DARESBURY LABORATORY

INFORMATION QUARTERLY for PROTEIN CRYSTALLOGRAPHY

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

Number 8 March 1982

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

THE REPRESENTATION OF PROTEIN STRUCTURES

A 1½ day meeting at York University, 6/7 January 1982

Guy Dodson (York)

The decision to hold a one-day meeting was made because this kind of time does not cause too much disruption, but still allows a good chance for getting together socially. Our impression was that the meeting was enjoyed - we found the discussions interesting and useful.

There was great enthusiasm about the use of colour in presentations and general agreement that colour computer graphics would become increasingly important. Views on many aspects of representing protein structures, for example the use of stereo figures and publication of coordinates, varied greatly between individuals. Fair enough too - it is an individual's decision.

A description of the meeting follows.

This meeting was organised to review the problems in describing protein crystal structures, particularly at this time when computer graphics techniques are developing so rapidly, and also to consider what types of representation are required by the different disciplines which exploit this knowledge.

There are fundamental difficulties in the simple description of a protein crystal structure which arise from the electron density map itself. David Blow surveyed this aspect, drawing particular attention to the sources of deficiencies in isomorphously phased maps, deficiencies which introduce inaccuracy into the description of the proteins. The extent and the nature of errors in isomorphously phased maps is now recognised from refinement studies carried out in many laboratories. Procedures to improve the heavy atom parameters include better data processing, incorporating the more accurate anomalous scattering differences and the greater resolution obtainable with synchrotron radiation.

The resolution and quality of the data and phasing (together with a knowledge of the sequence) will determine whether the description is detailed or whether it simply delineates the chain direction and identifies secondary structure. It is obviously important that the limitations in the description are made clear. Most laboratories now follow their initial structural determination by refinement calculations and often considerably extend (and modify) the details of the interactions, contacts, solvent structure and so on.

Phil Evans pointed out that this is of particular importance in enzyme-substrate studies where several workers have demonstrated the electron density can be distinctly improved by refining the coordinates of the complexed protein. His own work on the various liganded forms of PFK provided support, modest in Phil's view, for this procedure.

The prealbumin-thyroid interaction presented a subtle problem of interpretation. Although the native structure had been refined nicely, the complex could not be refined without proper allowance for the statistical disorder produced by the conformational changes associated with the hormone-protein interaction. The description of this structure relied upon interpretation of the dominating iodine positions and modelling of the hormone's light atom moiety. Clearly the limiting factor was the quality of the electron density; the analysis raised, not for the first time, the intractable problem of local disorder. Jane Burridge showed some slides where the atomic surfaces had been dotted in different colours: these gave a beautiful impression but the technique seemed limited to relatively simple surfaces.

The complexity in a detailed figure of a protein sometimes makes it less than useful. For those like Cyrus Chothia and Mike Sternberg, concerned with the folding problem and the structural and topological relationships between proteins, simpler representations are needed. Both Cyrus and Mike demonstrated how their analyses of the interactions between the structural elements in proteins reveal definite packing relationships which are important in directing and stabilising the protein's folding. They urge that the protein crystallographer should provide the C_α positions as soon as is 'reasonably' possible and that the secondary structure's organisation be made clear. There is no need for accuracy in these studies which will be little affected by the sort of revisions refinement makes to the atomic positions. However, these studies benefit greatly from knowledge of as many structures as possible.

The question of what is important in the description of a protein was considered further. Sub-unit interactions are very important and their representation is essential for general understanding. For structures like viruses or ferritin however, these interactions can be immensely complex. In her account of the TMV subunit contacts, Anne Bloomer made a convincing case for selection from the detail and argued for use of schematic figures. The figures presented by Arthur Lesk, using Raster graphics, illustrated beautifully the possibilities of exploiting new techniques and simplified figures. Some of the figure-drawing programmes that Arthur described are available now.

A gap in our picture of proteins is a graphic representation of the errors present in the atomic coordinates. In an extremely brief discussion, Guy Dodson pointed out that the correlation between the thermal parameter and the positional errors could be exploited. Figures in which the atomic radius was constructed proportional to B showed vividly the character of the atomic mobility within the protein structure. Also, the likely error in the atomic positions can be inferred from their radius. Obviously, this scheme can only be applied if the protein concerned has been refined.

The many developments in computer graphics systems were reviewed by Tony North. These have enabled protein crystallographers to interpret the protein's electron density more accurately and rapidly, and analyse powerfully almost all aspects of the structure. This technology moreover releases them from conventional crystallographic views. It was hoped that those implementing graphics systems would benefit from the discussions on protein representation and that new ideas would be discussed that could simplify and facilitate protein description.

The use of computer graphics in refinement was discussed by Alwyn Jones; he emphasised his principle of trying to make the operation of the system as uncomplicated as possible from the VDU. The development of the system at Uppsala is for the refinement of the satellite Tobacco Nervosis virus (60 subunits in the asymmetric unit) by real space procedures. Both Enrico Stura and Phil Evans described how colour pictures can be produced with filter systems on their Evans and Sutherland display units - and very clear and useful pictures they were.

A simple and cheap approach to computer graphics for use in refinement was discussed by Rod Hubbard. This system was based on a microcomputer which had been programmed to rotate and translate coordinates and to apply MODELFIT. The system is linked to a DEC 10 for storage and Fourier calculations. It runs a colour television set which enhances comparisons and simplifies stereochemical review and repair. The availability of colour enables colour contouring of the electron density and rapid production of differently constructed maps, i.e. $2\rho_o - \rho_c$, $\rho_o - \rho_c$, etc.

Ian Tickle presented the graphical displays he was using on the Evans and Sutherland in his docking programme. The problems of overlapping structures as the binding surfaces approach had not been overcome but even so the image of the approaching and contacting surface was clearly a great advantage. The more serious problem of predicting the conformational changes associated with interaction remains a fundamental hurdle in this field which Ian felt should not be underestimated.

It was a little difficult to gauge the general view on stereo figures and the use of colour. However advantageous stereo is, it is complicated to use with slides and colour is expensive for publication. It would be helpful if journals did use colour more - perhaps there is a role for the Proceedings of the Royal Society with its particularly high standards of figure production.

The matter of publishing protein coordinates came up several times. The obligation to make a newly solved protein's coordinates available as soon as reasonable is an important one. But there are two inhibiting factors. First, there is the wish for thorough exploitation of the structure within the laboratory where the analysis was carried out. Secondly, there is the recognition that initial coordinates are inaccurate and possibly misleading. Since refinement can provide

a satisfactory coordinate set, there are arguments for delaying the release of the atomic positions. This, however, hinders those whose work on such areas as protein structure comparison does not require atomic coordinates but does need as many structures as possible. Michael Rossman's threat to extract rough coordinates from stero figures should at least stimulate us to think seriously (and generously) about the release of coordinates.

Carl Brändén gave the final lecture in which he reviewed the design and structural organisation of the α-β proteins. In this study the structural similarities between the proteins were best seen in figures where the elements secondary structure and their connections were represented very simply in a notation like that described by Mike Sternberg. Survey of ligand binding to this disparate group of proteins revealed that the binding sites were consistently between antiparallel sheets in which the polypeptide chain from each strand was twisted apart. The combination of chain organisation and perhaps dipole effects evidently favoured interactions by various ligands to the protein. Thus we had a thoroughly interesting and timely demonstration of the value of comparison between proteins and the value of simple schematic figures.

THE SERC MICROFILM RECORDER

(K Crennell, Rutherford Appleton Laboratory)

The SERC maintains an Information International Inc (III) FR80 microfilm recorder for use by the research community at the Rutherford Appleton Laboratory. The recorder consists of a precision light source, several cameras, magnetic tape drives and a controlling computer.

The light source is an 80mm diameter, optically flat cathode ray tube, coated with a white phospher, with a raster of 16384 by 16384 addressable points on its surface.

Only one of the four cameras is mounted at a time. Each has a system of lenses to change the image to the appropriate size. In addition, the 35mm camera has a set of three coloured filters, cyan, magenta and yellow, allowing the direct production of colour films, and slides.

Cameras of the FR80

1. Microfiche camera unsprocketed black and white - uses 105mm film, with a 48x reduction of the image.

2. 35mm camera - uses sprocketed film, (black and white or coloured).

3. 16mm high precision camera - uses sprocketed black and white film.

4. Hard-copy paper camera - uses 12" wide rolls of unsprocketed paper up to 400 feet long.

There are special circuits to draw long vectors, plot points and output text at speeds of up to 40,000 characters a second. The intensity of the beam can be set to any one of 256 grey levels, but in practice we find that the emulsions on the films resolve only 32 levels and that on hard-copy paper only 16.

