DARESBURY LABORATORY

INFORMATION QUARTERLY for PROTEIN CRYSTALLOGRAPHY

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

Number 3

May 1980

Contents

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

EDITORIAL

PELLA MACHIN

Both of the working groups have met during the last couple of months, and the minutes of the group 2 meeting are included for information in this newsletter. It appears that the CCP is well under way now.

One suggestion made by the working groups was that a study weekend should be held. Plans for this are now going ahead and during the weekend of November 14-16th, 1980, a meeting will be held at Daresbury on the subject of 'Refinement of Protein Structure'. The format of a Daresbury Study Weekend is to assemble on the Friday evening for lectures and discussions through Saturday to the Sunday afternoon. Accommodation is usually at a local hotel. We hope to invite Wayne Hendrickson and Mike James to speak at the meeting.

Dr. Bricogne (Cambridge) and myself will be attending a workshop organised by NRCC in the USA at the beginning of June on MIR phasing. I will hope to report on such topics as, collaborative programming, the latest in MIR phasing and programming in RATMAC in the next newsletter.

Meanwhile, thank you for the contributions which I have received for this newsletter - we hope for even more reports from the universities in the next one.

MINUTES OF A MEETING OF WORKING GROUP 2 OF THE PROTEIN CRYSTALLOGRAPHY COLLABORATIVE COMPUTATIONAL PROJECT WHICH WAS HELD AT THE DARESBURY LABORATORY AT 11.30 ON 20TH MARCH, 1980

Present:	Dr. Alan Wonacott (Chairman)	Imperial College, London
	Dr. T. Bhat	" " "
	Dr. David Moss	Birkbeck College, London
	Dr. Ian Tickle	" " "
	Dr. Phil Bourne	Sheffield University
	Dr. Keith Wilson	Oxford University
	Dr. Phil Evans	MRC, Cambridge
	Mr. Don Akrigg	Leeds University
	Dr. John Helliwell	Keele University
	Dr. John Campbell	SRC, Daresbury
	Dr. Mike Elder	" "
	Ms. Pella Machin (Secretary)	" "

1. Progress reports were given by the various groups.

(i) David Moss reported that the diffractometer data processing programs were running satisfactorily. Most data were preprocessed on site at Birkbeck then transferred to Daresbury by magnetic tape. No other implementation work had been done but much time had been spent running the refinement program RESTRAIN (Moss, Morffew).

(ii) Bhat reported that the new IBM sort/merge was now available and in use. The density modification program had been implemented but required documentation. Work on the NA2 data formats had progressed and is reported in a later section (7).

(iii) At Daresbury, John Campbell had continued work with the Isaacs-Agarwal refinement program, testing the various versions, improving output formats etc. Versions are currently available for space groups

P1 (1) $P2_1$ (4) $P2_12_12$ (18)

$P2_12_12_1$ (19) $P4_12_12$ (92) $P3_121$ (152)

Phil Evans said that Cambridge had run some of these programs and noticed differences in the structure factor calculation results compared with those obtained with another program. It was agreed that John Campbell should compare the structure factor results with those from a classical structure factor program.

(iv) Pella Machin was implementing the PLUTO program as supplied by Phil Evans. This particular version can be used to produce contour maps and/or molecule drawings, superimposed where necessary. It was being tested with data supplied by the Sheffield group. It was suggested that if possible plotter output should be available at Daresbury (Versalec and Calcomp) and optionally on the FR80.

The Daresbury group were also implementing the Crowther rotation function program, CAD, and Eleanor Dodson's SEARCH program in association with a user at Birkbeck. In particular Mike Elder was implementing CAD, which will be important in file handling.

(v) Keith Wilson reported that the money for the Oxford workstation to Daresbury is now available.

(vi) Phil Bourne had been using programs and uncovering bugs in them! He requested that the program RECOM be available and Bhat reported that Imperial already have a version.

He also reported work done on the comparison of protein structures based on the Brookhaven data bank and suggested that the Brookhaven data should be available at Daresbury. This was agreed and it was decided that the secretary should write to Olga Kennard requesting a copy of it for the CCP.

It was suggested (Keith Wilson) that in future meetings progress reports would be handled better in association with a list of programs, with a view to checking the status of programs in an orderly fashion.

2. Documentation standards were discussed in association with John Campbell's discussion paper (CCP4/80/1).

(i) It was agreed that a standard was necessary and that in general those laid out in the paper were acceptable.

(ii) Bhat suggested that catalogued procedures should be used rather than CLISTS (because of the ease of concatenating data sets). However it was agreed that CLISTS should be provided when possible and that in special cases the JCL would additionally be listed.

(iii) Concerning output parameters it was agreed that the default MSGCLASS should be T and that the ROUTE parameter should be used for directing output to workstations.

(iv) Concern was expressed about the problem of using the NULLFILE option and the related necessity of complete file names. It was suggested by Ian Tickle that a command SET existed which would resolve this problem. It was agreed that the user support group at Daresbury (USG) be encouraged to implement the SET command as soon as possible.

It was agreed that available documentation should be sent to Daresbury where it could be entered to the word processor. A sample output would then be returned to the author for comment and amendment.

3. Mike Elder recommended the use of the PFORT verifier on any new code, to list deviations from standard FORTRAN. This was agreed. Unfortunately USG have not yet moved the program from the Rutherford Laboratory IBM's.

4. The possible use of the preprocessor RATMAC was discussed.

(i) Mike Elder said that this was being used in America and that it had several good features, for example, common block definition and general logical structure. He asked for a volunteer to write a new program using RATMAC.

(ii) Phil Evans said that very few programs were written from scratch, most were amalgamations of other programs which were then modified and extended. This meant that Fortran had to be used throughout.

(iii) Although no volunteers were forthcoming it was agreed that the possibility of mixing RATMAC and FORTRAN should be investigated and any advance in this area be reported to Mike Elder.

5. Possible uses of the CCP budget were considered. It was felt that firstly the budget should be used to cover working group meeting expenses and visits between universities. However it was also suggested that the CCP would benefit from various other activities if funds were available from the CCP budget or elsewhere.

(i) It was suggested that a Daresbury Laboratory study weekend be held on the subject of 'Refinement of Protein Structure' (covering least squares and immediately related model refinement rather than phase refinement). If possible the meeting should be held in September for 30-50 people. Suggested invited speakers from abroad were W. Hendrikson and possibly W. Steigman.

