yakov.
- 4- Ask Yakov.
- 6- Too late.
- 7- After an initial scaling problem, we discovered an error in our distribution function (the Fourier transform of which is naturally NOT its own complex Hermitian conjugate; we leave the proof of this to yourself as an exercise).
- 8- Six crowns, dummy!

Questionnaire Tips

(Free format): Title of course: Of course! Location: 56N, 10E, (Not in Europe any more), Duration/ Name of Participant/ Sex: (why didnt they just give one box?), Highest degree (of pleasure?), Nationality (i.e European or Danish), all other boxes may be left blank.

Unresolvable questions

Why should the use of low resolution data improve your chances of solving a structure (hint: because your highest resolution data is only four reflections at 2.3Å)?
What happened in the missing hour of Wednesday morning?
What is the structure of lysosyme?

Football results

Special reporter G. Bricogne has given us his unbiased maximum likelihood prediction of the results of the remaining matches of the European Football championships:

Semifinal 1: Sweden - Cyprus: 1-2
Semifinal 2: Denmark - Faeroerne: 0-0
Final: Zimbabwe - France: 6-12 (referee: Lennard Jones)

What have we learned ???

A summary of hints, insights, short cuts and bull that your editors have extracted from ten days of lectures. Together, these items form a "How to .."-guide to protein crystallography!

-how to obtain complete data-

- (0) Make sure that whatever protein you're working on, it's lysozyme.
- (1) Use INcomplete factorial designs to obtain whopping crystals within weeks.
- (2) Spend the next twelve years collecting data.
- (3) Custom-build a huge film cassette, but NEVER use it !!
- (4) Collect COMPLETE datasets. Use a sensible STRATEGY, for example: collect the entire hemisphere of data from 10000 to 0.9215272 Aengstroeuim (that's how it should be pronounced, by the way) at least four times.
- (5) Take good care when you scale your data: don't waste your time measuring crystal-to-film distances, set all your sigmas to zero (prevents a LOT of frustration later on).

- obtaining phases-

- (6) Solving the difference Patterson: use the PDB-structure of your protein (remember, it was lysozyme) to calculate phases; then do a difference Fourier and derive the position of your heavy atoms or ligands from the hand-out.
- (7) After you solved the structure (vide infra), use the final PDB-file to calculate beautiful rotation and translation functions with Merlot (pronounced: Merlot).
- (8) Phases CAN be refined, but if you think this is too difficult or just too much work, the worst that can happen is that your phases are 90 degrees out (that's only 25 % of the 360 degrees available).

-model-building and refinement-

- (9) Using "O", type the command `build_&_refine_lysozyme_structure`.
- (10) R-factors are still a problem for many people, but that's quite unnecessary, really. Recently, Spink, Spink & Spink (1992) proposed a new refinement protocol in which an ABS(Fobs - Fcalc) cut-off criterion is used to reject reflections which lie outside an acceptable range (doctors call these "outliers"). NOTE: Dodson has questioned the apparently large value of R_{free} and therefore Spink and Brunger suggested that the "bad" reflections could be used to calculate an unbiased R-factor, for which they suggest the name $R_{really,really\ bad}$.

-miscellaneous-

- (11) The bakery ("Peter Bager" just around the corner) sells hot breadrolls from 4:30 in the morning (butter extra).
- (12) At 5:30 kitchen personnel at the CPU will provide coffee free of charge.
- (13) Breakfast at 7:30 is a bad idea.
- (14) Denmark is a great country, provided you crave gammon and boiled potatoes every night.
- (15) MacDonald's is only ten minutes away from the UPC.
- (16) Something must be wrong with the slide projectors, because the other day I was seeing in stereo.
- (17) Beds in Denmark tend to be 1.60 meters long which is a trifle unfortunate if you happen to be a sweet, tall Dutchman.
- (18) Oh, by the way Ana, Neil asked if you could pick up your badge sometime. He found it in his (and Lluís's) room the other morning and almost stepped into it with his bare feet.
- (19) Four hours of lectures about Fourier theory is definitely NOT enough; eleven hours seems to be a reasonable minimum.

STOP PRESS: Danish X-rays strike French crystals.

STATUS REPORT FROM ESRF

Carl-Ivar Brändén

European Synchrotron Radiation Facility
Grenoble

The 844 m long 6.0 GeV electron storage ring is now built and is being commissioned. Before the summer shut-down, full intensity of the beam (100 mA) was achieved with a life-time of 5 hours. In July, an undulator was inserted into one of the straight sections to test the perturbation of the stored electron beam, and to analyze the X-ray radiation produced by the undulator. This X-ray beam was used for the first ESRF diffraction experiment from a diamond monochromator crystal, with very successful results. These diamond crystals are very promising as monochromators since they tolerate the severe heatload of the X-ray beams at ESRF, and in addition are transparent to non-diffracted photons which continue downstream and can be used for additional experiments.

We have recently started to build the first beamlines. There will be a total of 30 ESRF-funded beamlines, out of which 20 have already been assigned for specific applications. Five of these will be available for macromolecular crystallography.

A Beamline number 3 is a Laue beamline which will be shared between macromolecular crystallography and energy dispersive diffraction under high pressure. It is equipped with a 44 pole wiggler and a toroidal mirror for 1:1 focusing. In addition to conventional Laue crystallography, this beamline can also be used for single bunch experiments. Calculations show that in the 7.5 mA single bunch mode of the storage ring, a small 0.2 * 0.2 mm sample receives $3.2 * 10^{10}$ photons from the focused beam in a single bunch of 50 psec duration. This should be sufficient to record a reasonable

fraction of the reflections from well-diffracting crystals of small proteins, according to simulations made by the beamline responsible scientist, Michael Wulff.

B Beamline 4, the high flux Beamline, has been designed by P. Bösecke to serve two different user groups:

- (1) Macromolecular crystallography on weakly diffracting specimens, with a reasonably large sample cross-section (typically $0.1 * 0.3$ mm);
- (2) Time resolved small-angle X-ray scattering experiments.

The beam, which is produced by an undulator, will be monochromatized by a Si(111) double crystal monochromator and focused by a toroidal mirror to a size of $0.4 * 0.65$ mm at the sample in the crystallography hutch. The total flux of the beam at the sample position is calculated to be $8 * 10^{12}$ photons/sec for an energy bandwidth $\Delta E/E$ of $2 * 10^{-4}$ at 12 keV. Within the framework of a collaborative agreement between EMBL and ESRF, a newly recruited staff member of the EMBL Outstation in Grenoble, Bjarne Rasmusson, will be responsible for the macromolecular crystallography endstation of Beamline 4.

C A microfocus beamline, Beamline 1, has been designed by C. Riekell to cover several scientific areas, one of which is diffraction on single crystals down to the sub- μm range. The initial aim is to reach a focal spot of $10 * 10 \mu\text{m}$ of a beam from an undulator similar to that of Beamline 4.

These three beamlines all belong to the first set of 8 ESRF-funded beamlines which are planned to be ready for routine operation by outside users during 1994. The next set of beamlines will be ready one year later. There will be two beamlines among this set which can be fully devoted to macromolecular crystallography.

D Beamline 19 is now being designed by Andy Thompson, a staff member of the EMBL Outstation, for Multiple Anomalous Dispersion (MAD) experiments. This beamline is one of the few ESRF-funded beamlines on a bending magnet.

E Beamline 20 will be devoted to "state of the art" macromolecular crystal data collection. Proper design of this beamline has not yet started, but the preliminary plans include an undulator beam and a series of diamond crystal monochromators. Each diamond crystal will produce a monochromatic beam around 0.9 Å to a separate data collection station. The end station will be equipped with a Si-monochromator for a high energy beam around 0.3 Å. Since all the data collection stations on this beamline can operate simultaneously, the data collection capacity of the beamline will be very high.