Microfiche form a very compact media for the storage of large quantities of text and pictures. They can be cheaply copied photographically, and sent through the post as easily as a letter since they measure only 6" x 4". They are a reliable, cheap alternative to on-line databases and graphics terminals if you want to have access to a reference set of standard crystal structures drawn from various fixed angles. Computer simulations of photographic images on film make not only a better quality image, but are also much easier to handle than the same images made by overprinting on lineprinter paper. (See figure made using programs from the molecular computer graphics system developed by Mr D Richardson in the laboratory of Dr P Pauling (University College London) modified by A M Lesk (while a visitor at the MRC Laboratory of Molecular Biology Cambridge)).

The Open University has made films with the FR80 for use in their Biochemistry courses. The Cambridge Crystallographic Data Centre uses the hardcopy camera to prepare the camera ready copy for their annual bibliography, 'Molecular Structures and Dimensions, Organic and Organometallic Crystal Structures'.

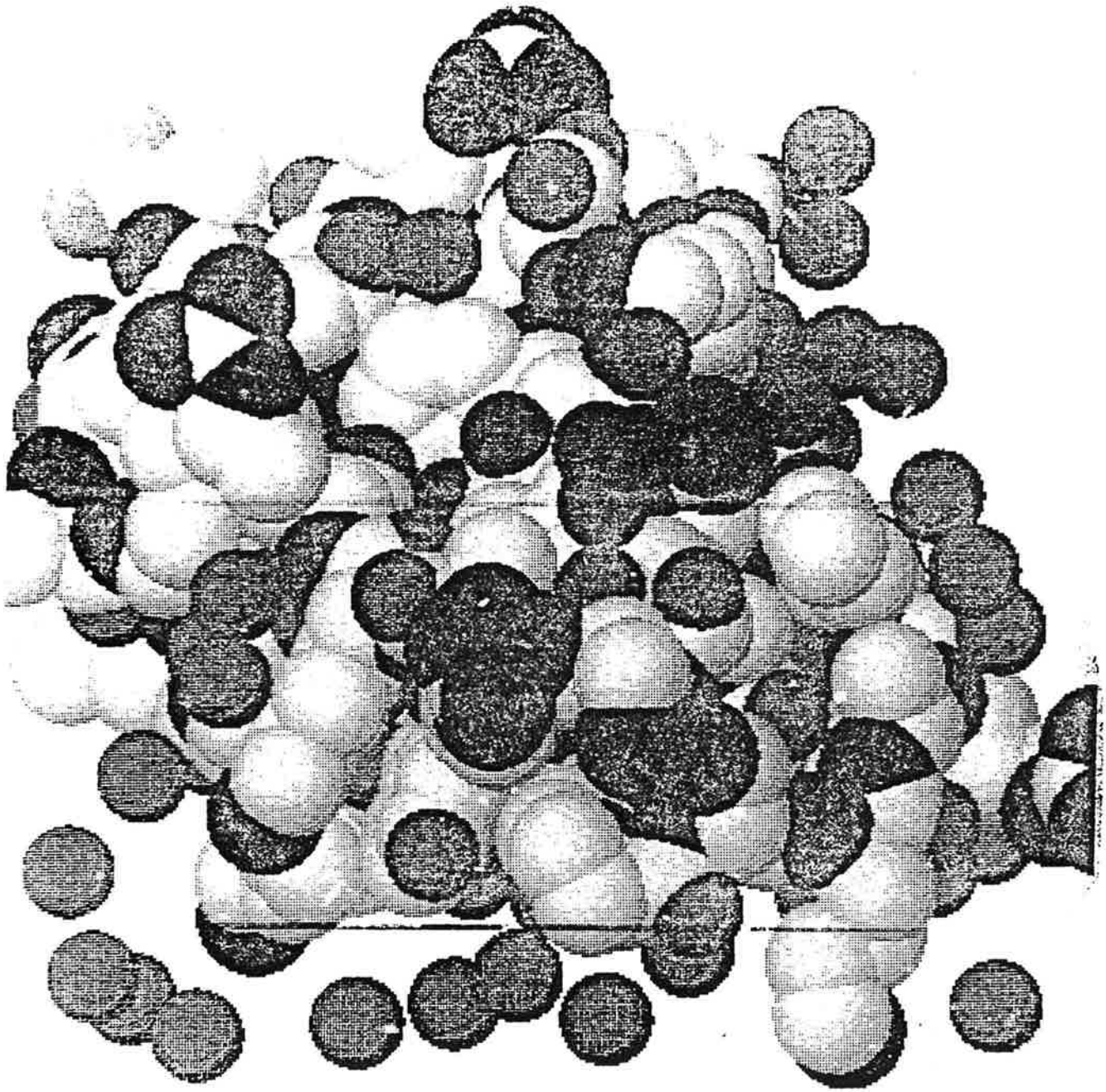
Access to the FR80

Intending users should request FR80 plotting time with their requests for SERC computer time in their grant application. The method of access is normally through the SERC network to the graphics spool of the central computers at the Rutherford Appleton Laboratory. Very large jobs may need to use magnetic tapes

instead. Initial enquiries about access methods should be made through your local Program Advisory Office. In case of difficulties, or if you want to discuss what kinds of diagrams the FR80 is capable of producing contact:

K M Crennell
Atlas Centre
Computing Division
Rutherford Appleton Laboratory
Chilton
Didcot
Oxon
OX11 0QX

Tel: (0235 21900 Ext 6397



Job NWCPKPIH (7Jun82) last frame.

AN IMPROVED STEREO VIEWING DEVICE

O S Mills (Manchester University)

Although some people are able to achieve naked eye stereopsis, the large majority require some form of optical assistance in order to merge two representations of an object and so obtain a three-dimensional effect. The most common of portable devices include coloured spectacles, devices with lenses to focus on the left and right-hand images (the simpler devices have no adjustment whilst others allow the lens separation to be adjusted to match the user's eyes and object separation), and devices with adjustable mirror(s) to facilitate 'merging' of the objects.

These devices need experience in use and even so many people never achieve a successful three-dimensional simulation.

A robust, simple and cheap device is on the market which is proving to be very successful especially with chemists many of whom admit that they now 'see' instantly, and often for the first time, a three-dimensional image. The left/right views are recorded on 35 mm film with a separation of 60 mm. The film may be either negative or positive. No special illumination is required, the average lighting level in a laboratory is quite adequate. Although the film strips can be made with a 35 mm camera, the future lies in using COM devices such as the Benson COM at UMRCC or the FR80 at the Rutherford Appleton Laboratory. If sprocketed devices are used, the nearest approximation to the ideal shift is 13 sprocket holes (= 61.75 mm); 12 sprocket holes (= 57 mm) correspond to three 'single-frame' movements (N.B. 'ordinary' 35 mm photographs are double frame). It is best to centre each image in the centres of the fields of view for each eye. Excellent results have been obtained for ORTEP drawings. The quality usually far exceeds that of drawings produced with pen plotters. The device is also extremely useful with colour pictures which can also be produced on COM devices. Although the viewer permits a field of 25 mm square, and the aspect ratio of a double-frame picture is 1.5:1, good results have been obtained with the GHOST package at UMRCC with a post-processor which uses a limited picture space of 1.46:1. When two molecules are required to be viewed for comparison, the two sets of images can be interlaced to give a small movement of the film.

The device costs only £1 (+ VAT) and is obtainable from:

K P Plastics (Bletchley) Ltd, Ward Road, Bilston Estate, Bletchley, Milton Keynes.

THE INAUGURAL MEETING OF THE U.K. MOLECULAR GRAPHICS GROUP

ANDY MORFFEW
IBM UK Scientific Centre
Athelstan House
Winchester, SO23 9DR

On the 16 September 1981 at Balls Park, Herts. (that is its name, honest) a very interesting meeting was being held by the A.I.O.P.I. (Assoc. of Information Officers to the Pharmaceutical Industry). Three of us, David White (Glasgow), Andy Vinter (Wellcome Research) and myself, were discussing usual conference topics, price of drinks, quality of food etc. The conversation came round to the number of conferences held by different groups, that we attend to either speak or hear about molecular graphics. Andy Vinter proposed that we formed a group on molecular graphics to bring together all those interested.

The suggestion was well-timed. Many groups are acquiring graphics systems or updating old hardware. Valuable lessons have been learnt from pioneering software systems and everyone realises that what is required is a platform for communicating all of this data. The wheels of administration were set in motion and the kick off meeting was planned for January 8th.

Now, this date may well set off some alarm in your memory cells. Friday the 8th of January was the day that Terry Wogan announced that Wales was closed until further notice and nobody believed him. Half of Britain was covered in a foot of snow, leather shoes were being ruined and pet owners had the difficult task of persuading Fido and Felix to venture into snow that was up to their ears. As I drove into Leeds (from the York meeting) I envisaged that maybe ten out of the ninety delegates may actually make it and that we would all have to eat nine roast beef dinners followed by chocolate gateaux and gallons of coffee.

