FRODO

USER'S GUIDE

Electron Density Fitting

And Modeling at UCSD

Updated: November 1985

UCSD Silicon Graphics Iris version:

Stuart Oatley
Department of Chemistry, B-017
University of California, San Diego

Original program:

Alwyn Jones
Wallenberg Laboratory
University of Uppsala, Sweden

NAME

prepmap - electron density map reformatting program for FRODO

SYNOPSIS

prepmap

DESCRIPTION

The PREPMAP program reads a standard formatted electron density map and writes a direct access file whose records are 'bricks' of electron density, i.e. 8 * 8 * 8 volume elements in crystal space. This output file is the DSN6 dataset used by FRODO. The map sections may be of up to 200 * 200 grid points.

The PREPMAP program prompts for:

1. Name of the input electron density map file.

2. Cell dimensions and number of grid intervals.

3. Name of the output brick file.

SEE ALSO

Descriptions of various aspects of use of the FRODO system may be found in these -me format files in directory doc : chat.me, menu.me, sam.me, dict.me, newfile.me

AUTHOR

Alwyn Jones, Wallenberg Laboratory, University of Uppsala, Sweden.
Stuart Oatley, University of California, San Diego

NAME

contour - electron density map contouring program for FRODO

SYNOPSIS

contour

DESCRIPTION

The CONTOUR program reads a standard formatted electron density map and contours it in bricks, i.e., volume elements in crystal space, which are typically between 4 and 10 map grid units along each unit cell axis direction. The brick size along an axis must be an exact divisor of the number of grid points along that axis.

When a new box to be contoured is selected in FRODO, the program first decides which bricks lie wholly or partly within the box, and these bricks only are searched for contours within the box. Hence large bricks incur the overhead of extra time searching for contours which turn out to be outside the box. On the other hand, small bricks produce short contour circuits; each circuit requires 6 + (3 x no. of vertices) words. Thus for short circuits the memory requirement approaches 3 words per vertex + 9 words per circuit.

The CONTOUR program prompts for:

1. Name of the contour file to be written; n.b. these files may well be large!

2. Scale to be applied to the input electron density values (usually 1.0).

3. Brick size on X, Y and Z, the cell axes.

4. Contouring flags for U, V and W, the map axes: U = columns, V = rows, W = columns (usually UVW; terminate with END).

5. Contour level; the program then returns to prompt 4. The FRODO CHAT command CCOM is used to select which contour levels are to be displayed and the line type to be used for each level.

6. Name of the input electron density map file.

SEE ALSO

Descriptions of various aspects of use of the FRODO system may be found in these -me format files in directory doc :
chat.me, menu.me, sam.me, dict.me, newfile.me

AUTHOR

CONTOUR(1G) Silicon Graphics (local) CONTOUR(1G)

Alwyn Jones, Wallenberg Laboratory, University of Uppsala, Sweden.
Stuart Oatley, University of California, San Diego

NAME

frodo - electron density fitting program

SYNOPSIS

frodo

DESCRIPTION

Before FRODO can be run, all the datasets used by the program must exist. Two of these, the names and dictionary datasets, are public; some, e.g., the coordinate and the map datasets, will have been created by other programs; and some are scratch files, which can be obtained by copying existing versions.

The constants dataset contains the names of all the other datasets as well as various constants that the user may set in CHAT mode and which are retained from one run of the program to the next. When staring a new project, make a new copy of file, and use the editor to update at least the DSN1 and DSN2 names contained in it; the other names can be changed in FRODO itself.

The DSN1 dataset contains the current contour (or background) object. This is always required since the current scale and screen centre are stored in it.

The DSNA dataset may contain a second contour object for use with the SWAP and COBA functions.

The DSN2 dataset is in direct access binary format and contains status flags and atom identifiers (atom name + residue number) for the complete protein sequence, whether the coordinates are known or not. It also may contain the coordinates, and is created by the program SAM.

The DSN3 dataset is a contour map, as output by the CONTOUR program.

The DSN4 dataset is the dictionary containing the standard bond lengths and angles for model refinement. The dataset lib/dict should be used unless there are good reasons for making a private version.

The DSN5 dataset is a scratch dataset required for model refinement.

The DSN6 dataset is a direct access dataset containing the electron density map, as output by the PREPMAP program.