All of these beamlines will be operated from UNIX workstations using a unified software system, such that once a user knows the system he should in principle be able to operate any beamline. Image plate detectors will be extensively used until other detector systems more suitable for 3rd generation synchrotron sources have been developed.

o0o

THE CRYSTALLOGRAPHY WORKSTATION FOR X-RAY STRUCTURAL INVESTIGATIONS OF SINGLE CRYSTALS AT THE TNK STORAGE RING

A.N.Popov, D.M.Kheiker, E.H.Harutyunyan

*Institute of Crystallography of the USSR Academy of Sciences
Moscow, USSR*

1. INTRODUCTION.

At the Institute of Crystallography, Moscow, a new instrument has been constructed for structure investigations of single crystals at the SR source. The station is intended for work at the TNK storage ring in the town of Zelenograd. Though a wide range of structure investigations is supposed to be carried out, the investigation of protein crystals structure is considered as the main task.

2. THE TNK STORAGE RING.

The TNK storage ring in Zelenograd (about 30 km from Moscow, developed at the Nuclear Physics Institute of Novosibirsk) is a dedicated SR source intended to solve problems encountered in sub-micron technology as well as to carry out various investigations in the wavelength range 0.2-2000 Å [1,2]. The source comprises two undulators for soft X-ray and vacuum ultraviolet radiation, five multipolar wigglers for soft X-ray lithography and two superconducting wigglers to work with hard radiation. The crystallography station will use radiation from superconducting wiggler. The main parameters of the TNK and the wiggler are presented in tables 1 and 2. The wiggler is constructed so as to obtain single radiation source image in the sample plane without additional collimation. The beam

from the wiggler (3° angle) is uniformly distributed among five users channels, with 3.6 mrad radiation for each channel. The distance from the source to the instrument monochromator is 15 m.

Table 1.

Parameters of the dedicated SR source TNK

Energy	- 1.2; 1.6; 2.0 GeV
Emittance	- 27 nm-rad
Stored current: total	- 300 mA
single bunch	- 100 mA
Lifetime	- 10 h

Table 2.

Parameters of the superconducting wigglers

Ondulator parameter	- 230		
Number of poles	- 5		
Length	- 130 cm		
Maximum field	- 80 kGs		
Interpole gap	- 3 cm		
Energy	1.2 GeV	1.6 GeV	2.0 GeV
λ_c (Å)	1.6	0.9	0.6
σ_x (mm)	0.58	0.43	0.54
σ_z (mm)	0.02	0.01	0.02
σ_x' (mrad)	0.09	0.07	0.09
σ_z' (mrad)	0.03	0.02	0.03
Power (Wt/mrad)	18	43	83

3. WORKSTATION COMPONENTS.

The radiation is monochromatized and focused by a horizontally dispersing triangular oblique cut monochromator prepared from perfect Ge and Si crystals. In front of the monochromator the slits are placed that restrict the white SR beam. Focusing in vertical

direction as well as cutting of high harmonics is provided by a segmented mirror that is positioned behind the monochromator. The monochromator and the mirror are followed by the slits to eliminate the background and for collimation. To carry out experiments with the crystals under study a special four-circle diffractometer was constructed with a horizontal axis of scanning ω , point scintillation detector and area detector. In front of a changeable collimator of the diffractometer there is a high-precision ionization chamber to monitor the beam. All units of the workstation following the monochromator are placed on the bench that can be turned around the axis, which coincides with that of the monochromator, within 11° - 45° range relative to the white SR beam. A special high-precision carriage is used to adjust the diffractometer. The possibility of fine adjusting the distance from the monochromator to the investigated crystal within the range 3-4 m is also envisaged.

4. OPTICAL SYSTEM.

The monochromator design is similar to that of an automatic goniometer head that allows to make adjusting rotations and displacements, bend the triangular crystal-monochromator as a round cylinder and to turn it around the vertical axis by $0.4''$ steps. Using of a set of crystals Ge (111), Si (111), Si (220) with variously cut angles (from 4° to 14° in 2°) and the possibility to change the distance to the sample under study provides Guinier focusing with the resolution ($\delta\lambda/\lambda$) reaching $(4-2) \cdot 10^{-4}$ in a wide wavelength range 0.4-2.5 Å, the intensity of the beam being high. The calculated horizontal size of focal spot is near 600, 350 and 250 μm for Ge (111), Si (111) and Ge (220) correspondingly. Using of Ge instead of Si provides more than two times gain in total intensity, the flux density increasing by only 10-20 % . The mono-

chromator is placed within cylindrical vacuum chamber that can be filled with helium to prevent heating effects.

The vertical size of the beam is about 4 mm at a distance 16 m from the source. The mirror comprises eight planar segments made of polished glass, each measuring $200 \times 50 \times 25$ mm³. The segments are positioned to form a surface imitating an elliptical cylinder. Tuning the mirror to the grazing angle of 2.4 mrad ($\lambda_{\min} \approx 0.9$ Å) leads to vertical focusing of the whole beam to 0.5 mm. Such a geometry will allow to work in the wavelength range 2.5-0.95 Å without readjusting the mirror when changing λ . To focus more hard radiation one will need plates with heavy metal coating. The evaluated density of the radiation flux for 0.15 A current is 10^{12} - 10^{13} photons/(s mm²) in the focus point.

5. DIFFRACTOMETER.

All four diffractometer axes are supplied with absolute angledetectors that allow to measure angles to 0.005° precision. Discreteness of rotation around all axes is not worse than 0.002°.

As an area detector a multiwire proportional chamber is used with a drift gap 100 mm long, constructed at the Joined Institute for Nuclear Research, Dubna. The detector is filled with Ar, Xe and methane gas mixture under the pressure 4 atm. Passive getter inside the detector provides long-lasting gas purity without refilling. The entrance window of Be 1 mm thick imitates part of the spherical surface of 120 mm radius and has the diameter 130 mm. Correcting drift electrodes practically eliminate parallax. The data is collected by fast delay-lines, counting rate being 10^6 Hz at 20 % losses. The number of pixels is 500x500, the pixel size - 0.25 mm, the detector resolution - 0.3 mm. Differential non-linearity does not exceed 5 % . The detector allows to work in a wide wavelength range

from 0.8 Å to 1.8 Å with detecting efficiency not worse than 60 %

The area detector is fixed on a special carriage that allows both its rotation by 45° around horizontal axis coinciding with ω axis of the diffractometer and changing the distance from the crystal under study to the detector from 100 mm to 800 mm.

6. CONTROL SYSTEM.

All the controlling and detecting units including the electronics of the area detector are manufactured according to CAMAC standards. As a controlling computer a PC AT is used.

REFERENCES

1. A.G.Valentinov et al. - Preprint INP 90-129, Novosibirsk, 1990.
2. G.N.Kulipanov - Synchrotron Radiation News, 1991, Vol.4, N°1, P.17-20.

Availability of the MOSFLM program suite for processing image plate and film data

Andrew G. W. Leslie, MRC laboratory of Molecular Biology, Cambridge, CB2 2QH

E-mail: ANDREW@UK.AC.CAM.MRC-LMB

The MOSFLM program suite (version 4.30) can be retrieved via anonymous ftp over Internet from the Alliant at LMB. The Unix version is stored as a compressed tar file, the VAX version as a backup saveset. The procedure for copying the files is as follows:

- 1) ftp 131.111.84.16
- 2) logon as user "ftp" password "ftp"
- 3) move to the correct directory:
 Under Unix: cd pub
 Under VMS: set def "pub"

Unix Version.