(ii) Mike Elder reported his correspondence with NRCC. It was suggested that Pella Machin should visit NRCC in the States to promote collaboration with them, to report on CCP activities and to attend the NRCC workshop. If possible she should take some data to run on the MIR program at the workshop. Various groups offered to supply such data.

(iii) The Oxford and Birkbeck groups suggested Alwyn Jones be invited to carry out further implementation of his program FRODO with the Evans and Sutherland graphics.

It was thought that in the long term this would benefit the whole CCP (most groups were aiming to carry out refinements and interactive graphics) but that as it was a rather biased application in the short term, Birkbeck and Oxford should contribute to the cost as well as the CCP.

The chairman agreed to write up these proposals and to submit them to the next meeting of Working Group 1.

6. The general topic of data formats was discussed again.

(i) Following the previous meeting, Bob Diamond reported the suggestions made by the working group to his international committee.

(ii) The group restated that the CCP agreed to differ from NRCC and did not wish to use the Xtal 80 data format, which was considered wasteful and inflexible compared with the proposed NA2 format. However it was agreed that a program should be available to extract relevant data from an Xtal 80 file.

(iii) The format of the MAP file was discussed. Only one format is currently in use and this was thought to be satisfactory in general, but could be improved by the addition of the following items to the initial title card (20A4):

NSEC	Number of sections		
IU	Axis order	Fast	where X=1, Y=2, Z = 3
IV		Medium	
IW		Slow	

It was suggested that a terminator was unnecessary and reading of sections would be based on NSEC or be terminated using the FORTRAN "END =" option.

7. The implementation of the NA2 format was discussed in detail in association with the paper "Details of binary data format for reflection data" (CCP4/80/4).

(i) The routines written by Bhat to handle the NA2 format are a mixture of FORTRAN, PL1 and assembler, the latter two being used for efficiency, ease of character manipulation, and dynamic store allocation. The group felt very strongly that for ease of implementation and compatibility all routines should be in FORTRAN if possible. Members of the group would be using non-IBM computers such as DEC, VAX, CRAY-1 and 2980 which would not support PL1 or IBM assembler.

It was agreed that nonstandard parts of the routines must be minimal, modular and well documented and that dynamic store allocation should be replaced by a fixed maximum record length array. (e.g. 400 words)

(ii) It was agreed that header assignments should be introduced in control data files by a card(s) labelled KEY (cols 1-3). Default values would be allowed for the following items:

H, K, L } Miller Indices h, k, ℓ

S — $4 \sin^2\theta/\lambda^2$

IC — Centric Flag where 0 = centric 1 = acentric

FP — Protein F value of the native

SP — Sigma F of the native protein

(iii) The value of $4 \sin^2\theta/\lambda^2$ be stored as an integer with a scale factor of 10,000.

(iv) The values of hkℓ would be stored as signed integers and no bias would be introduced. (sorting can be done on signed binary numbers).

(v) The introduction of additional headers was suggested for the following types of information:

Cell dimensions (Real, A^o and degrees)
Symmetry (as x, y, z etc.)
Lattice type (as a letter)
Batch numbers/or the number of batches

(vi) It was agreed that 4 byte integers were necessary for storing intensity values.

It was agreed that the Imperial group should implement these NA2 routines in the scaling and merging programs while the Daresbury group include them in the CAD program. CAD could also be used to convert files from old formats to NA2 and vice versa.

8. It was decided that as Phil Bourne was planning to attend ECM6 he should present a poster session on behalf of the CCP in order to state the aims of CCP4 and to encourage cooperation with groups in the future.

9. Phil Evans made a criticism of the Brookhaven coordinate format concerning the need to scan for ATOM cards, and the difficulty of identifying atom types. However, it was thought that programming methods could be devised to handle these problems and that the value of the standardisation of the file outweighed these complications.

10. Following from the previous meeting the state of the IBM 370/165 service was discussed.

(i) The general feeling was that the service had improved a lot but that there were still severe problems. Daresbury should be encouraged to continue improvements as quickly as possible such as by the implementation of HASP.

(ii) The job queue problems persist but will hopefully be cured by HASP.

(iii) It was noted that there were problems with shortage of disk space, archiving, and shortage of dial up lines.

(iv) Bhat suggested users should be encouraged to save LOAD library space, by completing the LINK step at run time. Storing fully linked programs was very wasteful of space as, for example, they will include copies of Fortran library routines.

11. Keith Wilson asked about the state of the CRAY-1 computer as the Oxford group had several refinements which they urgently wished to run. It was reported that the Logica software required to return CRAY data sets to the IBM was still not operational.

12. Mike Elder asked the group for advice concerning the request from IBM (Winchester) to license CCP programs for research purposes. It was thought that the matter should be referred to Working Group 1.

13. John Campbell reported on the existence of the program system ROCKS (CCP4/80/5). It was felt that the system had some good facilities and would be useful for comparison purposes. It was suggested that only a limited time be spent on it, though work should go ahead to at least try running the supplied IBM load module.

14. John Campbell described the message and log files which he had introduced to inform users of general activities and specific program changes.

15. The need for a 3 colour pen Calcomp was restated. It was agreed that Daresbury should be approached about this again and that if necessary working group 1 should be consulted about the possibility of purchasing a new plotter device.

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

BILL PULFORD (OXFORD UNIVERSITY)

The konnert refinement program has been tested on the CRAY using Tortoise Lysozyme data. After testing the program some useful refinement was carried out.

During a recent visit to Daresbury I ran many (over 15) test cycles to obtain some idea of the best parameters to use. I am now happy that at high resolution a value of W_F which gives at least

$$\Sigma\ W_F\ (F_o - F_c)^2 \ = \ 10\ X\ \Sigma\ W_d\ (d_o - d_c)^2$$

gives reasonable convergence. (See konnert writeup for details).

I ran seven real cycles of refinement on Tortoise Lysozyme which reduced its R factor for 1.6A data from 25.6 to 23.6. (Note that this is not representative for most konnert refinement which goes more quickly).