The DSN7 dataset is a backup of the main coordinate dataset. It should initially be a copy of the DSN2 dataset.

The DSN9 dataset contains standard residue and atom names. The dataset lib/names should be used unless you have any special requirements.

When running FRODO, the first prompt will request the name of the constants dataset. The next prompt will then be the '?' of the CHAT interface.

SEE ALSO

Descriptions of various aspects of use of the FRODO system may be found in these -me format files in directory doc : chat.me, menu.me, sam.me, dict.me, newfile.me

AUTHOR

Alwyn Jones, Wallenberg Laboratory, University of Uppsala, Sweden.
Stuart Oatley, University of California, San Diego

NAME

sam - coordinate dataset manipulation program for FRODO

SYNOPSIS

sam

DESCRIPTION

SAM exists as a standalone program and as an exit from the main graphics routine in FRODO. Its function is to manipulate the DSN2 dataset in various ways. It is used, among other things, to create an initial empty dataset, to fill it with data from various external formats and to list the atoms or sequence in the dataset.

When running the standalone program, the user is prompted for the name of the FRODO coordinate dataset to be used, and then for the name of the names dataset. When running FRODO, these prompts are not issued since the dataset names are already known (DSN2 and DSN9).

SEE ALSO

Descriptions of various aspects of use of the FRODO system may be found in these -me format files in directory doc :
chat.me, menu.me, sam.me, dict.me, newfile.me

AUTHOR

Alwyn Jones, Wallenberg Laboratory, University of Uppsala, Sweden.
Stuart Oatley, University of California, San Diego

CHAT INTERFACE COMMANDS

The CHAT interface produces a '?' on the terminal. The user can then type a four character command. An unrecognised keyword causes a listing of the current commands (as does HELP). All terminal input is in free format, i.e., numeric and character strings are separated by at least one space or a comma, but cannot be carried over to a new line. Numeric items do not require a trailing decimal point, ie, 1, 1. and 1.0 are all the same. Commas can be used to indicate numeric items whose values are to be unchanged, e.g. ,,,1 will reset the fourth item to 1, leaving the first 3 unchanged. Character string items include atom names, residue names (also referred to as residue types) (eg, GLY, TRP, HG, SOL) and residue numbers (eg, 1, -5, 101A, A50, XYZ1 etc).

When using FRODO for the first time you must give sensible values for ZONE, ACON and RZON. If contours are to be displayed from a vector DSN3 dataset, there must be suitable values for CCOM and CRAD.

Data set definition commands :

DSN1

The name of the 'background object' data set which is a description of the linear display list usually representing contours.

DSN2

The name of the direct access coordinate data set. The first record, the 'controller', is listed on the terminal (see the 'newfile' writeup for a description).

DSN3

The name of the data set of vectors usually made by pre-contouring a Fourier map at a number of levels with the program CONTOUR. This could also be a vectorised atomic model, where the C-COM's represent various models obtained in the course of a refinement, for example. Note that this by itself does not cause the new contour map to be read. Unless one of the keywords which initiate contouring (viz. BACK, SCAL, CCOM, CONT, ACON or CRAD) is also given, the old contour object will still be displayed.

DSN4

The name of the dictionary data set used by the regulariser (see the 'dictionary' writeup for a brief description).

DSN5

The name of the scratch data set used in refinement.

DSN6

The name of the direct access dataset containing the electron density map in 'bricks', as produced by the program PREPMAP.

DSN7

The name of the direct access coordinate data set used as the backup/restore data set (see the .SAVE description).

DSN9

The name of the 'names' data set used by SAM commands ABINITIO, INSERT, REPLACE. It is a source data set consisting of a residue type and number of atoms in the residue (format (A4,I5)) and then a list of the atom names (format (18A4)). The atomic order in this list becomes the order in the data set. There is no atomic type definition in this file, so that users who wish to leave the display with a Diamond data set must use SAM command ATOMTYPE to insert this in the DSN2 coordinate file (see SAM writeup).

DSNA

The name of the second DSN1 type data set used in conjunction with SWAP.

The following commands define which part of the DSN2 data set gets displayed (note that at present up to 2000 atoms can be handled):

ZONE

The prompt requires the alphanumeric names of 2 residues, specifying the limits of the display zone; the first must be before the second in the linear sequence. The connectivity is made on a residue to residue basis and with ZLNK connections. This is normally used at the start when working with the initial isomorphous map, later one normally works in SPHEre mode. The atoms can be picked from disk more quickly in ZONE mode then any of the other modes. SYM has no affect when in this mode.