Amazingly these fears were ill-founded as only fourteen of the delegates were finally unable to attend. The first part of the meeting set up the new group - The U.K. Molecular Graphics Group, setting the membership fees at £10 for this year (£4 for bona fide students).

The next event is the major meeting for 1982 and will be a three day meeting 22-24 March. It is called The Molecular Graphics Workshop and time will be given to overseas guest speakers, U.K.M.G.G. speakers (short papers), small group meetings and a hardware exhibition.

A committee was ratified for 1982 consisting of:-

M.S.G. Clark	(Beecham, Harlow)
D.J. Gilman	(ICI, Alderley Edge)
A.J. Morffew	(IBM, Winchester)
A.C.T. North	(Leeds University)
W.G. Richards	(Oxford University)
J.G. Vinter	(Wellcome, Beckenham)
D.N.J. White	(Glasgow University)

Then came the entertainment. Tony North gave an interesting review of molecular graphics covering both its development and some current work.

Tony Thomas (Roche Products) described their system, built on a Megatek 7000, and its use in drug design.

Tom Blundell described the molecular graphics programs that have been developed at Birkbeck. (I think this work is very exciting and the sheer magnitude and breadth of the software development shows that Tom's group is one of the leaders in the field.)

David White talked about energy minimisation algorithms and about the advantages of using a parallel processor.

Due to a shortage of time, I missed out my talk and instead distributed a newly prepared bibliography on molecular graphics.

The meeting then broke up with people slipping and sliding to the station or following snow ploughs to the M1.

I would like to thank all those who helped to arrange this meeting, especially Sandy Geddes, Leeds.

In 1982, I look forward to the development of an active group and I look forward to meeting those of you interested in any form of molecular graphics at the Workshop in March.

SCIENCE AND ENGINEERING RESEARCH COUNCIL

DARESBURY LABORATORY

POLARIZATION CORRECTION FOR THE SRS PROTEIN CRYSTALLOGRAPHY WORKSTATION

P.J. Duke and J.R. Helliwell

The polarization correction factor, P for an oscillation camera and a single triangular monochromator crystal on a synchrotron source, is of the form $P = P_o (1 - \tau^1 \alpha)$ where τ^1 is a source and optical system dependent parameter (we follow the notation here of Fourme and Kahn who derived the relation in 1979). For the SRS (at 1.8 GeV) protein crystallography station, Ge(111) monochromator (no mirror) τ^1 is 0.9523 at 1.488 Å and 0.9500 at 1.608 Å. We have also calculated with our program the same values as are accepted for the DORIS and DCI set-ups (which gives us some measure of confidence !). When you collect data on the mirror + monochromator system on X-ray 7 at the SRS τ^1 will be quite a different value (i.e. beam less polarised). Watch this space.

AN EDITOR FOR 'LCF' REFLECTION DATA FILES

John W Campbell (Daresbury Laboratory)

Introduction

This note describes the program 'EDITLCF' which has been written for editing/ inspecting the reflection data in an LCF format reflection data file. It may be used either in an interactive mode or in a batch mode. The reflection data are normally processed in sequence in a single pass through the file using a sequence of local edit commands. During the pass, a limited number of reflections are held at any stage in an on-line buffer and the user may move up and down the reflections which are currently in the buffer. Also, if required, the user may complete the current pass through the file and prepare to start another pass through the file by using the 'TOP' command. Global edit commands which make a complete pass through the file are available. The edit session is completed using the 'END' command.

Format of an Edit Command

In general, the format of an edit command is as follows:

Function code Reflection list Values list

Function Code : This defines the type of edit to be performed, e.g. change, insert, delete, list, etc.

Reflection List : This defines the reflection or classes of reflection on which the edit is to be performed. Reflections are defined by giving the indices and/or a list of items with values or value ranges of the form label(i) or label(i,j) ('label' is a label from the LCF file, 'i' is an integer value and 'i' and 'j' are minimum and maximum integer values).

e.g. 2 3 10

or H(1,5) k(0) L(1,5) CRYST(1)

or 5 6 8 CRYST(2,3)

When indices are given explicitly, they must be given immediately after the function code (i.e. in fields, 2, 3 and 4 of the command). The labelled items may be in any order.

Value List : This defines the new values to be used in conjunction with a change, scale or insert command. It consists of a list of labelled values of the form label=x where 'x' is the value to be assigned to the column labelled 'label' (or the scale factor to be applied for a scale command).

e.g. FOBS=120 SIGFO=15 PHI=-168

Values for scale factors may be real, e.g. FOBS=1.12

An edit command may be continued on the next line by putting a dash (minus sign) at the end of the current line of the command. The command may be spread over as many lines as required.

Types of Edit Command

There are basically three types of edit command available, local edit commands, global edit commands and special edit mode commands. In the descriptions below 'refln' or 'reflns' indicates a reflection list and 'value' indicates a value list as described above. 'n' is the number of reflections to be operated on, and 'm' indicates an operation on every 'm'th reflection.

Local Edit Commands : These operate on a single reflection or a requested number of reflections in sequence starting with a given reflection. The available functions are outlined below:

Function	Command format(s)	
Find reflection	F refln	
List reflection(s)	L refln	(Ln Ln/m)
Print (hard copy)	P refln	(Pn Pn/m)
Expanded list with labels	X refln	(Xn)
Change values	C refln values	(Cn)
Scale values	S refln values	(Sn)
Delete reflection(s)	D refln	(Dn)
Insert after reflection	I refln	
Move down buffer	+n	
Move up buffer	-n	

A blank command is equivalent to the command +1 i.e. moves one reflection down the file.

The reflection list for a local edit command should define a single reflection. If the definition is ambiguous, the first reflection encountered, which fits the given specification, will be taken. If the reflection list is omitted, the current reflection is assumed. If the reflection list is present the command is equivalent to a Find followed by an edit on the current reflection.

Global Edit Commands: These commands cause a complete pass through the file with edits being performed on the reflections specified by the reflection list. If the reflection list is omitted, all reflections will be assumed. The available functions are outlined below:

Function	Command format
List reflections	GL reflns
Print (hard copy)	GP reflns
Change values	GC reflns values
Scale values	GS reflns values
Delete reflections	GD reflns

Edit Mode Commands: These form a miscellaneous set of commands controlling the way in which the edit program operates. The 'END' command, which is used to complete the edit session, is included in this category. Each edit mode command has its own format. The functions available are as follows:

Function	Command format
Return to top of file	TOP
Complete the edit session	END
Select items for listing	SELECT label1 label2 ...
Hard copy output	PRINT ON/OFF
Verify edits	VERIFY ON/OFF
Non-standard index names	HKL labelH labelK labelL

Additions by Phil Evans

Phil Evans has implemented a version of the editing program on the VAX at Cambridge. He has made the following additions to the program and it is intended that these will be incorporated in the Daresbury version.

1. As an alternative to defining a range of values when specifying an item in a reflection list, a list of values, separated by slashes, may be given.

   e.g. BATCH(1/15/16/25)

2. An extra edit command to 'delete all but' the specified reflections has been added.

3. Provision has been made for handling LCF files containing batch title records.

Status of the Program at the Daresbury Laboratory

The program EDITLCF is available for use using the CLIST PCA.CLIST(EDITLCF) (batch mode) or PCA.CLIST(EDITLCF$) (interactive mode). Program documentation is available in manuscript form and is being word processed. When the word-processing is complete the program etc will be transferred to PCZ and will be included officially in the CCP program suite.

BROOKHAVEN DATA BANK STRUCTURE FACTOR TAPES

Pella Machin (SERC, Daresbury Laboratory)

The atomic coordinate holdings of the protein data bank have been available at Daresbury for nearly two years, and they are regularly updated. Details of their whereabouts have been described in a previous newsletter (at DL see PCZ.PROTDIR.DATA).

It was suggested to me that some people might also like access to the structure factor tapes. The structure factor entries do not all fit on one 1600 bpi magnetic tape so Brookhaven issue firstly NONST1TP containing the "older" structure factor entries which are not expected to change and secondly NONST2TP which will grow as new entries are added. I have requested and received a copy of NONST1TP and this is now available at Daresbury on magnetic tape 74SB13 (NL, 1600 bpi, blksize = 4800).

If anyone wants further information on the format or content of the file let me know.

I can also obtain copies of the new structure factor entries NONST2TP if required. I will not however take any action on this till I have some requests for it.

COMPUTER GRAPHICS AT BIRKBECK

Garry Taylor (Birkbeck College)

Since its installation in 1979 the computer graphics system at Birkbeck (based on a PDP 11/60 with 128K memory and an Evans and Sutherland Picture System II with 64K memory) has been exhaustively used both by members of the department and by visiting scientists. The system is used at least 12 hours per day, most weekends and frequently through the night; the hardware being more durable than the users!