The program which currently resides on the IBM 370/165 (as file DPW.PROLSQT.CRAYJCL) is set up for space groups 4, 5, 18, 19 and is of sufficient size to handle a protein of TIM size (3000) without further alteration. The routine for space group $P4_32,2$ (96) has now been written (Guiseppe Zanotti) and debugged (Peter Aibymiuk) and I will put it on the CRAY-1 at the next opportunity.

There are 3 other programs associated with the main refinement program. These are stored in files on the IBM as DPW. PROTIND. CRAYJCL, DPW. FORMR. CRAYJCL. DPW. UNBIN. CRAYJCL. Anyone who wants further details about this should contact me.

REPORT ON COMPUTING BY THE BIRKBECK PROTEIN GROUP

DAVID MOSS (BIRKBECK COLLEGE, LONDON)

1) Data Processing

The following programmes have been used successfully at Daresbury.

(i) DIFVAL which validates diffractometer data derived from paper tapes. Currently, we transfer our paper tape data to magnetic tape locally and send the magnetic tape to Daresbury by post. Soon we hope to be able to use the paper tape reader on our workstation.

(ii) HRS which scales together diffractometer data acquired from several crystals by the method of least squares.

(iii) DSCALE which applies several scaling methods, including Kraut scaling, to data from heavy atom derivatives.

2) Phase Determination

We have implemented and are using:-

(i) REFINE which refines the positions of heavy atoms in protein derivatives.

(ii) PHASE which calculates phases and figures ofmerit from these positions.

(iii) Programmes originally written by G. Bricogne for determination of Hendrickson and Lattman coefficients and their application to the combination of phases from several sources.

3) Refinement

Five protein structures are in the course of refinement by restrained least squares using the programme RESTRAIN.

The refinement of ribonuclease began with a reasonably precise model built from a 2.5Å map. The behaviour of the temperature factors at 2.3Å gives hope that it will be possible to proceed with our 1.5Å data with little rebuilding of the model.

Three insulin refinements have been carried out using the co-ordinates of porcine insulin as a starting point. The refinement of native bovine insulin using 2.3Å data proceeded uneventfully to R = 21%, excluding solvent molecules. Two modified bovine insulins have been refined. In one case about six residues proved to be outside the radius of convergence of the refinement, thus requiring further graphics work, and in the other case the high temperature factors ($U \simeq 0.4 Å^2$) indicate a need to review the scaling of the data.

The refinement of an acid proteinase (EPR) from the fungus Endothia Parasitica began from a rather crude model built from a 2.7Å map. Refinement cycles were carried out at resolutions starting at 3.5Å. Areas of the structure which were quite wrong, were often forced right off the electron density when resulting maps were re-examined. Good areas were often improved. The model is now being re-examined in the graphics system before further refinement is attempted.

In order to improve the convergence properties of restrained least squares further experiments are being carried out at 3.5Å resolution. These will centre on the weighting of the geometric restraints and the use of isomorphous phases as restraints.

In terms of computer time, most of our computing at Daresbury will involve least squares refinement. We are currently implementing RESTRAIN on the CRAY as larger refinements will be beyond the resources of the IBM 370. We hope to run tests on ribonuclease using the CRAY before the end of June.

4) Contouring Electron Density Maps

We are currently exploring two methods of contouring our maps.

(i) PLUTO with output on the Versatec electrostatic plotter at Daresbury.

(ii) CONTOUR with microfilm output on the FR80 at Rutherford. This can be copied to Xerox paper or projected onto transparent film for minimaps.

5) Evans and Sutherland Graphics System

Two programmes, BILDER and FRODO, have been successfully implemented and used on our graphics system for building models of protein molecules from contoured electron density maps.

BILDER has been used to construct a preliminary model of EPR. This work has shown that in the hands of a skilled user, BILDER offers most of the facilities which are required for building a protein molecule into an electron density map.

FRODO has been used for the construction of models of several proteins. There are two advantages of FRODO which are much appreciated. Firstly the user can see on the screen all the molecule and associated electron density within a sphere of given radius about a specified point enabling him to investigate and adjust interchain contacts. Currently this radius is not always large enough and we intend to acquire more Picture System memory to enable the use of larger radii. Secondly, amino acid residues can easily be inserted, deleted, translated, rotated or have their identity changed. This is a most essential facility when there are any doubts about the residue sequence or the course of the chain in the map. We have also used the facility to build a trial model of turtle ribonuclease from the backbone of bovine ribonuclease when no electron density map was available.

In consultation with its author, Alwyn Jones, we plan to progressively enhance the Evans and Sutherland version of FRODO.

6) Data Formats

Conversion of programmes to accept atomic co-ordinate data in the format of the Brookhaven Data Bank is taking place. FRODO now produces output in this format and RESTRAIN already uses it. The ordering of atoms produced by FRODO differs from that required in the Data Bank so we have written a programme to convert from FRODO to Brookhaven ordering.

LEEDS UNIVERSITY AMDAHL V7

DONN AKRIGG

The University of Leeds has recently installed an Amdahl 470/V7 computer to replace the ICL 1906A as a central university facility. The V7 is intended as a central machine which will be connected to other main frames to form a campus network. The first machine installed as part of this network is a Prime 750, which consists of a Prime 750 2 megabyte main frame, 2 300 megabyte disks, a 9-track 800/1600 BPI magnetic tape deck, 64 asynchronous communications ports, 4 synchronous communcations ports and a medium speed pointer.

The Amdahl is compatible with the IBM 370 and 303X series machines, and consists of 6 megabytes of memory (capable of addressing up to 16 megabytes of virtual memory), ten 317 megabyte disks on two controllers, four 9-track 6250/ 1600 BPI magnetic tape decks, 1 Memorex 1270 communications controller with 96 ports, 2 line printers (one capable of two case printing), a 3270 controller for connecting Memorex 1377 VDU's to a card reader, and paper tape reader/punch station.