- 4) Set correct mode so EITHER: type binary
 OR: binary
(this is machine dependent)
- 5) Pull the compressed tar file across: get mosflm.tar.Z

Vax Version (from a VAX host)

- 4) Set correct mode: set type ascii
- 5) Pull over the backup file: get/fdl mosflm.bck

The get/fdl facility is DEC specific, and actually copies two files. The first is an fdl file (ascii) which specifies the attributes for the backup saveset file which is then copied over automatically (the mode is automatically changed to binary for the backup file). The fdl file is actually called MOSFLM.BCKfdl. If you just try to "get" the MOSFLM.BCK file then the file will not be recognised as a backup saveset. The "get/fdl" facility will only work if ftp on your own machine is supported by the DEC software UCX. The simplest way to test this is simply to type "UCX" and you should get a prompt if it is installed on your machine. There is some third party software to run ftp on DEC machines which does not support the get/fdl facility, in which case you need to speak to your computer manager (It is apparently possible to get around this problem but I do not know how !).

There is additional documentation on the programs in 2 files. The file "mosflm_notes" gives an overview of the operation of the programs, while the "mosflm_user_guide" gives specific information on running the programs. I would like to thank all those who sent comments on the draft versions of this documentation.

A new version of the program (Version 5.0) is currently undergoing tests at LMB. The major changes are to the profile fitting algorithms. The areas over which the standard profiles are determined are now under user control, and improved algorithms to determine the standard profiles and derive the profile fitted intensity have been implemented. In addition, the optimum background mask parameters (defining the X and Y rims and corner cutoff) are determined automatically by the program. This version will only be released following tests at a limited number of sites, but anyone wishing to test the program should contact me.

ON THE RELIABILITY OF ATOMIC MODELS IN THE PDB : STEREOCHEMICAL POINT OF VIEW

A.G. Urzhumtsev

*Institute of Mathematical Problems of Biology, Russian Acad.Sci.,
Pushchino, Moscow region, 142292, Russia
(current address: 15, rue Descartes, UPR de Biologie Structurale,
IBMC, 67000 Strasbourg, France)*

Abstract.

The reliability of a spatial structure is determined by the reliability of all its components. But usually only average characteristics are reported in the literature and they may "hide" large distortions of stereochemical parameters. An analysis of some files of the Protein Data Bank has confirmed this hypothesis.

The atomic model of protein is a complex system whose reliability depends on the reliability of all its components. In particular, the presence of only one strongly distorted bond length in the main chain is enough to raise doubts in reliability of the whole structure because usually local improvement of a model is hardly possible. This, in turn, makes doubtful the speculations concerning the mechanism of protein activity.

Unfortunately, the majority of published papers on protein structure determination only contain average information on stereochemical distortion, which is absolutely insufficient. For example, for a protein of 100 amino acid residues, the average error in bond lengths of about 0.01Å may be a result of errors in two polar situations:

- a) all bond lengths contain errors of about 0.01Å;
- b) all bond lengths are absolutely correct, except the one which has an error of about 3Å.

It is clear that the former type of models is quite good, which cannot be said of the latter one.

We have looked through a number of Protein Data Bank files (Bernstein et al., 1977) to analyse maximal distortions in stereochemical values. Of course, the quality of a model depends on many factors, of which the most important are the resolution, the refinement program used, the protein size, the date of structure determination, etc. Note that the proper geometry may be shown both by a well-refined structure and by a non-refined structure which is in bad agreement with experimental X-ray

data. Because of a very complex interrelation of these factors the models were analysed without classifying them into groups, we only separated refined structures from the unrefined ones.

The first 99 files were picked up from the available PDB version. Upon removal files with RNA structures, with incomplete models, with unrefined and partially refined models, etc, the errors $\langle \Delta d \rangle$ of bond lengths were analysed for the rest of 50 files. Other stereochemical values were not analysed so thoroughly although it was noted, for example, that errors in bond angles were distributed roughly as in bond lengths.

Other 4 models were excluded intentionally from the analysis: two of them had very large average errors, $\langle \Delta d \rangle$, of about 0.16 \AA , and for the other two the large maximal errors, Δd_{max} , exceeded 1 \AA .

An analysis has confirmed that the average errors $\langle \Delta d \rangle$ are small enough both in refined and in majority of unrefined models (Fig.1). However, 18 refined models, almost 40% of the total, have main chain errors Δd_{max} larger than 0.1 \AA and 6 of them have errors larger than 0.2 \AA (Fig.2). When side chain errors are taken into account, the number of such models increased up to 30 and 9, respectively. Of course, it should be noted that some models including several unrefined ones have perfect characteristics, both average (Fig.1) and maximal (Figs.2,3).

The analysis which has been carried out shows that special efforts should be undertaken during refinement to avoid large stereochemical distortions. For example, the refinement complex FROG (Urzhumtsev et al., 1989) provides a user with a histogram of distribution of different stereochemical characteristics but not with only average values.

Thus, the selective analysis of PDB files has revealed some PDB models with unrealistic stereochemical parameters. That is why one should be careful when dealing with macromolecular models. As a preventive step it will be very useful to ask the authors to publish not only the average but also the maximal distortions in macromolecular atomic models described.

Addition in proof : when the paper has been prepared (September 1990) for publication the author was informed that the similar work has been done by Dr. M.Basharov, Institute of Biological Physics, Pushchino.

References :

- Bernstein, F.C., et al. (1977) *J.Molec.Biol.*, **112**, 535-542
- Urzhumtsev, A.G., Lunin, V.Yu. & Vernoslova, E.A. (1989) *J.Appl.Cryst.*, **22**, 500-506

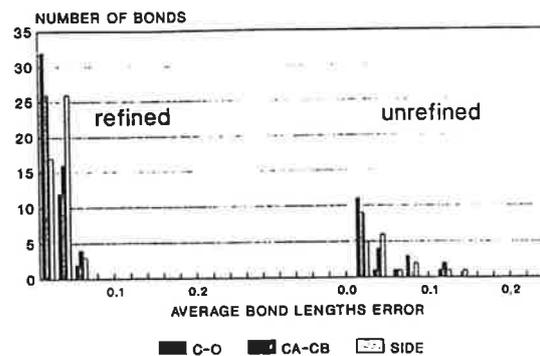
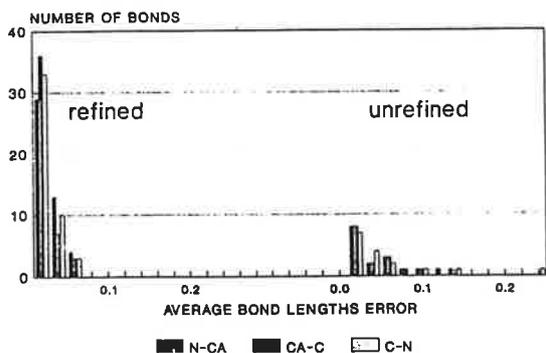


Fig. 1. The histogram of distribution of average bond-length errors, $\langle \Delta d \rangle$, in refined and unrefined models :
 a) main chain bonds, b) other bonds.

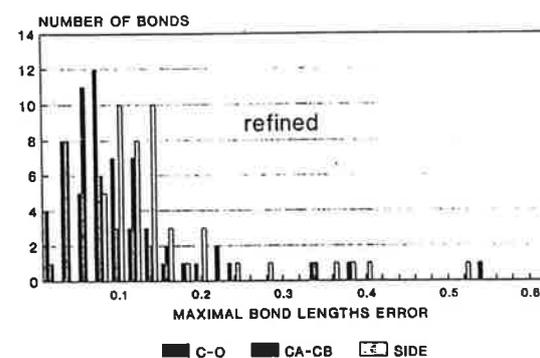
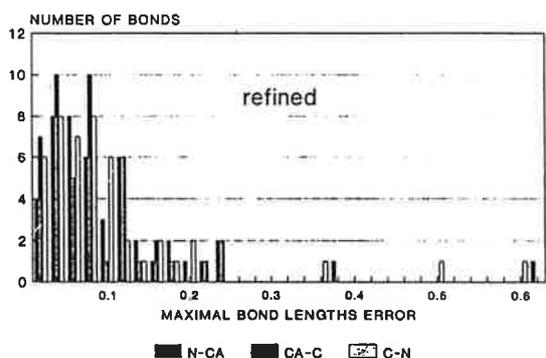


Fig. 2. The histogram of distribution of maximal bond-length errors, Δd_{\max} , in refined models :
 a) main chain bonds, b) other bonds.