Most of the time is spent interpreting electron density maps or model building using either Dr R Diamond's BILDER or Dr A Jones' FRODO, although the latter is used more often as its philosophy lends itself well to the department's particular problems. Proteins interpreted include γ-crystallin II, endothia pepsin, ribonuclease, glucagon, avian pancreatic polypeptide (aPP) and Ao-glu insulin by Birkbeck users; tRNA synthetase by Dr Bhat; catalase by Dr W Melik-Adamyam from Moscow and a carboxypeptidase by Dr O Diddeberg from Belgium. In addition FRODO has been used to model build β-crystallin, insulin like growth factors (IGF's), relaxin, bovine pancreatic polypeptide, various insulins, chymosin, pro-chymosin, pepsin and pepsinogen.

Not all of the time is spent using BILDER or FRODO however, as several new programs have been written for the Evans and Sutherland for investigating surface representations, molecular docking, comparisons of structures etc. All of these programs utilise the capabilities of the picture system to rotate, translate, create perspective window and depth cue under interactive control. Following is a brief description of each of these programs.

MIDAS

MIDAS (Molecular Interactive Display program with Atom Selection), written by Ian Tickle, was initially a program for displaying regions of interest in molecular models, but more recently has been extended by him to include an algorithm which allows the docking of two rigid molecular surfaces.

The program has the facility to display any selection of atoms, with or without labels, according to name, atom type, residue type, residue hydrophobicity, etc or any combination of these. This is effected by the use of a flexible syntax to describe subsets of atoms, e.g. *, !/NEUT means all charged residues while */PHOB,/HIS,15A:19,165 means all hydrophobic residues, histidines and residues

15A to 19 inclusive plus residue 165. Contacts (e.g. H-bonds) between selected atom types can also be displayed (Fig. 1 - which shows the main chain and hydrophobic side chains of aPP with every 5th C_α labelled and with main chain H-bonding shown).

Molecular surfaces can be displayed using an algorithm, SURGEN, developed by Laurance Pearl in the department. The surface is represented by dots generated on a 0.5Å grid from the atomic van der Waal's spheres. This is considerably simpler and faster to compute than Connolly's algorithm used in Dr R Langridge's laboratory, but can lead to slightly 'square' atoms.

The docking algorithm of MIDAS uses an automatic one-dimensional search along a specified trajectory; pairwise interatomic contacts are searched for and are sorted according to increased displacement from the starting position. Then the computer brings the docking molecules together for each contact in turn and recomputes the pairwise contact distances, this time giving the spheres a softness dependent on their hydrogen bonding ability or allowable flexibility; finally the best contact situation is displayed which gives the maximum intercalation with the maximum buried surface area.

Figure 2 shows van der Waal surface representations of the α-helix and proline helix (separated) of aPP showing the complimentary hydrophobic surfaces which pack together, and which can be made to do so using the docking algorithm.

MIDAS with its FRODO like menu and psuedo analogue to digital converters for rotations, etc, compliments FRODO extremely well as a tool for understanding the interpreted molecular model. Ian has also incorporated a tracker ball into the system which allows manual rotation about any axis.

BILBO

This program with a ring of anthropomorphism about its name has been written by Anne-marie Honegger from the Biochemisches Institut, Zurich who has been a visitor with us for the last six months.

Bilbo is designed for the study of protein-protein interaction such as dimerisations, sub unit interactions and receptor binding of polypeptide hormones, as well as for the comparison of protein surface topology, charge distribution, etc.
Van der Waal and solvent accessible surfaces are generated in a similar way to

SURGEN, using bit mapping. These dot-surface representations can be superpositioned using a least squares best fit algorithm (e.g. Dr A Maclachlan's EASIFIT) and different logical operations (inclusive or, exclusive or/and) can be performed between their bit (dot) maps to highlight similarities and differences in shape and volume. Accessible surface areas of the whole protein or of individual residues can be listed on demand.

Surfaces can be displayed as clouds of dots, but for a stationary picture in order to obtain a better sense of depth a grid can be dropped over the surface of the molecule (Fig. 3) or surfaces can be represented as a furry cloud of vectors perpendicular to the atomic surface at each bitmap point (Fig 4) - both of these latter representations use much more memory. Different surface textures can be chosen to highlight hydrophobic etc., residues and H-bond donors and acceptors (Fig 4 shows the hydrophobic surfaces 'furred' and hydrophilic dotted).

For docking, so far only a simple hard sphere docking routine is available (Fig 5), and one can 'slab' through the interacting surfaces to look for details.

FITZ

FITZ, written by Garry Taylor as a tool for the manipulation and comparison of molecular structures has three main functions:

(a) Up to four objects (Cα backbones, atomic models, dot surfaces) can be displayed and independently moved allowing comparison of structures, simple solid body docking or the simultaneous display of several views of the same object. Ian Tickle sweated through the non-trivial task of writing the software to allow rotations and translations relative to fixed screen axes and in FITZ this is generalised for n independent objects. An object can also be 'cut' into parts to allow, e.g. independent movement of protein domains. Finally Dr A Maclachlan's EASIFIT algorithm has been incorporated to allow the best fit between two objects to be obtained on the screen; equivalent pairs of atoms are either explicitly given or are taken automatically by proximity after an initial visual best fit on the screen. Figure 6 shows a comparison of the Cα backbones of endothia pepsin and penecillopepsin before and after fitting, based on a fit between three catalytically important residues' Cα atoms which were deemed to have the same spatial relationships in both proteins.

(b) Crystal packing can be investigated by the display of a unit cell box and the generation of symmetry related molecules whose centres of gravity fall within specified volume limits. This facility has been useful as an adjunct to several molecular replacement problems in the laboratory and the program has an option allowing the input of the orientational parameters of a molecule and then filling a unit cell with symmetry related molecules. The centre of gravity of the molecules can be changed via the pen and tablet and the related molecules move in real time giving a rapid means of investigating sensible packing arrangements (Fig 7).

(c) A 'balsa wood' model can be rapidly built up in 3-D using the tablet as a drawing pad, entering contours from the electron density sections. This was found to be useful in the interpretation of a low resolution MIR map of mucor chymosin. Using the ability to manipulate several objects, the best visual fit of the C_α backbone of a structurally related enzyme endothiapepsin into the molecular envelope of mucor chymosin was found - thus solving the molecular replacement problem in real space (Fig 8).

DOCKER

Dr Bernard Busetta from Bordeaux wrote a program to investigate the binding of a substrate to an enzyme. The program allows complete flexibility of the polypeptide substrate and flexibility of the side chains of the active site. A molecule can be docked into the active site, close contacts being displayed when they occur, until a satisfactory fit has been obtained. This configuration then acts as the starting point for energy minimisation using simple 6-12 non bonded potentials, the resulting conformation of the system being displayed. Solvent accessibility calculations can also be performed.

POLY et al

John Quinn in the group interested in water structure has used the E & S as a very effective display tool. Figure 9 shows a three-dimensional energy plot of the molecular interaction for drivers, and the program has made it possible to examine simply and efficiently large regions of the 'phase space' available to the molecules interacting with a given potential.

The E & S has also been used for examining different infra red spectra by entering the graphical data via the tablet and this has proved to be the simplest way of digitising the spectra for subsequent analysis.

A molecular dynamics display program has been written which shows the configuration of single centre Lennard Jones particles on the screen while the simulation runs on the 11/60. The simulation can then be followed visually and was written for spotting 'bottlenecks' in the simulation or for observing phenomena which take place within small timescales.

FUTURE ?

We are hoping to obtain a colour screen in the near future which will provide an exciting new dimension in the use of molecular computer graphics. Colour can be used to highlight regions of molecular models in obvious ways and will clearly show overlap when docking or comparing structures when e.g. red and green overlap to give yellow. All of the programs described will gain significantly from the use of colour.