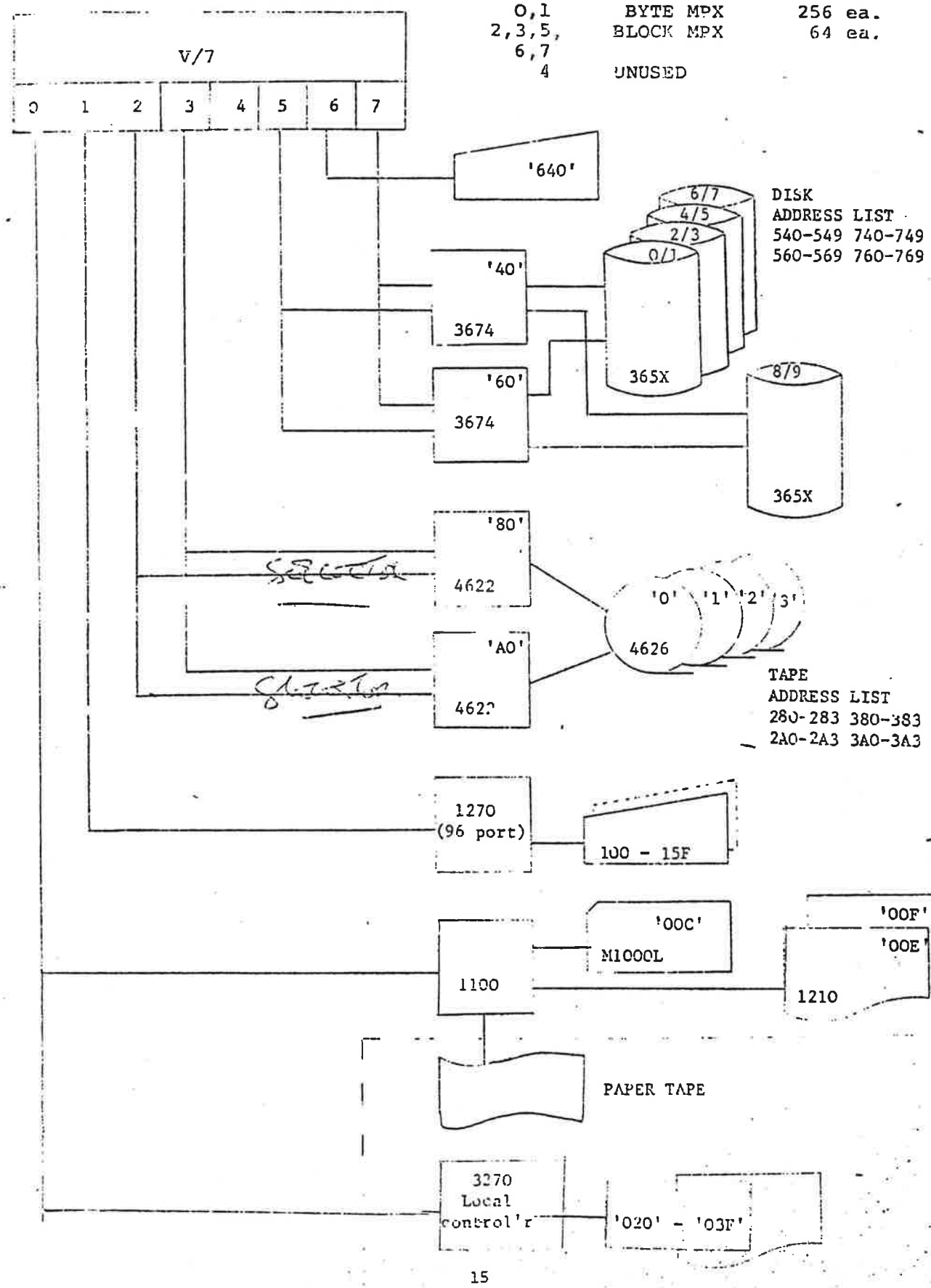
The Amdahl V7 runs under the conversational monitor system of IBM's virtual machine facility (VM/CMS). The image presented to the user by VM/CMS is that of his own personal machine with card reader, printer, punch and his own disks. These are virtual peripherals and are mapped on to real peripherals by the VM operating system. It is intended to add a second 1270 communications controller, an IBM 1403 line printer, 2 ampex 30 megabyte disk drives and 2 ampex 60 megabyte exchangeable disk drives.

The Fortran compilers on the Amdahl are supplied by IBM and consist of FORTHXE (FORTRAN extended enhanced) and FORTGI; compilers for ALGOL 60, ALGOL 68, PL1 and Pascal are also available.

The Leeds protein crystallography group intend to mount programs on the V7 which are currently running on IBM installations at Cambridge and Daresbury. The departmental PDP11/45, which is used for interactive graphics work, has been connected to the Amdahl, and software is under test for running a HASP work-station emulator on the PDP11/45 to allow file and job transfer between the 11/45 and the Amdahl.

LEEDS UNIVERSITY V/7 PERIPHERAL ADDRESSES

CHAN.	TYPE	SUBCHAN
0,1	BYTE MPX	256 ea.
2,3,5, 6,7	BLOCK MPX	64 ea.
4	UNUSED	



BROOKHAVEN PROTEIN DATA BANK

PELLA MACHIN (DARESBURY LABORATORY)

At the meeting of the working group on 20th March, 1980, a suggestion was made that the Brookhaven Protein Data Bank should be available on magnetic tape at Daresbury. As secretary of the group I wrote to Dr. Olga Kennard (Cambridge) requesting a copy of the files for the CCP and have since received the data files.

Magnetic Tape Details

The data are available on tape XBB704

9-track
1600 bpi
NL

Each record is 80 characters in length and the block size is 4800 bytes.

There are 138 files, as described below.

The JCL necessary to access the tape (for example from a FORTRAN program, using unit 2, reading file 11) is

```
//FT02F001 DD VOL=SER=XBB704,UNIT=DEN1600,
//    DISP=(OLD,KEEP),LABEL=(11,NL,,IN),
//    DCB=(RECFM=FB,LRECL=80,BLKSIZE=4800)
```

Tape Directory

Files 1-12 of the tape contain FORTRAN programs:

File	Name	Purpose	Author
1	TAPDIR	Print directory of tape contents	H. Bernstein, F. Bernstein
2	BENDER	Parameters for bent-wire models	G. Williams
3	CONECT	General full connectivity	F. Bernstein
4	CONTCT	Intermolecular contacts	L. Andrews
5	DGPLOT	Diagonal plots on printer	E. Swanson, F. Bernstein
6	DSTNCE	Calculate distances from CONECT records	F. Bernstein
7	FISIPL	Phi/psi plots on the printer	F. Bernstein
8	NAMOD	Ball and stick model display	Y. Beppu
9	PHIPSI	Main chain torsion angles	L. Andrews, G. Williams, F. Bernstein
10	STEREO	Extract X, Y, Z from stereo diagrams	M. Rossmann
11	TORSRU	Complete torsion angles	G. Reeke
12	TOTALS	Validation of Master record	L. Andrews, F. Bernstein

The remaining files (13-138) are the protein data bank entries. The following table describes the contents of each tape file.