SPHE(re)

This is the normal working mode. The prompt asks for a sphere radius (e.g. 7A). All atoms in the data set within this distance from the current screen centre are displayed. If an atom is not connected to any other atom on the screen, it appears as a 3-D cross. The user must be careful when modifying large groups of atoms on the edge of the screen since large discontinuities can be introduced in the coordinates in the DSN2 data set, e.g. if the user applies .FBRT to a piece of a lysine side chain where the NZ atom is outside the sphere radius, then the CE-NZ distance could be very large.

The program decides on which atoms are in the sphere by first checking if the residue centre is in the sphere, in which case it checks all the atoms in the residue.

MIX

A maximum of 10 zones plus a sphere can be defined. The program continues to prompt for residue names until the user types END. It then prompts for the sphere radius (if the user types 0.0 no sphere search is carried out).

SMIX

As MIX but the joining routine has been modified so that hydrogen, iodine and sulphur atoms are checked by name and then have different bonding distances (1.3A for H-X, 2.2A for I-X and 2.5A for S-S). Methyl groups containing hydrogens do not look like little pyramids with this option. The algorithm checks for atoms starting with the characters H, I and S.

For ZONE, SPHE, MIX and SMIX, if no input is given, then the parameters required will default to the last values specified.

N.b. - ZONE, SPHERE and MIX are mutually exclusive. ZONE mode is generally used only in the initial stages of model fitting. Unless LINK is specified, atoms with a residue are bonded if <1.9A apart, and the bond linking one residue to the next is always made. SPHEre or MIX mode in essential in later stages to check for inter-residue and inter-molecular contacts. Bonds are made between all pairs of atoms <1.9A apart (unless LINK is specified).

The following commands are relevant to producing background display lists:

BACK

Allows the user to display a background. This assumes the existence of a DSN3 data set.

NOBA(ck)

This stops the displaying of a background object. There is no need for a DSN3 data set.

SCAL(e)

Input the picture scale; this should normally be set to 1.0, but can be changed in case of overflow.

CONT

The prompt expects the coordinate to use as the next screen centre for use in SPHEre. If only return is typed, the centre defaults to the position of the last atom identified on the screen.

ACON

The prompt expects a residue and atom name to define the next centre. Omitting the atom name forces CA as default. Error messages warn if the residue or atom does not exist in the current DSN2 data set.

After the above two commands, typing GO will cause recontouring around the new centre if 'BACK' is on.

VECT

Requests the background object to be selected from the vector DSN3 dataset. If previously MAP was selected, the program asks whether both should be used to construct the background.

MAP

Requests the background object to be calculated from the map DSN6 dataset. Prompts are issued for the size of the box to be contoured, in A about the picture centre; the contour levels and their associated line types (see CCOM) and (if more than one level) colours; and the number of directions in which to contour, usually three. These contour colours are stored separately from that for a single level background object (see COL). The maximum number of grid points allowed in the box is 650 in each section, and 16000 in total. If previously

VECT was selected, the program asks whether both should be used to construct the background.

CRAD

This command is only relevant if VECT has been set. The prompt expects values for the contour radius and cover radius.

During the selection of the atoms to be displayed, the program keeps track of the extreme values of the coordinates in each orthogonal axis direction. The size of the box defined by the extreme values is then increased all round by the cover radius (e.g. 1A, to ensure that the atoms near the edges of the box are covered with contours. The volume of the resultant box is compared with 8*(r**3) where r is the contour radius (e.g. 6A). If it is greater then the box volume is reduced to 8*(r**3) by first reducing the greatest extreme (from the screen centre) to the 2nd greatest, then the 2nd to the 3rd, the 1st to the 3rd and so on. Since one should normally be working in SPHERE mode, one would normally choose a contour radius smaller then the sphere radius, in which case the cover radius is unimportant.

The electron density map on DSN3 is pre-contoured in bricks, i.e., volume elements in crystal space typically between 4 and 10 map grid units along each unit cell axis direction. Only whole bricks are allowed. The program first decides which bricks lie wholly or partly within the box, and these bricks only are searched for contours within the box. Hence large bricks incur the overhead of extra time searching for contours which turn out to be outside the box. On the other hand, small bricks produce short contour circuits; each circuit requires 6 + (3 * no. of vertices) words. Thus for short circuits the memory requirement approaches 3 words per vertex + 9 words per circuit. The contours found inside the box are copied to DSN1.