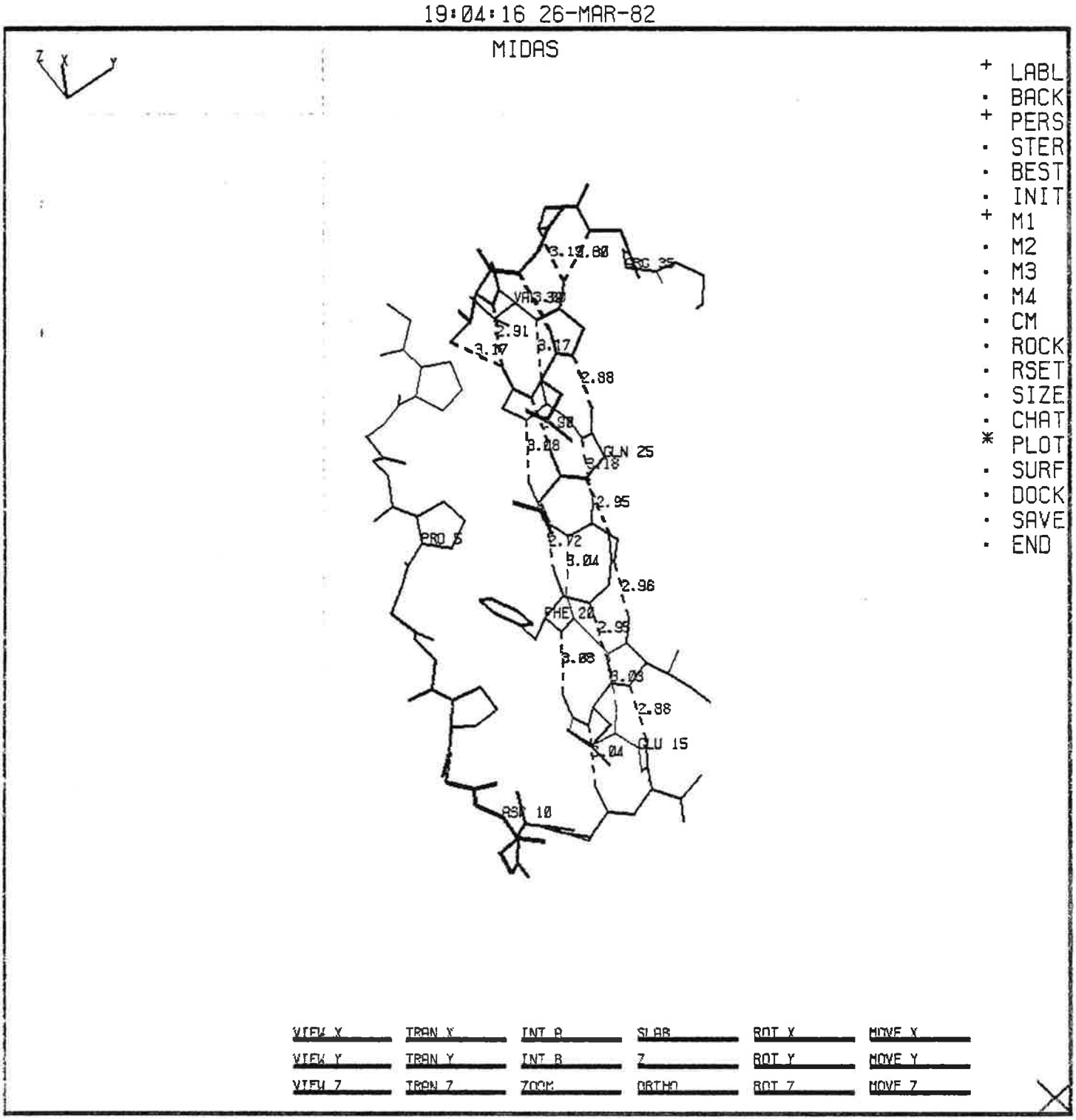

19:04:16 26-MAR-82
MIDAS
+ LABL
· BACK
+ PERS
· STER
· BEST
· INIT
+ M1
· M2
· M3
· M4
· CM
· ROCK
· RSET
· SIZE
· CHAT
* PLOT
· SURF
· DOCK
· SAVE
· END
GLN 25
PRO 5
PHE 20
GLU 15
ASP 10
VIEW X
TRAN X
INT A
SLAB
ROT X
MOVE X
VIEW Y
TRAN Y
INT B
Z
ROT Y
MOVE Y
VIEW Z
TRAN Z
ZOOM
ORTHO
ROT Z
MOVE Z

Fig 1.

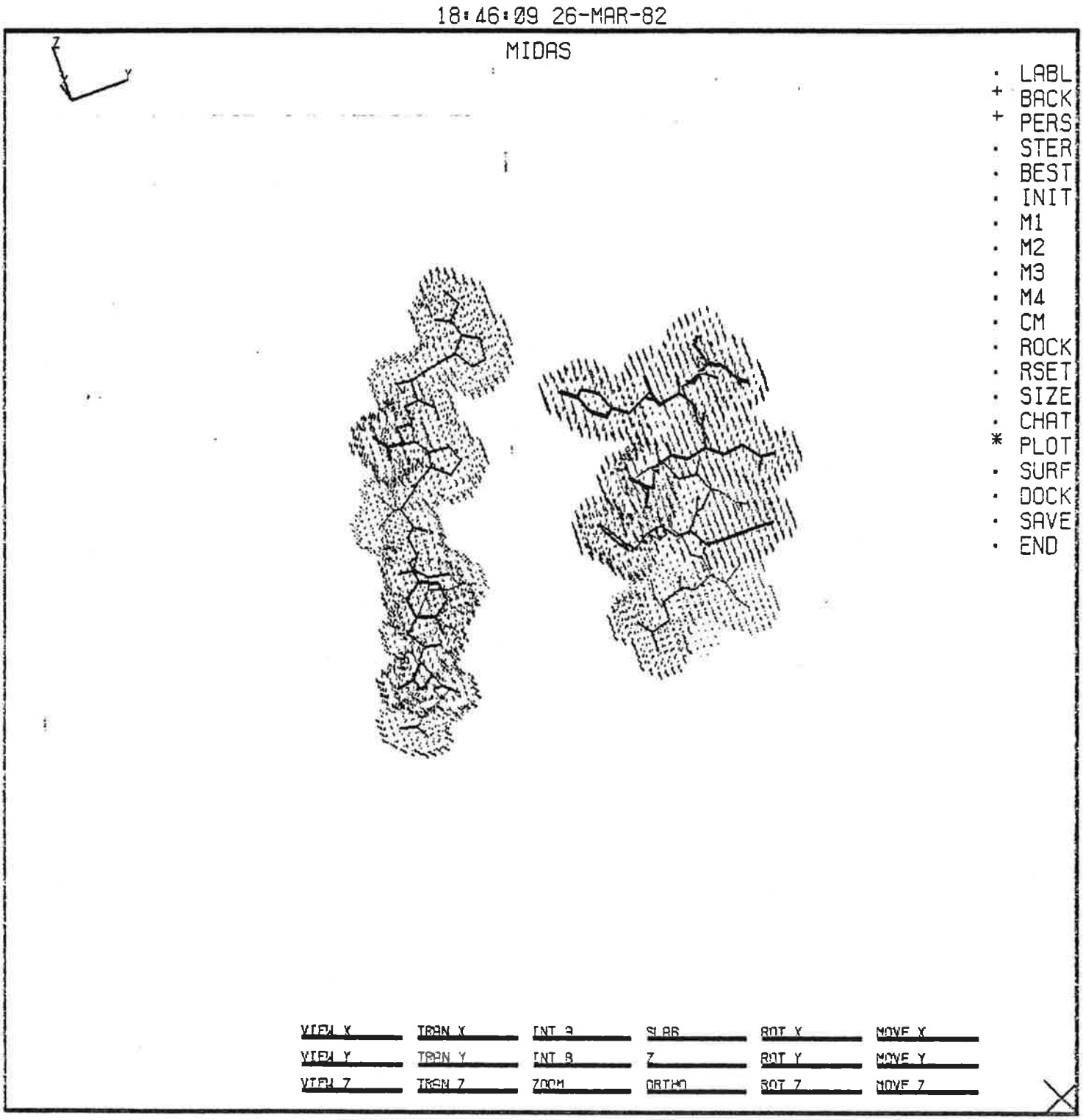

18:46:09 26-MAR-82
MIDAS
Z
Y
· LABL
+ BACK
+ PERS
· STER
· BEST
· INIT
· M1
· M2
· M3
· M4
· CM
· ROCK
· RSET
· SIZE
· CHAT
* PLOT
· SURF
· DOCK
· SAVE
· END
VIEW X
TRAN X
INT 3
SLAB
ROT X
MOVE X
VIEW Y
TRAN Y
INT 8
Z
ROT Y
MOVE Y
VIEW Z
TRAN Z
ZOOM
ORTHO
ROT Z
MOVE Z

Fig 2.