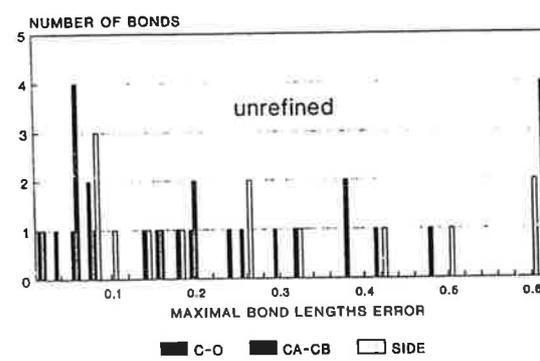
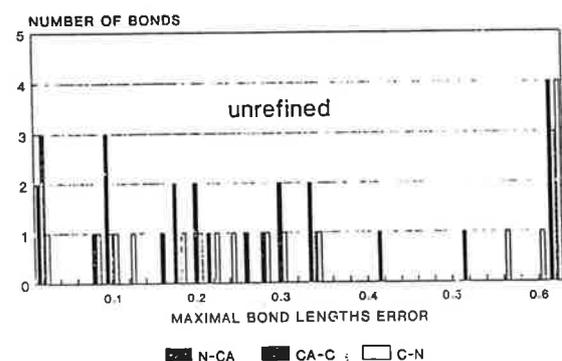


Fig. 3. The histogram of distribution of maximal bond-length errors, Δd_{\max} , in unrefined models :
 a) main chain bonds, b) other bonds.

PROCHECK: A program to check

the stereochemical quality of protein structures

Roman A Laskowski(1), Malcolm W MacArthur(1,2),
David S Moss(2), & Janet M Thornton(1)

(1) Biomolecular Structure and Modelling Unit, Department of
Biochemistry and Molecular Biology, University College, Gower Street,
London WC1E 6BT, England.

(2) Crystallography Department, Birkbeck College, Malet Street,
London WC1E 7HX, England.

We have developed a suite of programs called PROCHECK which provides a detailed check on the stereochemistry of protein structures. It is of use not only for checking protein structures in the process of being solved, but also for checking existing structures and those being model-built.

The input to PROCHECK is a coordinate file in the standard Brookhaven format. The checks applied include the various stereochemical parameters of Morris et al. (1992), together with an up-to-date set of standard bond lengths and bond angles (Engh & Huber, 1991).

The outputs produced by PROCHECK comprise a number of plots and a detailed residue-by-residue listing. The plots clearly show how well the structure compares with well-refined structures of the same resolution, and how good the stereochemistry is on a residue-by-residue basis. The plots are output in PostScript format, so can be printed off on a PostScript laser printer or displayed on a graphics screen if you have the appropriate software.

The source code and operating instructions are available from the authors, free of charge to academic institutions. The programs are supplied with script files for running on unix operating systems, and command files for running under VAX VMS. Requests for the programs can be made by post, or by e-mail to roman@uk.ac.ucl.bioc.bsm.

References

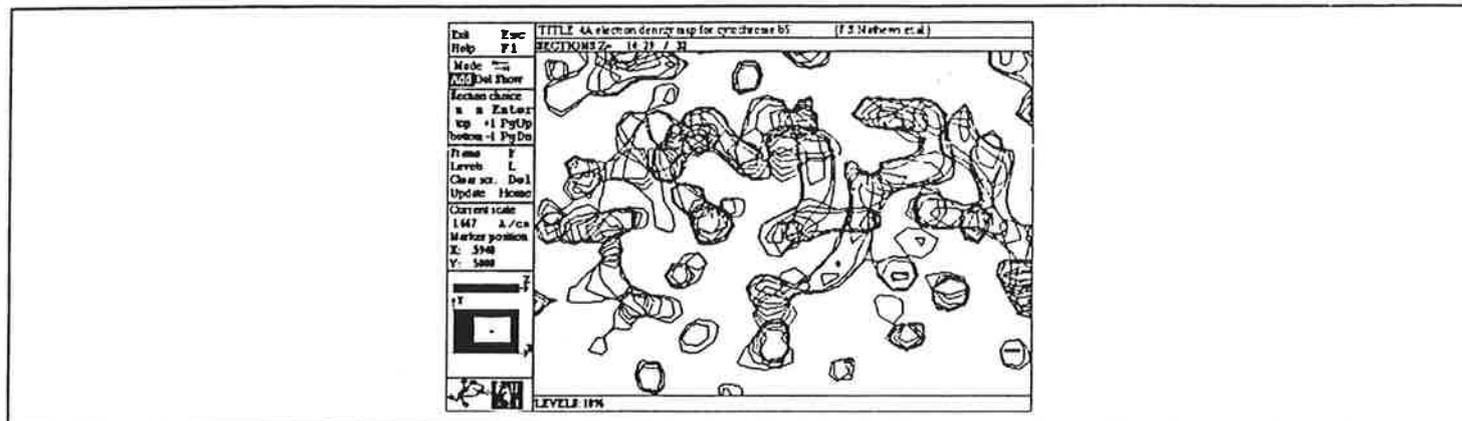
Engh R A & Huber R (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst.*, A47, 392-400.

Morris A L, MacArthur M W, Hutchinson E G & Thornton J M (1992). Stereochemical quality of protein structure coordinates. *Proteins*, 12, 345-364.

FAN

Scalar Field Analysis

FAN is a utility program especially useful in crystallography research. FAN was developed at the USSR Academy of Sciences especially for studying lattice functions which depend on three variables such as scalar fields and three dimensional distributions. FAN is the first microcomputer program of its type to let you display functions by isolines in two-dimensional sections of real physical space. Using FAN you can interactively explore the structure of your function by changing perspective or scale. FAN also lets you map functions in space by periodicity and display functions defined in oblique coordinate systems.



FAN works by letting users graphically explore functions. In the FAN graphical interface you can

- ⇒ change the set of displayed isoline sections,
- ⇒ change the scale values and viewing frame position,
- ⇒ control the colors of the displayed isoline sections,
- ⇒ view function sections in any orientation,
- ⇒ remove, replace or add single isoline sections,
- ⇒ magnify or reduce of function images.

FAN has built-in algorithms to process very large data sets and it can convert ASCII and binary data files to its own compact file structure.

FAN was written by Drs. E.Vernoslova and V. Lunin, of the Research Computing Center, Academy of Sciences of the USSR.

IBM PC or compatible;
DOS 3.3+; 512K RAM;
Hercules/ /EGA/VGA;
Math coprocessor.

Cost: \$99
Site license: \$195



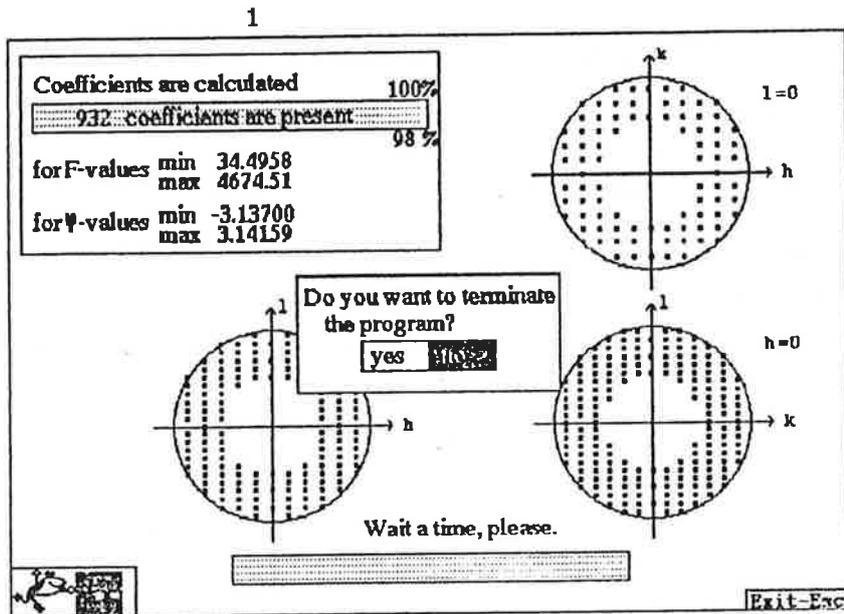
Edited by
Jeffrey A. Millstein,
Applied Biomathematics

Available from: Exeter Software, 100 North Country Road, Setauket, New York 11733, USA.
Phone: (800)-842-5892 or (516)-751-4350. FAX: (516)-751-3435.