13	1APE	ACID PROTEINASE (E.C.3.4.23.10), ENDOTHIAPEPSIN	T.L.BLUNDELL,B.T.SEWELL,J.A.JENKINS,I.J.TICKLE
14	1APR	ACID PROTEASE (E.C.3.4.23.9)	R.BOTT,M.LIU,E.SUBRAMANIAN,D.R.DAVIES
15	1ACT	ACTINIDIN (SULFHYDRYL PROTEINASE) (E.C. NUMBER NOT ASSIGNED)	E.N.BAKER
16	2ADK	ADENYLATE KINASE (E.C.2.7.4.3)	G.E.SCHULZ AND *CO-*WORKERS
17	1WGA	WHEAT GERM AGGLUTININ	C.S.WRIGHT
18	1ADH	LIVER ALCOHOL DEHYDROGENASE (E.C.1.1.1.1) COMPLEX WITH ADENOSINE DIPHOSPHATE-RIBOSE	C.-I.BRANDEN,H.EKLUND,B.NORDSTROM,E.ZEPPEZAUER,
19	2ADH	LIVER ALCOHOL DEHYDROGENASE (E.C.1.1.1.1) COMPLEX WITH ORTHOPHENANTHROLINE	C.-I.BRANDEN,H.EKLUND,B.NORDSTROM,E.ZEPPEZAUER,
20	4ADH	APO-LIVER ALCOHOL DEHYDROGENASE (E.C.1.1.1.1)	C.-I.BRANDEN,H.EKLUND,B.NORDSTROM,E.ZEPPEZAUER,
21	1ALP	ALPHA LYTIC PROTEASE (E.C.3.4.21.12)	G.D.BRAYER,L.T.J.DELBAERE,M.N.G.JAMES
22	1ATC	ASPARTATE CARBAMOYLTRANSFERASE (ASPARTATE TRANSCARBA (E.C.2.1.3.2)	J.L.CRAWFORD,H.L.MONACO,W.N.LIPSCOMB
23	2BCL	BACTERIOCHLOROPHYLL-A PROTEIN	B.W.MATTHEWS,R.E.FENNA,M.C.BOLOGNESI,M.F.SCHMID
24	1CPV:	CALCIUM-BINDING PARVALBUMIN B	P.C.MOEWS,R.H.KRETSINGER
25	2CPV	CALCIUM-BINDING PARVALBUMIN B	P.C.MOEWS,R.H.KRETSINGER
26	3CPV	CALCIUM-BINDING PARVALBUMIN B	P.C.MOEWS,R.H.KRETSINGER
27	1CAB	CARBONIC ANHYDRASE FORM B (CARBONATE DEHYDRATASE) (E.C.4.2.1.1)	K.K.KANNAN,B.NOTSTRAND,K.FRIDBORG,M.PETEF,
28	1CAC	CARBONIC ANHYDRASE FORM C (CARBONATE DEHYDRATASE) (E.C.4.2.1.1)	K.K.KANNAN,I.VAARA,A.LILJAS,P.BERGSTEN,K.FRIDBORG,
29	1CPA	CARBOXYPEPTIDASE A (E.C.3.4.12.2)	F.A.QUIOCHO,W.N.LIPSCOMB
30	1CPB	CARBOXYPEPTIDASE B (E.C.3.4.12.3) FRACTION II	M.F.SCHMID,J.R.HERRIOTT
31	2CHA	ALPHA CHYMOTRYPSIN A (TOSYLATED) (E.C.3.4.21.1)	J.J.BIRKTOFT,D.M.BLOW
32	3CHA	ALPHA CHYMOTRYPSIN A (E.C.3.4.21.1)	A.TULINSKY,R.L.VANDLEN
33	1GCH	GAMMA CHYMOTRYPSIN A (E.C.3.4.21.1)	G.H.COHEN,D.R.DAVIES,E.W.SILVERTON
34	1CHG	CHYMOTRYPSINOGEN A	S.T.FREER,J.KRAUT,J.D.ROBERTUS,H.T.WRIGHT,
35	2CNA	CONCANAVALIN A	G.N.REEKE *JUNIOR,J.W.BECKER,G.M.EDELMAN
36	3CNA	CONCANAVALIN A	K.D.HARDMAN,C.F.AINSWORTH
37	2B5C	CYTOCHROME B5 (OXIDIZED)	F.S.MATHEWS,E.W.CZERWINSKI
38	156B	CYTROCHROME B562 (E. COLI, OXIDIZED)	P.H.BETHGE,E.W.CZERWINSKI,F.S.MATHEWS
39	1CYT	CYTOCHROME C (OXIDIZED).	R.SWANSON,B.L.TRUS,N.MANDEL,G.MANDEL,O.B.KALLAI,
40	2CYT	CYTOCHROME C (REDUCED).	T.TAKANO,B.L.TRUS,N.MANDEL,G.MANDEL,O.B.KALLAI,
41	1CYC	FERROCYTOCHROME C	N.TANAKA,T.YAMANE,T.TSUKIHARA,T.ASHIDA,M.KAKUDO
42	1C2C	FERRICYTOCHROME $C=2=	F.R.SALEMME,S.T.FREER,N.H.XUONG,R.A.ALDEN,J.KRAUT
43	155C	CYTOCHROME C550	R.TIMKOVICH
44	251C	CYTOCHROME C551 (OXIDIZED)	R.J.ALMASSY,J.L.CHAMBERS,R.E.DICKERSON
45	1EST	TOSYL-ELASTASE (E.C.3.4.21.11)	L.SAWYER,D.M.SHOTTON,H.C.WATSON
46	1ECD	HEMOGLOBIN (ERYTHROCRUORIN, DEOXY)	W.STEIGEMANN,E.WEBER
47	1ECO	HEMOGLOBIN (ERYTHROCRUORIN, CARBONMONOXY)	W.STEIGEMANN,E.WEBER
48	1ECA	HEMOGLOBIN (ERYTHROCRUORIN, AQUO MET)	W.STEIGEMANN,E.WEBER
49	1ECN	HEMOGLOBIN (ERYTHROCRUORIN, CYANO MET)	W.STEIGEMANN,E.WEBER
50	1FDX	FERREDOXIN	E.T.ADMAN,L.C.SIEKER,L.H.JENSEN
51	1FXC	FERREDOXIN	K.FUKUYAMA,T.TSUKIHARA,Y.KATSUBE,Y.MATSUURA,N.TANAKA
52	3FXN	FLAVODOXIN (OXIDIZED FORM)	M.L.LUDWIG
53	4FXN	FLAVODOXIN (SEMIQUINONE FORM)	M.L.LUDWIG
54	1GCN	GLUCAGON (PH 6 - PH 7 FORM)	T.L.BLUNDELL,K.SASAKI,S.DOCKERILL,I.J.TICKLE
55	1PGI	D-GLUCOSE 6-PHOSPHATE ISOMERASE (E.C.5.3.1.9)	H.MUIRHEAD
56	1GPD	D-GYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (E.C.1.2.1.12)	D.MORAS,K.W.OLSEN,M.N.SABESAN,M.BUEHNER,G.C.FORD,
57	1HRB	HEMERYTHRIN B	W.A.HENDRICKSON,K.B.WARD
58	1HMN	HEMERYTHRIN (MET,AQUO)	R.E.STENKAMP,L.C.SIEKER,L.H.JENSEN,J.E.MC*QUEEN JUNI
59	1HDS	HEMOGLOBIN (SICKLE CELL)	E.L.AMMA,R.L.GIRLING
60	2MHB	HEMOGLOBIN (HORSE, AQUO MET)	R.C.LADNER,E.G.HEIDNER,M.F.PERUTZ
61	2DHB	HEMOGLOBIN (HORSE,DEOXY)	M.F.PERUTZ ET AL.
62	1HHB	HEMOGLOBIN (DEOXY)	G.FERMI,M.F.PERUTZ
63	1HCO	HEMOGLOBIN (CARBONMONOXY)	J.M.BALDWIN
64	2HCO	HEMOGLOBIN (CARBONMONOXY)	J.M.BALDWIN