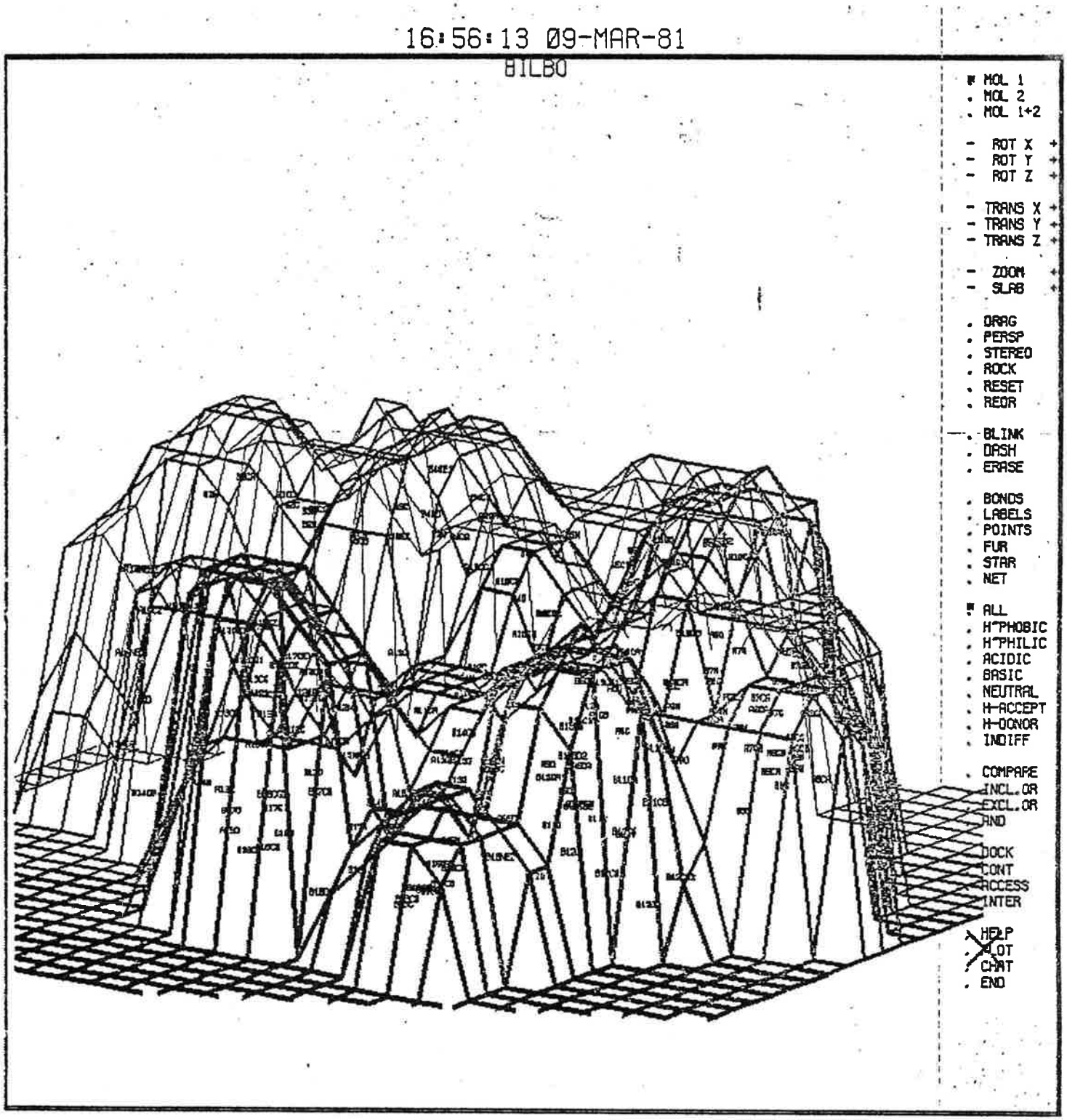

16:56:13 09-MAR-81
BILBO
MOL 1
MOL 2
MOL 1+2
ROT X
ROT Y
ROT Z
TRANS X
TRANS Y
TRANS Z
ZOOM
SLAB
DRAG
PERSP
STEREO
ROCK
RESET
REOR
BLINK
DASH
ERASE
BONDS
LABELS
POINTS
FUR
STAR
NET
ALL
H^PHOBIC
H^PHILIC
ACIDIC
BASIC
NEUTRAL
H-ACCEPT
H-DONOR
INDIFF
COMPARE
INCL.OR
EXCL.OR
AND
DOCK
CONT
ACCESS
INTER
HELP
PLOT
CHAT
END

Fig 3.

Hydrophobic surfaces showed as "fur", hydrophilic as points

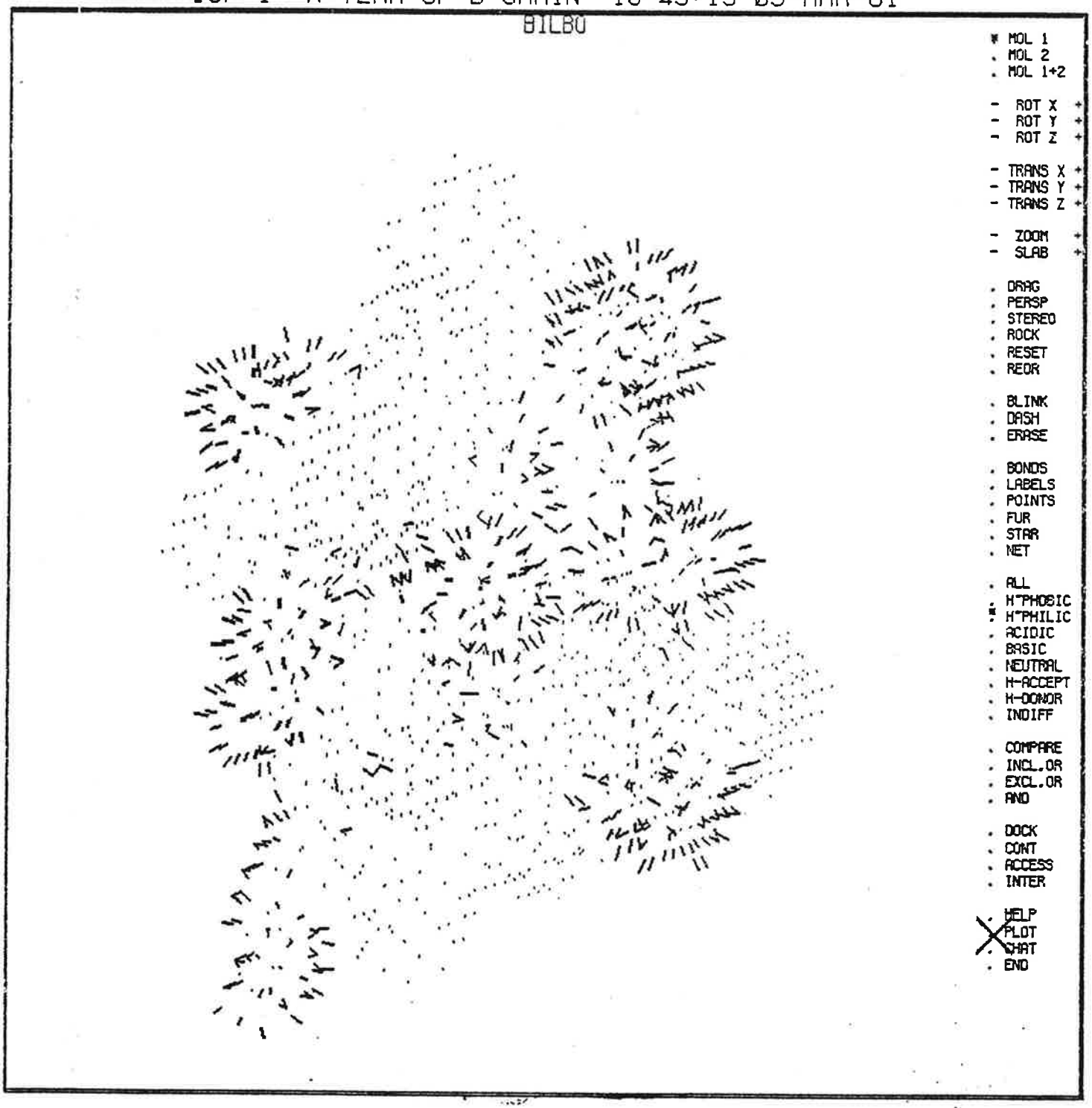


Fig 4.