CCOM

This command is only relevant if VECT has been set. The vector DSN3 data set may contain data for a number of different contour levels. Each level is called a C-com, and this command is used to request which ones should be used and what their line types and colours should be. For vectorised coordinates, each C-com corresponds to a particular model and a number of these can be requested with this command, for example to follow the course of a refinement.

The program prompts for a line of C-coms and then for a line of associated line types. Normally only one contour level will be displayed, in which case either the default colour or that supplied

for the background via COL will be used. If more than one contour is requested, the program also prompts here for their colours; these do not affect that selected for the background in COL. Line types are:
0 = solid line (default);
1 = dashed line;
2 = dots at vertices;
3 = blinking;
4 = blinking dashed;
5 = blinking dots.
The last three options are not currently available on the Iris 2400 system.

SWAP

This makes the data set defined by DSNA the new DSN1 data set and vice-versa, i.e. we get what was in the DSNA data set as a new background object.

COBA

This enables a composite background object to be made from both the DSN1 and DSNA datasets. The DSNA object will be shifted to the screen centre of the DSN1 object, though overflow problems may occur if these are greatly different. The colours used for the DSNA object will be those selected when the object was created. The option can be cancelled by a 'NO' response to the prompt.

The following commands are to do with the regulariser options:

MODE

The program prompts for one of three possible modes:

1. MODL: To carry out normal regularisation. If an atom has the magic coordinate (1500.,1500.,1500.) given by SAM to signify that it is not determined, an attempt will be made to place the atom, given the standard dictionary bond lengths, angles and torsion angles (which may be set by ANGL). To be successful the algorithm requires at least 3 connected atoms. For example, if SAM has been used to REPLACE a residue of type ALA for one of type PHE, then one requires one pass through the regulariser to pick up the ring and the CG.

2. EXTD: To extend the chain from the first residue in the RZON. This option does no regularisation, being a model making option. The atoms in the rest of the RZON are added to the first residue using the connectivity, and bond lengths and angles defined in the dictionary

data set, DSN4. By using the ANGL option one can define values for torsion angles to make any desired conformation.

3. RESI: To build a residue or zone of residues from scratch. It does not require coordinates for the first residue in the RZON. The algorithm searches for 3 connected atoms in the dictionary, it then puts the first atom at the current screen centre, the second along the X-axis and the third in the X-Y plane. The remaining atoms are added as with EXTD.

RZON

Input one or two character string items defining the residue number limits for the regularisation. There is a maximum of 150 atoms.

NCYC

Input the number of cycles of refinement. The method of minimisation requires a number of passes through the RZON. During each pass the atoms are shifted to satisfy their local conditions. After each cycle there is a listing on the terminal defining the largest shift from the starting coordinates, the largest error in bond lengths, bond angles and torsion angles found in the last cycle. One also gets an internal count for which atom had the error. 10-20 cycles seem to give reasonable results. Use 0 to build in standard conformation.

SHIF

This defines the maximum allowed shift during each cycle. If the program calculates a larger shift during refinement, then the applied shift is in the calculated direction but is equal to the value defined by SHIF. Suggested values are 1.0A when using the 'explosion' method of building an initial model, or when starting from an MIR map; otherwise 0.3-0.5A for refined structures.

MWT

The regularisation requires 4 starting weights and 4 final weights. The weights correspond to the shift from the starting coordinates, the bond length error, the bond angle error and the torsion angle error. During each cycle, the weights get updated according to an exponential function, starting with the initial values and finishing with the final values. The program prompts for the 4 starting values (usually use 1., 1., 1., 1.), and then for the the 4 final values (Alwyn Jones suggests 1., 30., 20., 10.; Ian Tickle

suggests 1., 10., 20., 1.). These give high initial weights to the input coordinates, but as the cycle number increases, other factors become more important. If refining badly distorted structures, a lower weight to the input coordinates is better; perhaps initial weights of 1., 100., 100., 100. and final weights of 1., 200., 200., 200..