65	1FDH	HUMAN FETAL DEOXYHEMOGLOBIN F /II	J.A.FRIER *JUNIOR
66	1LHB	HEMOGLOBIN(MET)-CYANIDE V	W.A.HENDRICKSON,W.E.LOVE,J.KARLE
67	2YHX	YEAST HEXOKINASE B (E.C.2.7.1.1) COMPLEX WITH ORTHO-TOLUOYLGLUCOSAMINE	T.A.STEITZ,C.M.ANDERSON,R.E.STENKAMP
68	1HIP	OXIDIZED HIGH POTENTIAL IRON PROTEIN (HIPIP).	C.W.CARTER *JUNIOR,J.KRAUT,S.T.FREER,N.-H.XUONG,
69	1HYA	HYALURONIC ACID (POLY D-GLUCURONIC ACID-N-ACETYL-D -GLUCOSAMINE). THE GLUCURONIC ACID-GLUCOSAMINE LINKAGE IS BETA(1,3) AND THE GLUCOSAMINE-GLUCURONIC ACID LINKAGE IS BETA(1,4)	S.ARNOTT
70	1FAB	LAMBDA IMMUNOGLOBULIN FAB*	R.J.POLJAK,L.M.AMZEL,B.L.CHEN,R.P.PHIZACKERLEY,
71	1MCG	IMMUNOGLOBULIN, LAMBDA-*TYPE BENCE-*JONES DIMER MCG	E.E.ABOLA,A.B.EDMUNDSON,K.R.ELY,R.L.GIRLING,
72	1REI	BENCE-*JONES IMMUNOGLOBULIN /REI$ VARIABLE PORTION	O.EPP,E.E.LATTMAN,P.COLMAN,H.FEHLHAMMER,W.BODE,
73	1RHE	IMMUNOGLOBULIN, LAMBDA-TYPE BENCE-*JONES DIMER RHE	B.C.WANG,C.S.YOO,M.SAX
74	1KGA	2-KETO-3-DEOXY-6-PHOSPHOGLUCONATE (/KDPG$) ALDOLASE (E.C.4.1.2.14)	A.TULINSKY
75	4LDH	LACTATE DEHYDROGENASE (E.C.1.1.1.27) APO ENZYME M4	J.L.WHITE,M.L.HACKERT,M.BUEHNER,M.J.ADAMS,
76	3LDH	LACTATE DEHYDROGENASE (E.C.1.1.1.27) M4 ENZYME, TERNARY COMPLEX WITH /NAD$ AND PYRUVATE	J.L.WHITE,M.L.HACKERT,M.BUEHNER,M.J.ADAMS,
77	1LDX	LACTATE DEHYDROGENASE (E.C.1.1.1.27), ISOENZYME C=4	=W.D.L.MUSICK,M.G.ROSSMANN
78	1HBL	LEGHEMOGLOBIN (ACETATE,MET)	B.K.VAINSHTEIN,E.H.HARUTYUNYAN,I.P.KURANOVA,V.V.BORI
79	1LZM	LYSOZYME (E.C.3.2.1.17)	S.J.REMINGTON,L.F.TEN *EYCK,B.W.MATTHEWS
80	1LYZ	LYSOZYME (E.C.3.2.1.17)	R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH
81	2LYZ	LYSOZYME (E.C.3.2.1.17)	R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH
82	3LYZ	LYSOZYME (E.C.3.2.1.17)	R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH
83	4LYZ	LYSOZYME (E.C.3.2.1.17)	R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH
84	5LYZ	LYSOZYME (E.C.3.2.1.17)	R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH
85	6LYZ	LYSOZYME (E.C.3.2.1.17)	R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH
86	7LYZ	LYSOZYME (E.C.3.2.1.17) TRICLINIC CRYSTAL FORM	J.MOULT,A.YONATH,J.SUSSMAN,O.HERZBERG,A.PODJARNY,
87	8LYZ	LYSOZYME (E.C.3.2.1.17) IODINE-INACTIVATED	C.R.BEDDELL,C.C.F.BLAKE,S.J.OATLEY
88	1MDH	MALATE DEHYDROGENASE (CYTOPLASMIC) (E.C.1.1.1.37)	L.J.BANASZAK
89	1MLP	MUREIN LIPOPROTEIN	A.D.MC*LACHLAN
90	1MBN	MYOGLOBIN (FERRIC IRON - METMYOGLOBIN)	H.C.WATSON,J.C.KENDREW
91	2MBN	MYOGLOBIN (MET)	T.TAKANO
92	3MBN	MYOGLOBIN (DEOXY)	T.TAKANO
93	1MBS	SEAL MYOGLOBIN (MET)	H.SCOULOUDI
94	1MHR	MYOHEMERYTHRIN	W.A.HENDRICKSON,K.B.WARD
95	8PAP	PAPAIN (E.C.3.4.22.2)	J.DRENTH,J.N.JANSONIUS,R.KOEKOEK,H.M.SWEN,
96	1PAD	PAPAIN (E.C.3.4.22.2) -ACETYL-ALANYL-ALANYL-PHENYLALANYL-METHYLENYLALANYL DERIVATIVE OF CYSTEINE 25 (/ACAAPACK)	J.DRENTH,K.H.KALK,H.M.SWEN
97	2PAD	PAPAIN (E.C.3.4.22.2) -CYSTEINYL DERIVATIVE OF CYSTEINE-25 (/PAPSSCYS)	J.DRENTH,K.H.KALK,H.M.SWEN
98	3PAD	PAPAIN (E.C.3.4.22.2) -CYSTEINE-25 OXIDIZED (/PAPSX)	J.DRENTH,K.H.KALK,H.M.SWEN
99	4PAD	PAPAIN (E.C.3.4.22.2) -TOSYL-METHYLENYLLYSYL DERIVATIVE OF CYSTEINE-25 (/TLCK)	J.DRENTH
100	5PAD	PAPAIN (E.C.3.4.22.2) -BENZYLOXYCARBONYL-GLYCYL-PHENYLALANYL-METHYLENYLGLYCYL DERIVATIVE (/ZGPGCK)	J.DRENTH,K.H.KALK,H.M.SWEN
101	6PAD	PAPAIN (E.C.3.4.22.2) -BENZYLOXYCARBONYL-PHENYLALANYL-METHYLENYLALANYL DERIVATIVE (/ZPACK)	J.DRENTH,K.H.KALK,H.M.SWEN
102	1PEP	PEPSIN (E.C.3.4.23.1)	N.S.ANDREEVA,A.A.FEDOROV
103	1PGK	PHOSPHOGLYCERATE KINASE (E.C.2.7.2.3)	H.C.WATSON
104	2PGK	PHOSPHOGLYCERATE KINASE (HORSE,MUSCLE) (E.C.2.7.2.3)	R.D.BANKS,C.C.F.BLAKE,P.R.EVANS,R.M.MASER
105	1PGM	YEAST PHOSPHOGLYCERATE MUTASE (E.C.2.7.5.3)	J.W.CAMPBELL,H.C.WATSON,G.I.HODGSON
106	2PAB	PREALBUMIN (HUMAN PLASMA)	S.J.OATLEY,C.C.F.BLAKE
107	1RLX	RELAXIN	N.W.ISAACS,G.DODSON,A.C.EVANS,A.JACK,A.C.T.NORTH,R.J
108	2RLX	RELAXIN	N.W.ISAACS,G.DODSON,A.C.EVANS,A.JACK,A.C.T.NORTH,R.J