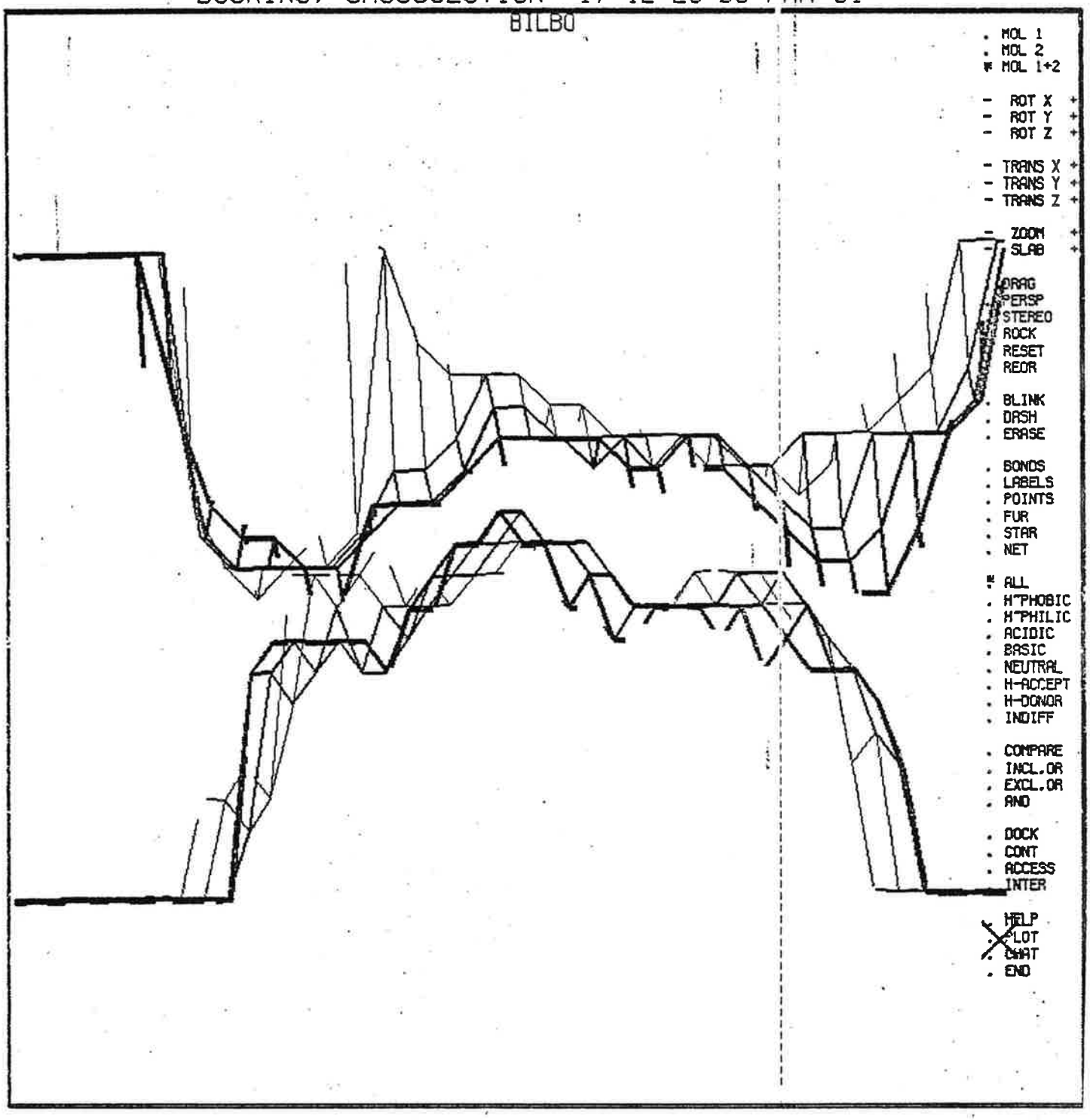

DOCKING, CROSSSECTION 17:12:25 09-MAR-81
BILBO
. MOL 1
. MOL 2
* MOL 1+2
- ROT X +
- ROT Y +
- ROT Z +
- TRANS X +
- TRANS Y +
- TRANS Z +
- ZOOM +
- SLAB +
DRAG
PERSP
STEREO
ROCK
RESET
REOR
. BLINK
. DASH
. ERASE
. BONDS
. LABELS
. POINTS
. FUR
. STAR
. NET
* ALL
. H"PHOBIC
. H"PHILIC
. ACIDIC
. BASIC
. NEUTRAL
. H-ACCEPT
. H-DONOR
. INDIFF
. COMPARE
. INCL.OR
. EXCL.OR
. AND
. DOCK
. CONT
. ACCESS
INTER
HELP
PLOT
CHAT
. END

Fig 5

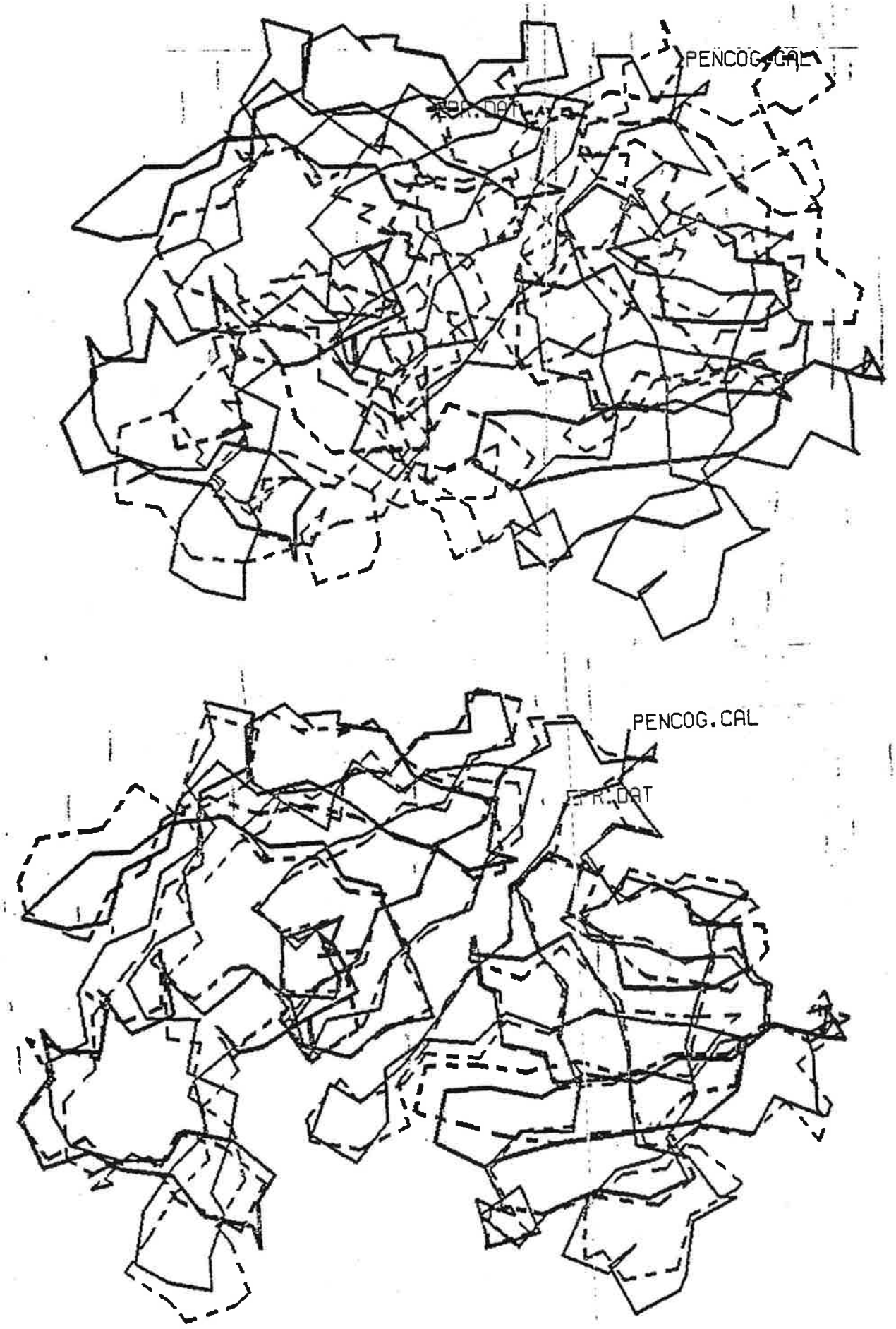


Fig 6.

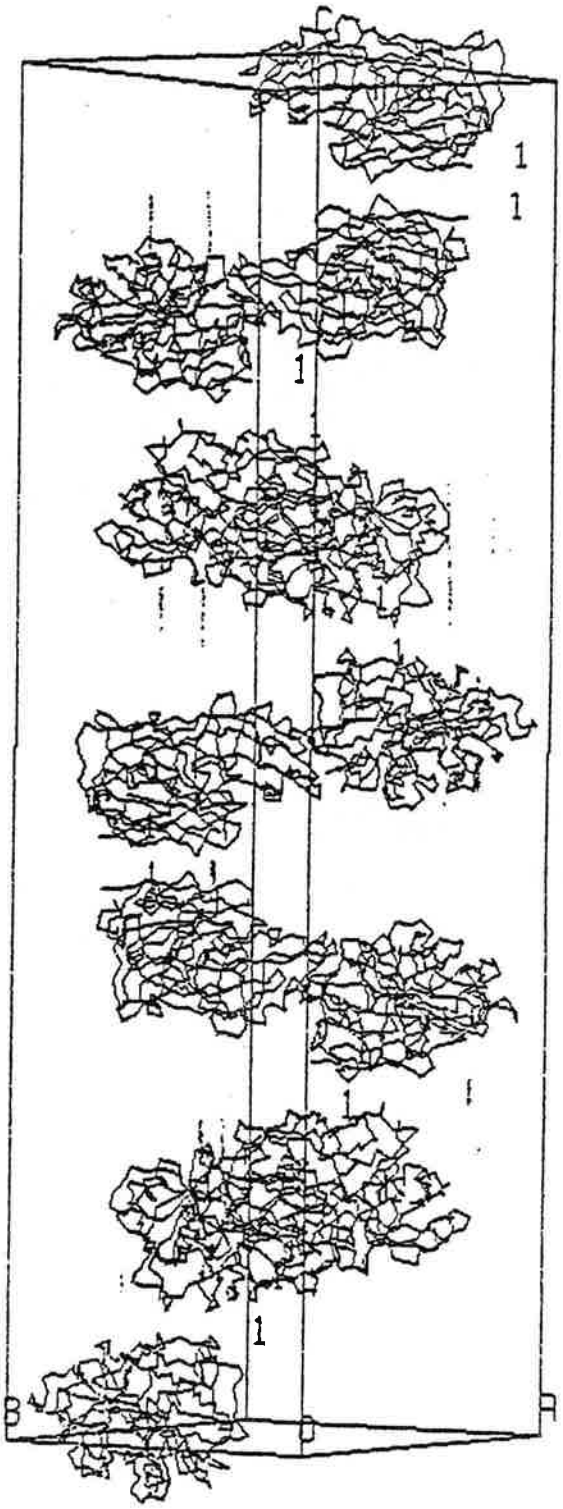


Fig 7.