ADD

This option must be used with care. The positions of all dummy atoms (IA=2,4) are not used in the refinement, but are calculated from the dictionary connectivity, bond lengths and angles. It is neccesary to have 3 connected atoms somewhere in the RZON. This option is not required any more, having been made redundant by the MODE options.

ANCH

This option applies an anchor at the ends of the RZON to maintain connectivity along the chain. In fact it applies fixes to atoms named N, CA, and CB in the first residue and CA, C and O in the last residue of the RZON. It is therefore only useful with protein molecules (other types must be anchored with explicit .FIX).

DIST

This allows one, when the .REFI option is selected on the screen menu, to insert a maximum of 10 distance restraints between atoms in the RZON and other atoms. The program prompts for the residue name and atom name of the moving atom, and then for the names for the fixed atom. The regulariser treats the distance of the first atom (whose position may change during the regularisation) to the position defined by the second atom before regularisation (i.e. it does not change during the regularisation) in the same way as a bond restraint on the first atom. If both atoms are in the refinement zone, the constraint is still to the initial position of the fixed atom.

ANGL

This allows one, when the .REFI option is selected on the screen menu, to guide special conformations during regularisation or model making. It can be used for example to force a piece of chain to adopt the phi and psi angles of an alpha helix during the regularisation or can set free normally 'fixed' torsion angles (e.g. omega angles). If this option is on, before entering the regulariser the user gets a message, requesting a four character option name. Typing END causes the program to continue.

Other options are:

1. ALPHA: To define the phi and psi angles for an alpha helix. The program then prompts for a zone of residues over which to apply the torsional restraint and for a weight. The weight is the fraction of the fixed torsion angle weight at any time in the regularisation. So typing 1.0 ensures that the normally flexible phi and psi angles are treated as fixed torsion angles.

The next 3 options have the same prompt.

2. ABET: To define anti-parallel pleated sheet angles

3. PBET: To define parallel pleated sheet angles

4. PROL: To define polyproline dihedral angles.

The next 3 options have the same prompt.

5. PHI: The prompt is for a residue range, a value and a weight. The value refers to the value to which the phi angle should be restrained.

6. PSI: Similarly for a zone of psi angles.

By using PHI and PSI the user can, for example, define main chain angles for 3-10 helices.

7. OMEG: To allow the user to set free the omega angle for high resolution structures.

8. TAU: To allow the user to set free the tau angle (defined N-CA-C).

9. Unrecognised options are assumed to refer to torsion angles. An atom must be defined which has 3 back chained atoms in the dictionary description of the residue. These 4 atoms define the torsion angle. The prompt asks for residue, atom, value and weight. As an example, one would define a CHI1 angle by defining the CG and a CHI2 angle by defining the CD atom.

NOAD

Switches off the ADD option; all atoms, whether real or dummy are used in refinement.

NOAN

Switches off the ANCH option.

NODI

Switches off the DIST option.

NANG

Switches off the ANGL option.

REAL

All atoms are made real after regularisation.

STAY

There is no change in the atomic status after regularisation (usual).

The remaining options are:

GO

To exit from the CHAT interface. The coordinates to be displayed are loaded from the DSN2 dataset; if BACK is on and a new picture centre has been specified, a new contour object is computed. The present system values get written to the control data set.

HELP

This causes a listing of the presently allowed commands (as does an Unrecognised command).

NOW

This causes a listing on the terminal of the present values of the dataset names and the important program constants.

DBP

Specifies whether the display will operate in single buffered or double buffered mode. Input is 1 for single, and 2 for double. This will normally be left set to 2 on the Iris 2400 system.

MUVA

This is used to move atoms to new postions. The program prompts for residue and atom names, and then for the new coordinates. This continues until the user presses just a return at the first prompt. If the atom name is not typed, the program defaults to CA. The new coordinate values get written to the DSN2 data set. If the user has a large list of data, it may be more convenient to use the SAM option UPDATE which can also define the IA flag.

LINK

This is used to look at a data set made of just CA atoms. The connectivity is defined so that the first atom is connected to the second, the second to the third etc..

NOLI(nk)

This is the normal option. The connectivity is then formed by various distance criteria.

ROCK

Input the rocking amplitude (degrees) and the rocking speed (degrees/frame). Suggested values are 5 and 0.5. Rocking is initiated by the ROCK screen menu option.

SYMM

This forces the program to check if symmetry related atoms occur within the viewing sphere. The DSN2 data set must have been updated with SAM option SYMMETRY.