109	3RLX	RELAXIN	N.W.ISAACS,G.DODSON,A.C.EVANS,A.JACK,A.C.T.NORTH,R.J
110	4RLX	RELAXIN	N.W.ISAACS,G.DODSON,A.C.EVANS,A.JACK,A.C.T.NORTH,R.J
111	1RHD	RHODANESE (E.C.2.8.1.1)	W.G.J.HOL,J.H.PLOEGMAN,K.H.KALK,G.DRENT
112	2RSA	RIBONUCLEASE A (E.C.3.1.4.22)	A.WLODAWER
113	1RNS	RIBONUCLEASE-S (E.C.3.1.4.22)	F.M.RICHARDS,H.W.WYCKOFF
114	2RXN	RUBREDOXIN (OXIDIZED, FE(III))	L.H.JENSEN,K.D.WATENPAUGH,L.C.SIEKER,J.R.HERRIOTT
115	1SNS	STAPHYLOCOCCAL NUCLEASE (E.C.3.1.4.7) COMPLEX WITH 2*-DEOXY-3*-5*-DIPHOSPHOTHYMIDINE	F.A.COTTON,E.E.HAZEN *JUNIOR
116	1SGA	PROTEINASE A FROM STREPTOMYCES GRISEUS (/SGPA) (E.C. NUMBER NOT ASSIGNED)	G.D.BRAYER,L.T.J.DELBAERE,M.N.G.JAMES
117	2SGB	PROTEINASE B FROM STREPTOMYCES GRISEUS (/SGPB) (E.C. NUMBER NOT ASSIGNED)	L.T.J.DELBAERE,G.D.BRAYER,M.N.G.JAMES
118	1SSI	STREPTOMYCES SUBTILISIN INHIBITOR	Y.MITSUI,Y.SATOW,Y.WATANABE,Y.IITAKA
119	1SBT	SUBTILISIN /BPN$* (E.C.3.4.21.14)	R.A.ALDEN,J.J.BIRKTOFT,J.KRAUT,J.D.ROBERTUS,
120	2SBT	SUBTILISIN NOVO (E.C.3.4.21.14)	J.DRENTH,W.G.J.HOL,J.N.JANSONIUS,R.KOEKOEK
121	1SOD	CU,ZN SUPEROXIDE DISMUTASE (E.C.1.15.1.1)	J.S.RICHARDSON,K.A.THOMAS,D.C.RICHARDSON
122	1TLN	THERMOLYSIN (E.C.3.4.24.4)	B.W.MATTHEWS,L.H.WEAVER,W.R.KESTER
123	2TLN	THERMOLYSIN (E.C.3.4.24.4)	B.W.MATTHEWS,L.H.WEAVER,W.R.KESTER
124	1SRX	THIOREDOXIN REDUCTASE (/NADPH$) (E.C.1.6.4.5) (OXIDIZED FORM)	B.-O.SODERBERG
125	4TNA	TRANSFER RIBO-NUCLEIC ACID (YEAST,PHE) $T/RNA	A.JACK,J.E.LADNER,A.KLUG
126	6TNA	TRANSFER RIBO-NUCLEIC ACID (YEAST,PHE),$T/RNA	J.L.SUSSMAN,S.R.HOLBROOK,R.W.WARRANT,G.M.CHURCH,S.-H
127	8TNA	TRANSFER RIBO-NUCLEIC ACID (YEAST, PHE), TRNA	M.SUNDARALINGAM AND CO-WORKERS
128	1TIM	TRIOSE PHOSPHATE ISOMERASE (E.C.5.3.1.1)	D.W.BANNER,A.C.BLOOMER,G.A.PETSKO,D.C.PHILLIPS,
129	1PTN	BETA-TRYPSIN (NATIVE AT $P*H 8) (E.C.3.4.21.4)	H.FEHLHAMMER,W.BODE,P.SCHWAGER
130	2PTB	BETA-TRYPSIN (BENZAMIDINE INHIBITED) AT $P*H 7 (E.C.3.4.21.4)	H.FEHLHAMMER,W.BODE,P.SCHWAGER
131	1PTC	BETA-TRYPSIN (E.C.3.4.21.4) COMPLEX WITH PANCREATIC TRYPSIN INHIBITOR	A.RUEHLMANN,D.KUKLA,P.SCHWAGER,K.BARTELS,R.HUBER,
132	3PTI	TRYPSIN INHIBITOR	R.HUBER,D.KUKLA,A.RUEHLMANN,O.EPP,H.FORMANEK,
133	3PTP	BETA TRYPSIN, DIISOPROPYLPHOSPHORYL INHIBITED. (E.C.3.4.21.4)	J.L.CHAMBERS,R.M.STROUD
134	1TGP	TRYPSINOGEN (E.C.3.4.21.4) COMPLEX WITH PANCREATIC TRYPSIN INHIBITOR	W.BODE,P.SCHWAGER,R.HUBER
135	1TPI	TRYPSINOGEN (E.C.3.4.21.4) COMPLEX WITH PANCREATIC TRYPSIN INHIBITOR AND ILE-VAL	W.BODE,P.SCHWAGER,R.HUBER
136	1TGA	TRYPSINOGEN (E.C.3.4.21.4) FROM MGSO4	W.BODE,H.FEHLHAMMER,R.HUBER
137	1TGB	TRYPSINOGEN-CA (E.C.3.4.21.4) FROM PEG	W.BODE,H.FEHLHAMMER,R.HUBER
138	1TGN	TRYPSINOGEN (E.C.3.4.21.4)	A.A.KOSSIAKOFF,R.M.STROUD