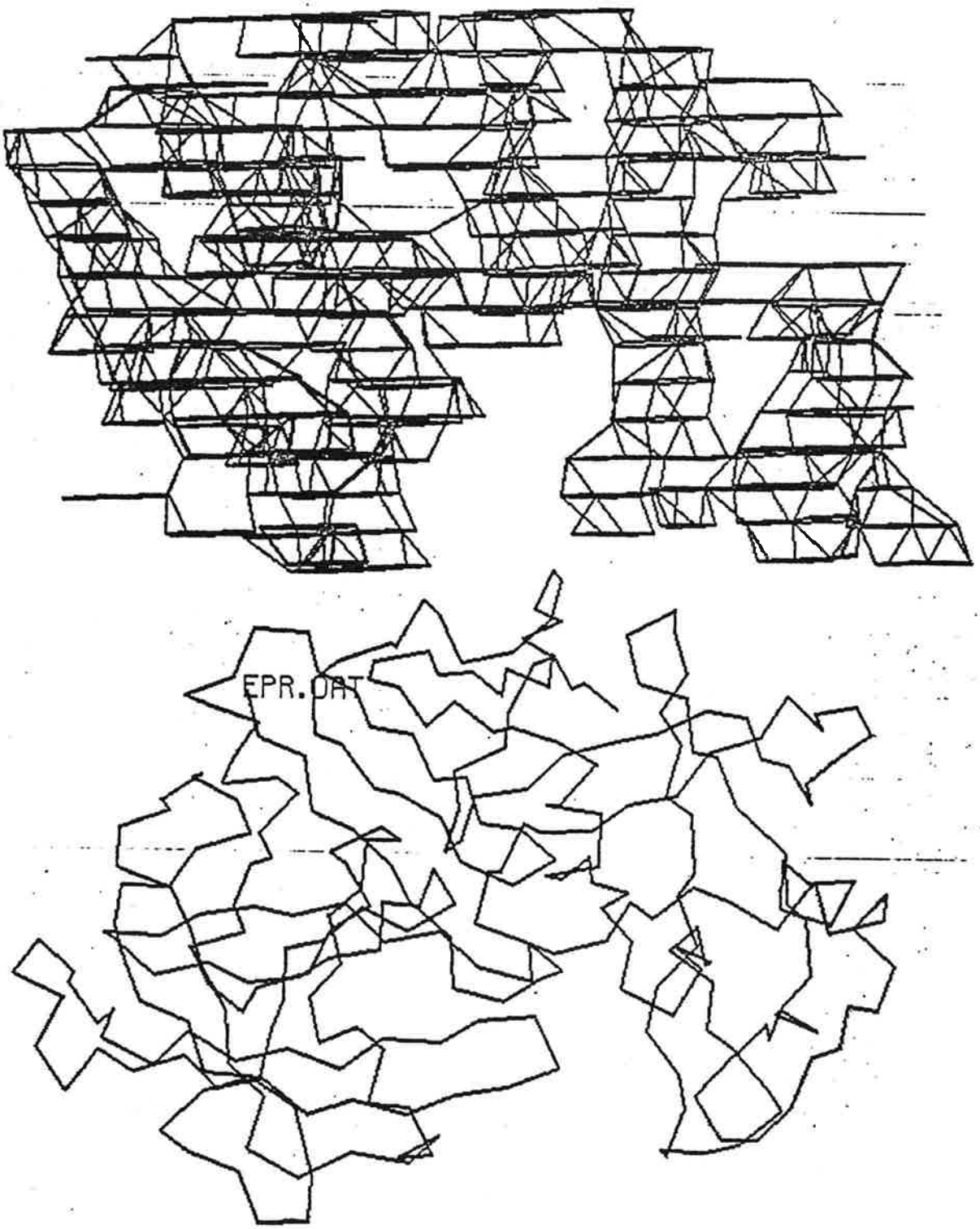


Fig 8.

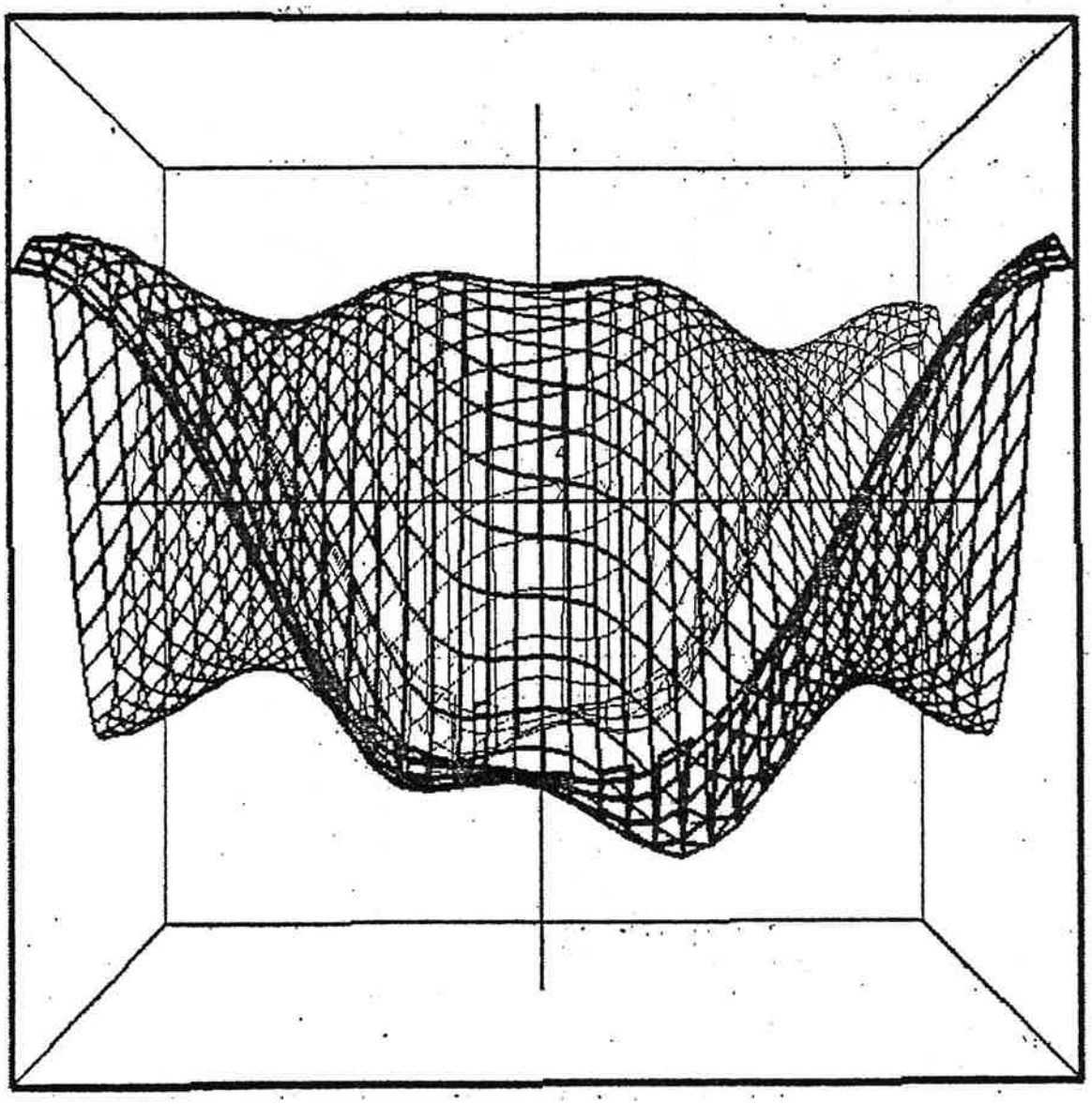


Fig 9.