NOSY(m)

This switches off the symmetry checking.

ZLNK

When in ZONE mode the program forms a bond between a named atom in one residue and another in the next residue. These two atoms are defined with ZLNK. The program first prompts for the atom in the i'th residue and then the (i+1)'th residue. For proteins, these are C and N.

SAVE

It is only necessary if you have altered any screen coordinates and forgotten to use the .SAVE screen menu option. It does the same

thing as .SAVE, i.e. it writes the current screen coordinates (but not symmetry generated atoms) to the DSN2 data set.

AXES

This option causes a set of axes to be drawn in the top right corner of the screen. It is the recommended option.

NOAX

No axes are drawn on the screen.

NAYB

This sets the contact distance criterion (e.g. 3.5A) for the .NAYB screen menu option.

TITL

Input a title (max. 50 chars) to be displayed at the bottom of the screen, above the pseudo A-D's; the date and time are appended to the input title. This can be useful for identifying slides taken from the screen and is also included on output produced by the .PLOT screen menu option. If the PF1 key is toggles off, the screen menu, pseudo A-D's and text scratch area are not displayed and the title moves to the bottom of the screen. The title can be removed by typing just carriage return in response to the prompt.

RIGI

This allows one to apply a rotation /translation operation to a zone of residues which best fits the current DSN2 coordinates to a set of guide coordinates. The least squares algorithm was written by Jim Remington. The program prompts for residue and atom name, and then for the guide position. Input is terminated by typing END. A maximum of 100 atoms can be used in the matching process. If the user does not wish to apply the transformation, he should answer the prompt TO WHAT RESIDUE RANGE IS IT APPLIED ? with END.

RSR

Input the real-space refinement parameters. The program prints the current values and then prompts whether an update is required. Supplying a zero value for some parameters will give you a default value. The first prompt is for the number of cycles of refinement to be performed (default 1; maximum 5), and their corresponding translational search values. The second prompt is for the number of

nearest grid points over which the grid sum is to be performed (default 2), the atomic radius to be used in obtaining the calculated electron density (default 1.5) and the starting scale factor between the observed and calculated electron densities. The third prompt is for the atomic numbers corresponding to the various atom types defined by the atom type record (IEND(3)) in the DSN2 dataset. This gets set to 0 or 1 for all atoms if the file has been created from PDB or Hendrickson format coordinates, so you may wish to update your file with the SAM option ATOMTYPE before selecting either .RSR, .RSFX or .RSCL on the screen menu; otherwise give all atoms a compromise value of 7 here.

COL

Allows one to alter the default colours of the various objects displayed on the screen. Reply 0 to the prompt to get a list of objects and colours on the screen. Object numbers are:

1. Text
2. ID's
3. Background
4. Moving atom
5. Neighbours
6. Fbrt
7. Tor
8. Axes
9. Molecule
10. Fbrt ID's
11. Dots/Stars
12. Box, Menu
13. Title, A-D's
14. Menu hi-lite
15. Cursor
16. Display background (normally black)
17. Menu background
18. Menu hi-lite background

SAM

Enters the SAM interface for manipulation of the DSN2 dataset.

REFI

Enters the REFI model regularization routines (see the screen MENU description). These do not need the picture system active, but in this case they cannot be terminated by pressing the middle button of the mouse.

MOL

Enters the MOL routines for making a background object from the DSN2 data set.

SYS

Allows the user to issue a Unix shell command from the terminal.

END

Saves the constants dataset and terminates the program.

FRODO DICTIONARY DESCRIPTION

The dictionary used by FRODO is described by Hermans & Ferro in Biopolymers, Vol. 10, pp. 1121-1138. The connectivity is represented as a tree structure of backward and forward pointers. Bond lengths and angles are defined via the back pointers. The full dictionary of Jan Hermans also has the possibility of defining non-bonded atom energy coefficients.

The first record is a title.

The next record defines the residue type and the number of atoms in the dictionary description of the residue. The format is (A4,6X,I5). A residue name END terminates the file.

Each atom has one record defining the name, backward chain, lowest forward chain, highest forward chain, a flag to define fixed or flexible torsion angle, a bond length, a bond angle, a torsion angle. The format is (A4,2X,4I2,3F5.0).