FFT

Discrete Three-Dimensional Fourier Transform

FFT (Fast Fourier Transform) is a convenient tool for calculating discrete three-dimensional Fourier transform in crystallography. FFT enables you to calculate Fourier and Patterson syntheses, difference Fourier and difference Patterson syntheses in a several formats. Program has a user-friendly interactive regime for input of all necessary parameters. FFT can calculate synthesis in any space crystallographic group, in any crystallographic axes orientation. The missing of structure factors is under control when synthesis is calculated.



Requirements: IBM PC or PS/2 compatible, DOS 3+, 512K RAM, and EGA, VGA or Hercules monochrome graphics monitor.

FFT was written by E.A. Vernoslova of the Pushchino Computing Center of the Academy of Sciences, USSR, and is one of the FROG-series of crystallographic IBM PC compatible programs.

C A N

Comparative Analysis of Three-Dimensional Functions

CAN is a microcomputer program especially useful in crystallography research, designed for visual investigation independently or at a time images of three objects:

- two lattice scalar function depending on three space variables;
- a protein molecule's atomic model.

The analyzing function is performed in the form of the set of function values calculated at the points of a grid in three dimensional space. The atomic model is determined by a set of atoms coordinates and displayed as a skeleton model's projection into the screen, where atoms are presented by globules and chemical bonds - by segments.

CAN makes it possible:

- to display at the screen the picture of the specified isolines in the one chosen section or in a number of sections;
- easy change the set of displayed sections (add/delete/replace);
- to change the scale of the picture (to "magnify" / "diminish" desired fragment) and move the frame into the desired region;
- easy vary the set and colours of displayed isolines (to change isoline levels for any values);
- to determine relative coordinates of any point of the screen;
- to depict a function by sections along any of coordinate axes;
- to display the skeleton model at the screen;
- to move and to rotate the model's image as a rigid body with respect to the analyzed main function.

Requirements: IBM PC or PS/2 compatible, DOS 3+, 512K RAM, EGA or VGA graphics monitor.

CAN was written by E.Vernoslova and V.Lunin, Research Computing Center, Academy of Sciences, Pushchino, Moscow Region, RUSSIA, and is one in FROG-series of crystallographic IBM PC compatible programs.

3rd European Workshop on crystallography of Biological Macromolecules

Como, May 24th - 28th, 1993

Under the auspices of the European Science Foundation Network on Crystallography of Biological Macromolecules

ORGANIZERS

T. Alwyn Jones (Upssala - Sweden)

Paola Spadon (Padova - Italy)

Giuseppe Zanotti (Padova - Italy)

PURPOSE AND ORGANIZATION OF THE WORKSHOP

The workshop is the third of a series of meetings held in Como every two years. It will be open to about 150 participants and will focus on the application of X-ray crystallographic methods to the study of proteins and nucleic acids, with emphasis on the latest developments. Experimental and computing methods, but also new interesting structures will be discussed. The participants of the workshop are encouraged to propose titles for a contribution.

LOCATION

The workshop will be held at the Centro di Cultura Scientifica "A. Volta", Villa Olmo, lake Como, Italy. Villa Olmo is located near the Swiss-Italian border, at 10 minutes walking distance from the centre of Como. Como can be reached comfortably by train (30-40 min. from Milano; 3-4 hours from Zürich and Basel) or by car via the A9-E36 motorway.

APPLICATION

Applications, including a short c.v., proposed titles for scientific contributions and any additional relevant information should be sent before January 10th, 1993 to:

Prof. Giuseppe Zanotti

Dip. Chimica Organica

Via Marzolo 1

35131 Padova, ITALY

tel. (49)831229 or 831327 - fax: (49)831222

E-mail: COMO@PDCHOR.UNIPD.IT

Further information will be provided to the applicants.

ORGANIZERS ADDRESSES:

Prof. T. Alwyn Jones, Dept. of Molecular Biology, Biomedicum Centre, Box 59, S-75124 Upssala, Sweden

Prof. Paola Spadon, Dipartimento Chimica Organica, Università di Padova, Via Marzolo 1, 35131 Padova, Italy

Prof. Giuseppe Zanotti, Dipartimento Chimica Organica, Università di Padova, Via Marzolo 1, 35131 Padova, Italy

CCP4 OVERVIEW

in relation to programs for protein crystallography

STAGE 1: data acquisition

OSCILLATION DATA

film or image plate

AREA DETECTOR DATA

FAST or Xentronics

LAUE DATA

DIFFRACTOMETER

STAGE2: data processing

Initial corrections e.g., L.P. applied.

Output file contains a single record for each observation .

Use programs such as:

ABSCALE (MOSFLM o/p)

ABSURD (FAST o/p)

ROTAPREP (multivarious o/ps)

to prepare an MTZ multi-record file.

SORTMTZ

ROTAVATA

AGROVATA

TRUNCATE

Produce a scaled, merged MTZ file containing h k l, F, sigF

STAGE 3: decision

Either:

Data merging preparatory for
MIR phasing, mutant inspection etc.

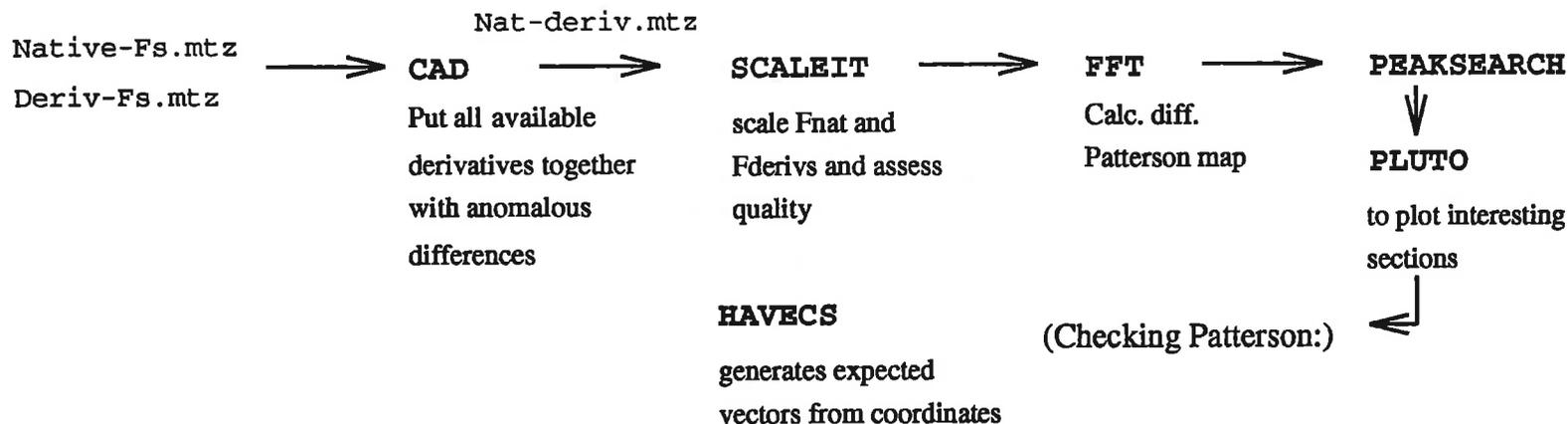
Or:

Matching amplitudes to a model
e.g., molecular replacement
refinement

Multiple Isomorphous Replacement (MIR) Structure Solution

1) Find heavy atom sites

a) Solving Patterson for first derivative to give major site(s):



40

b) Use Direct methods to find site(s):

SHELX
MULTAN

2) a) Heavy atom refinement

VECREP, HEAVY-MTZ

b) Heavy atom refinement and phasing

MLPHARE

Several options for heavy atom refinement and phasing

Maximum Likelihood PHasing
and REfinement

3) Difference Fouriers to find new heavy atom sites in first and other derivatives

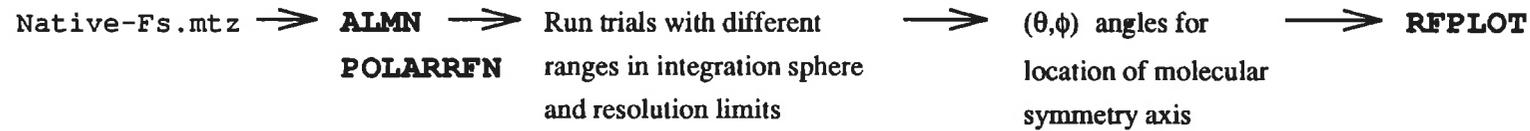


4) Back to calculate and check Pattersons for other derivatives

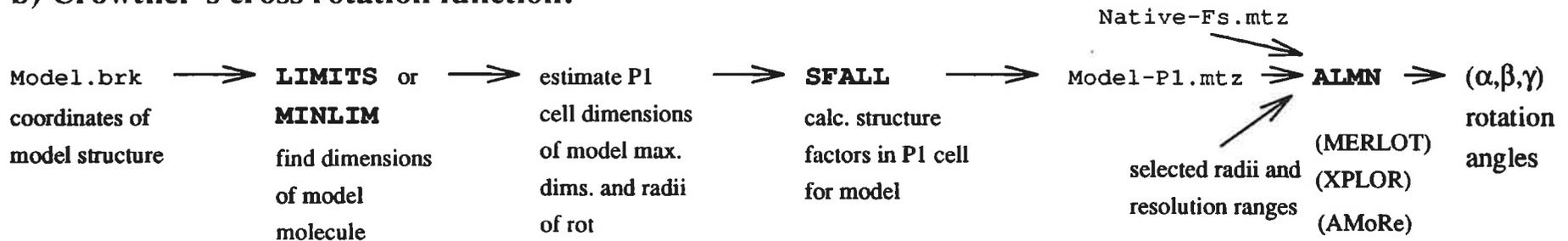
[It is advisable to do the same procedure starting with a different derivative to check for errors or bias; do try to start on the same origin.]

MOLECULAR REPLACEMENT

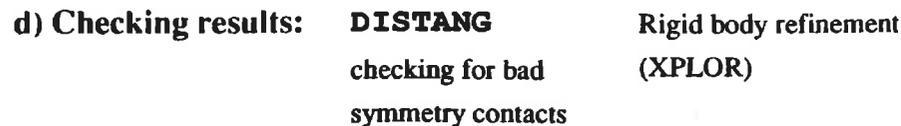
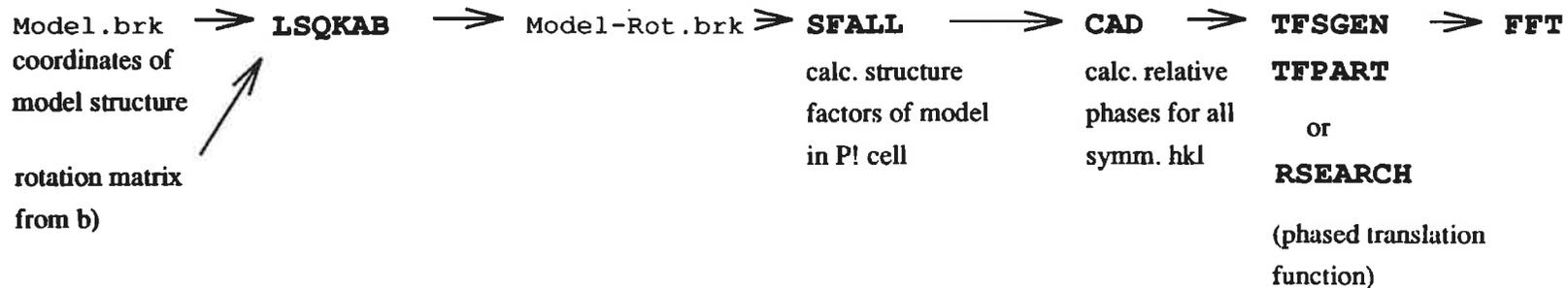
a) Detection of non-crystallographic symmetry using Crowther's FFT rotation function



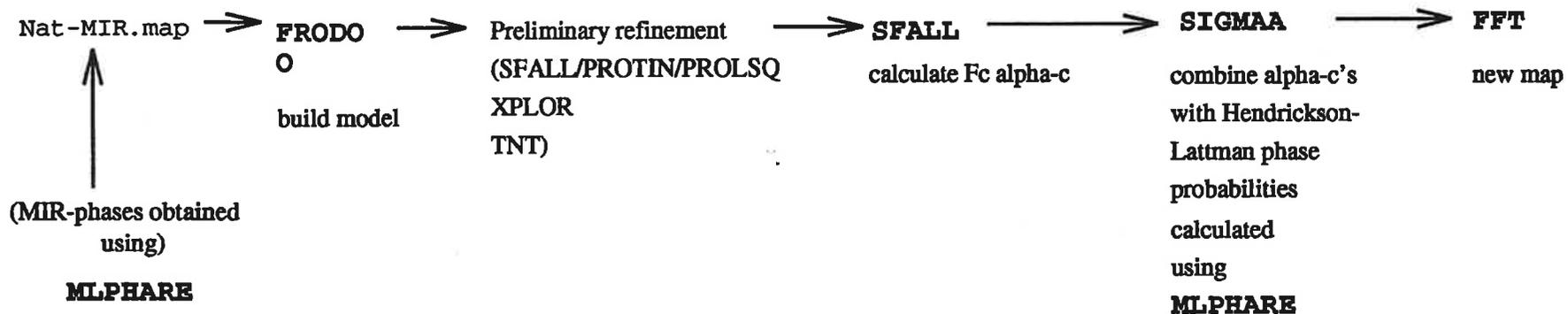
b) Crowther's cross rotation function:



c) Translation function:*

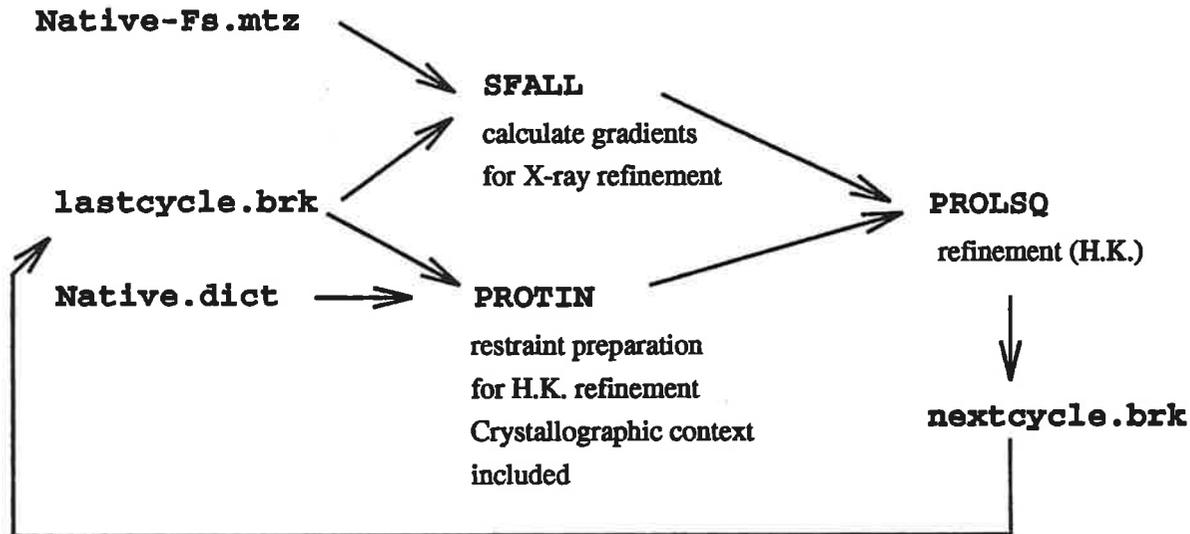


COMBINATION OF PARTIAL STRUCTURE AND ISOMORPHOUS PHASE INFORMATION



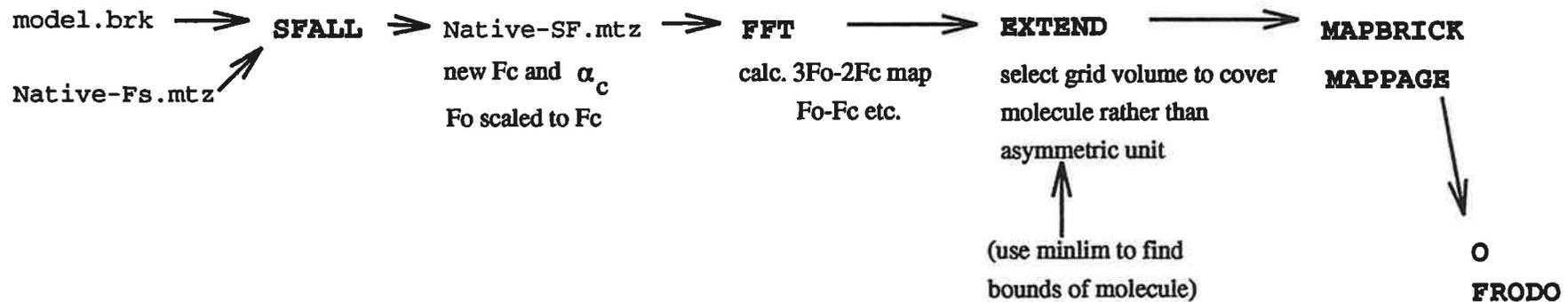
REFINEMENT AND DIFFERENCE ELECTRON DENSITY MAPS

a) Cycles of restrained least-squares structure factor refinement (Hendrickson-Konnert)



43

b) Difference map generation



Main, supported CCP4 programs

This is the list of main supported CCP4 programs. For each program the name is given of the person to contact in the case when expert advice is required, this person is not necessarily the program author.

1. **ABSCALE**- film and image plate scaling.
Andrew Leslie, MRC, Cambridge.
2. **ABSURD**- initial processing of FAST intensity files from MADNES.
Phil Evans, MRC, Cambridge.
3. **AGROVATA, ROTAVATA**- scaling data in batches.
Phil Evans, MRC, Cambridge.
4. **ALMN**- Crowther's rotation function using FFT.
Eleanor Dodson, York University.
5. **ASC_2_P84, P84_2_ASC**- converts ASCII file to PLOT84 type meta file and reverse.
Peter Daly, ESRF, Grenoble.
6. **BAVERAGE** averages temperature factors over main and side chain atoms.
Eleanor Dodson, York University.
7. **CAD**- collecting assorted reflection data.
Eleanor Dodson, York University.
8. **COMPLETE**- produces statistics on completeness of the data set.
Andrew Leslie, MRC, Cambridge.
9. **CONTACT**- calculates various contacts in protein structure.
Peter Brick, Imperial College, London.
10. **COORD_CONV**- interconverts various coordinates formats.
Eleanor Dodson, York University.
11. **DISTANG**- calculates distances and angles in protein molecule.
Peter Brick, Imperial College, London.
12. **DUMPRESO**- more elaborate than MTZDUMP display of contents of the MTZ file.
Sandra McLaughlin, EMBL Outstation, Hamburg.
13. **ECALC**- calculates normalised structure amplitudes.
Ian Tickle, Birkbeck College, London.

14. **ENVELOPE**- generates a molecular envelope to use in solvent flattening.
Andrew Leslie, MRC, Cambridge.
15. **EXTEND**- extends the map to cover the specified volume of the unit cell
Phil Evans, MRC, Cambridge.
16. **F2MTZ**- converts a formatted reflection file to MTZ file.
Eleanor Dodson, York University.
17. **FFT**- crystallographic fast Fourier transformation.
Eleanor Dodson, York University & Phil Evans, MRC, Cambridge.
18. **FTTBIG**- crystallographic Fourier transformation in real memory.
Eleanor Dodson, York University & Phil Evans, MRC, Cambridge.
19. **FHSCAL**- scaling derivative to native data by Kraut's method.
Ian Tickle, Birkbeck College, London.
20. **FLATMAP**- use molecular envelope to flatten solvent region of electron density map.
A. Leslie, MRC, Cambridge.
21. **HEAVY**- heavy atom refinement and phasing program.
Ian Tickle, Birkbeck College, London.
22. **HKL PLOT**- plots a section of reciprocal space using reflection file.
Eleanor Dodson, York University.
23. **ICOEFL**- scaling of multiple Fc (or Ic) to Fobs. Phil Evans, MRC, Cambridge.
24. **LCF2MTZ**- converts an LCF reflection file to MTZ one.
Sandra McLaughlin, EMBL Outstation, Hamburg.
25. **LSQKAB**- optimises fit of atomic coordinates from two data sets.
Eleanor Dodson, York University.
26. **MAPEXCHANGE**- converts binary map file into ASCII one and reverse.
Phil Evans, LMB, Cambridge.
27. **MAPSIG**- print statistics on signal/noise for translation function map.
Ian Tickle, Birkbeck College, London.
28. **MAPTONA4**- converts binary map file to na4 format and reverse.
Ian Tickle, Birkbeck College, London.

29. **MLPHARE**- phase calculation and refinement.
Eleanor Dodson, York University.
30. **MTZ2VARIOUS**- makes ASCII reflection files from MTZ file.
Eleanor Dodson, York University.
31. **MTZDUMP**- displays the contents of the MTZ file.
Sandra McLaughlin, EMBL Outstation, Hamburg.
32. **MTZTONA4, NA4TOMTZ**- transforms MTZ file to transferable
ASCII file and reverse.
Ian Tickle, Birkbeck College, London.
33. **MTZUTILS**- reflection data file utility program.
Kim Henrick, MRC, Cambridge.
34. **OVERLAPMAP**- calculates the correlation coefficient or overlap
between maps.
Eleanor Dodson, York University.
35. **PDBSET**- various useful manipulations on coordinate files.
Phil Evans, MRC, Cambridge.
36. **PEAKMAX**- search for peak in the electron density map.
Phil Evans, LMB, Cambridge.
37. **PLUTO**- plots molecules end electron density maps.
Kim Henrick, LMB, Cambridge.
38. **POLARRFN**- fast rotation function in polar coordinates.
Phil Evans, LMB, Cambridge.
39. **PROLSQ**- structure refinement.
Eleanor Dodson, York University.
40. **PROTIN**- prepares file with restrains for PROLSQ.
Eleanor Dodson, York University.
41. **PRMAP**- prints sections of electron density maps.
Phil Evans, LMB, Cambridge.
42. **REINDEX**- reindex MTZ reflections file.
Eleanor Dodson, York University.
43. **ROTAPREP**- produces multirecord MTZ file.
Eleanor Dodson, York University.
44. **ROTMAT**- interconverts CCP4/MERLOT/XPLOR rotation an-
gles.
Eleanor Dodson, York University.