PROTEIN DATA BANK

Publication Notice

In order to help us document the use of data supplied by the Bank, please indicate below any publications reporting work making use of the Bank's services. In citing the Protein Data Bank in print, we suggest that a reference be included to F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, J. Mol. Biol. 112, 535-542 (1977).

AUTHOR(s)

JOURNAL

YEAR

VOLUME

PAGE NUMBERS (inclusive)

TITLE

We would appreciate receiving reprints.

Please return this notice to

Frances C. Bernstein
Chemistry Department
Brookhaven National Laboratory
Upton, New York 11973

PROGRAM MESSAGE AND LOGGING FACILITY

JOHN CAMPBELL (DARESBURY LABORATORY)

Introduction

One of the problems in attempting to create a system from a pool of programs, used by different people in different parts of the country, is that of communication. This note describes two areas in which an attempt has been made to ease this problem:-

1. The first area, of interest to the project as a whole, concerns finding out to what extent and in what areas the system is being used. This is of use both in monitoring the project generally and for finding out who is currently likely to be using a program which is about to be modified in some way.

2. The second area involves the requirement for immediate communication of program modifications or enhancements to those who are using the program. Regular updates of documentation, sent to all users, cannot in themselves adequately cope with this problem.

The proposed solution involves the use of a log file, a short message file and programs which access these files. In particular, each major program in the system will be run with an initial job step which calls the program PCLOG. This program has two main functions:-

1. To make an entry in the log file giving details of the program being run.

2. To output a section, at the start of the line printer output for the program, listing relevant messages from the short message file. In particular such messages will relate to program modifications, bugs or

enhancements. Messages of general interest or of interest to a particular user may, however, also be printed.

The Log File

When a job calls PCLOG as its first step, the following entries will be made in the log file:-

The program name.
The space group number (if input as a CLIST parameter).
The version code.
The user I.D.
The account number.
The time requested for the program.
The core requested for the program.

The log file is filled in a cyclic manner with a maximum of 999 entries being held. Thus, when the file is full, it will always contain details of the latest 999 program runs. A facility is available to list the contents of the log file and it is proposed for future development to write a program which may be used to make a summary of the file contents over a given period.

The Short Message File

The short message file has been designed to hold messages of three different types:-

1) General messages. These are messages which will be of interest to anyone using the system.

2) User messages. These are messages directed towards a particular user e.g. SUS.

3) Program messages. These are messages concerned with a particular program e.g. PLUTO.

As mentioned above, the program PCLOG will print the relevant entries from the short message file at the start of the printer output of a program. Normally messages from all three categories would be printed, though the options may be set otherwise in the PCLOG job step. The user messages printed could only be those directed towards the person using the program and, similarly, the program messages printed could only be those relating to the program being run.

The interactive program MESSAGE is available for adding entries to the message file. It may also be used to list, modify or delete existing entries. The maximum length of a message is 72 characters and the maximum capacity of the message file is 200 messages. For longer messages, other text files may be set up with a reference being entered in the short message file.

In addition to the text of the message, the following items of information are held and printed when the message is output:-

A message number.
The date of the message.
The I.D. of the user sending the message.

Extending the Facility

It would be possible to extend the use of the short message file as a means of communication between users. This could be achieved by automatically calling a program as part of the system LOGON procedure which would list messages directed towards the user logging on together with general messages if required.