45. **RSEARCH**- rfactor search. Eleanor Dodson, York University.
46. **RSTATS**- scales two data sets.
Peter Brick, Imperial College, London.
47. **SCALEIT**- various (including anisotropic) derivative and native scaling.
Eleanor Dodson, York University.
48. **SFALL**- structure factor calculations using FFT.
Eleanor Dodson, York University.
49. **SIGMAA**- phase combination.
Eleanor Dodson, York University.
50. **SORTMTZ**- sorting MTZ file.
Peter Daly, ESRF, Grenoble.
51. **SURFACE**- molecular surface calculations.
Kim Henrick, MRC, Cambridge.
52. **TRACER**- reduced cell calculations.
Kim Henrick, MRC, Cambridge.
53. **TFFC**- translation function.
Ian Tickle, Birkbeck College, London.
54. **TRUNCATE**- truncates reflection data.
Phil Evans, MRC, Cambridge.
55. **UNIQUE**- generates unique reflection data set.
Andrew Leslie, LMB, Cambridge.
56. **VECREF**- vector space refinement of the heavy atoms.
Ian Tickle, Birkbeck College, London.
57. **VECTORS**- generates Patterson vectors from atomic coordinates.
Phil Evans, LMB, Cambridge.
58. **WATERSORT**- assign waters to nearest protein atoms.
Phil Evans, LMB, Cambridge.
59. **WATERTIDY**- rearranges water position.
Eleanor Dodson, York University.
60. **WILSON**- makes Wilson plot.
Eleanor Dodson, York University.

Miscellaneous and unsupported CCP4 programs

1. **ANGLES**- bonds and dihedral angles from Diamond and PDBfiles.
2. **AREA**- solvent extended accessible area.
3. **AREAIMOL**- solvent area and symmetry related molecules.
4. **AXISSEARCH**- changes axis and cell.
5. **BPLOT**- plot average B factors along the chain.
6. **BRKSORT**- sorts atoms within each residue in PDB file.
7. **COMPAR**- comparison of coordinates.
8. **DERIV**- calculates gradients from the difference electron density map.
9. **DIFAREA**- compares solvent accessibilities on two data sets.
10. **DIFRES**- compares two files of atomic coordinates in PDB format.
11. **HASSP**- Patterson search program.
12. **HAVECS**- generates Patterson vectors from atom coordinates.
13. **HBOND**- calculates possible main chain H-bonds.
14. **HELIXANG**- distances and angles between helices.
15. **LIMITS**- calculates limits of x,y,z values for a given set of coordinates.
16. **MAPSTAMP**- writes MAP and machine stamp into map file.
17. **MAPBRICK**- prepares bricked maps for MRC FRODO.
18. **MAPCONT**- contours density map in brick format for MRC FRODO.
19. **MAPREPLACE**- combines parts of two maps.
20. **MATHINGS**- various operations on matrices from rotation functions.
21. **MATROTS**- changes matrices according to different programs.
22. **PDBROT**- rotate PDB file to match RIBBON picture.
23. **PLTOUT**- driver for POSTPLOT files.
24. **POSTPLOT**- part of RIBBON.

25. **REFORM**- various on-line coordinates transformations.
26. **RESAREA**- prints out solvent accessible areas for each residue.
27. **RIBBON**- just RIBBON
28. **RIBROT**- rotates RIBBON picture.
29. **SPLITD**- splits protein in the POSTPLOT file into number of files which can be coloured in FRODO.
30. **SUPERPOSE**- superposition of two molecules.
31. **SYMGEN**- generates symmetry equivalents.
32. **VECSUM**- program for automated Patterson solution.
33. **WATERAREA**- solvent accessibility.
34. **WATPEAK**- selects peaks from peakmax and puts them close to the respective protein atoms.
35. **XPLOR2PDB**- changes atom format from XPLOR to PDB one.
36. **ZEROED**- sets part of the map to zero.

**List of the CCP4 programs which are very likely to be removed
from the distribution kit in the next release of the CCP4
package**

1. **ACCSA**- accessible surface area.
2. **ATCONV**- converts various formats to PDB.
3. **ATPEAK**- takes peaks from PEAKSEARCH and finds atom in the atom file which are close to them.
4. **BAVSC2**- as BAVERAGE.
5. **BRK2DIA**- converts PDB to Diamond format.
6. **BRKMAT**- adds MATRIX cards to PDB file.
7. **BURGESS**- secondary structure prediction.
8. **CADPRO**- process CAD4 data to be suitable for profile fitting program PROFIT.
9. **CONVRNXP**- converts various formats to PDB.
10. **COORDSWAP**- as above.
11. **COORD_CONV**- as above.
12. **CROSSEC**- calculates f' and f'' at different wavelenghts.
13. **DABS**- smooths diffractometer absorption curves.
14. **DEG**- linear decay correction to reference reflections.
15. **DIFFOP**- argus diffractometer data processing program.
16. **MAPFORM**- converts map to formatted version.
17. **MINLIM**- calculates limits of x,y,z values for a given set of coordinates.
18. **O2BRK**- converts O type PDB file into standard one.
19. **PATVEC**- generate Patterson peaks from atomic positions.
20. **PEAKSEARCH**- just a peaksearch.
21. **RESECTION**- changes order of axes in the map.
22. **RESLICER**- as above?

23. **SKEWBRK**- applies matrix operations to PDB atom file.
24. **TIDYTEXT**- John Campbell text formatting program.
25. **XYZLIM**- calculates limits of x,y,z coordinates for the given set of atoms.
26. **XYZ_MODIFY**- transforms coordinates

Postdoctoral Research Positions

for two years, to perform

Fundamental Studies on Protein Crystal Growth

One position is at Nijmegen University, The Netherlands

Required: experience in crystal growth or solid state physics or chemistry

Info: Prof. Dr. P. Bennema

Laboratory for Solid State Chemistry,

Toernooiveld, 6525 ED Nijmegen,

tel.: 31-(0)80-653070

The other is at Groningen University, The Netherlands

Required: experience in molecular dynamics simulations

Info: Prof. Dr. H.J.C. Berendsen

Laboratory for Biophysical Chemistry

Nijenborgh 4, 9747 AG Groningen

tel.: 31-(0)50-634323 / 634320

Information can also be obtained from: Prof. Dr. J. Drenth,

Laboratory for Biophysical Chemistry,

Nijenborgh 4, 9747 AG Groningen,

tel.: 31-(0)50-634382; Fax: 31-(0)-50-634800

RESEARCH ASSISTANT

A post is available in a multidisciplinary team studying physical properties of filamentous bacteriophages. These are simple nucleoprotein complexes that assemble as they extrude through bacterial membranes. The research of the team includes study of the following properties of the bacteriophage particle and its precursors in assembly:

1. Structure and dynamics, using synchrotron X-ray fibre diffraction (with Dr C Nave at the SERC Daresbury Laboratory).

2. Electronic and electro-optical properties using techniques and concepts developed to study organic polymers (with Dr R H Friend at the Cavendish Laboratory, University of Cambridge).

3. Directed genetic modification to create defined changes in the properties studied under (1) and (2) (with Professor R N Perham at the Department of Biochemistry, University of Cambridge).

Candidates should be trained in one of the physical or chemical sciences. Duties will include preparation and purification of bacteriophage for physical studies, and training in the necessary microbiological techniques will be given if required. Salary 12129-14359 pounds per annum (under review). Preliminary enquires may be directed to Dr D A Marvin, Department of Biochemistry, University of Cambridge (E-mail: UK.AC.CAM.BIO.MB1::DAM4) or to any of those named above. Applications, including a CV and the names of 2 referees, and marked "molecular electronics" should be sent to: Professor R N Perham, Department of Biochemistry, University of Cambridge, Tennis Court Rd, Cambridge CB2 1QW, UK.