THE ATOM NAME

The atom name contains 4 characters. Names with X or Y in them have special significance and are used, especially, to close rings. These special atoms act as ghosts of the atom which is named without the X or Y, and they will have the same position as the real atom during regularisation.

THE BACKWARD CHAIN

The chain values act as pointers to atoms defined in the dictionary. For example, the value 1 signifies the first atom in the residue being described, and 5 would be the fifth. The backward chain represents a connection from the present atom (call it the i'th) to an earlier atom in the list (call it the h'th).

THE LOWEST FORWARD CHAIN

It signifies a connection between the present atom and a later atom in the list (call it the j'th). Then $i < j$ and there must be no backward chain connections to atom i between the i'th and j'th atoms in the residue.

THE HIGHEST FORWARD CHAIN

It signifies a connection between the present atom and a later atom in the list (call it the k'th). Then there can be no backward chains to i from atoms in the list beyond k. There can be other connections to the i'th atom from atoms in the list between positions j and k, but these will be described by backward chains.

THE BOND LENGTH

It is the value for the bond defined by the backward chain from this atom.

THE BOND ANGLE

It is the value of the bond angle defined by 2 backward chains from this atom.

THE TORSION ANGLE

It is the value of the torsion angle defined by 3 backward chains from this atom.

THE TORSION ANGLE FLAG

It takes the value 0 for fixed dihedrals and greater then 0 for flexible ones. The angle concerned is best described with an example :
The angle PHI for an amino acid is defined by the atoms C, N, CA, C. It is flexible and its flag is carried around with atom CA. Likewise PSI is defined as flexible at atom C and OMEGA at atom N.

THE COMPLICATION

This comes for atoms with more then one forward chain. The highest forward chain atom is straight forward. The torsion angle value held at the lowest forward chain atom now takes on a special meaning and defines how much this atom is out of the plane defined by h, i and k as defined above. In particular, if we hold h towards us and look down the i to k bond, it is the angle made by line hi and ij in projection. It is quite easy to get this value wrong by 180 degrees (or wrong sign) when making a new type of residue.

WARNING

This is meant as a help to someone who has to make a new entry in the dictionary. It is not a full description. I advise you to look at the existing entries and draw out how they have been built up. With a little practice it is not difficult (honest!).

The distributed version of FRODO skips over all atoms begining with H.

NEW FRODO FILE DESCRIPTION

An EMPTY FILE is :
1. First record
2. Empty sequence records with residues name '$$$$' for all ('blank sequence' record).
3. Extra information, '****' ('blank extra' record)
4. No cooordinate data

A First Record shall contain:('the controller')
1. Number of sequence entries
2. Number of extra information entries
3. Record number of next free entry, for coordinate data
4. Number of sequence entries presently used; shall include first residue named '????' in record 2, i.e. in an 'empty file' it is 1
5. Number of 'extra entries' in use
6. Number of residues presently defined
7 - 12. Not used
13 - 15. Year, day, month when the data set was last used with FRODO

A FULL FILE is
1. All sequence records are used or
2. All extra records are used
The file is useable but cannot be extended

An 'IN-USE-FILE' shall contain
1. A correct first record - 'the controller'
2. An empty '$$$$' in record 2
3. At least record 3 shall not be a blank sequence record
4. Extra entries may be blank

A blank sequence record shall contain just '$$$$'

A blank extra record shall contain just '****'

A non-blank sequence record shall contain
1. Residue name 4 bytes, right justified
2. Residue type 4 bytes, left justified
3. Start and end record of the coordinate data for this residue (2 integers, each 2 bytes)
4. Spherical centre and radius of residue each as floating point variables 16 bytes

A Residue is
1. A logical grouping of atoms.
2. The atoms within a residue may of may not be covalently bound.

3. For molecules made up of a sequence of smaller units (such as the case in proteins and nucleic acids), sequence records should follow the molecular sequence. This is only a requirement when external computer programs (e.g regularisation of a zone by the start and end residue, defines that the intermediate residue in the molecular sequence should also exist in file sequence.

Extra records are defined by 4 byte key. Their content will depend on the key.
1. LATT and one byte describing lattice type, e.g. P
2. CELL unit cell of the structure in 6 floating point numbers.
3. SYMM and 12 integers representing one crystallographic symmetry operation. The first 9 numbers define the rotation matrix a1-a9, the last 3 define shift vector b1-3 in units of 1/24 of the unit cell.
e.g. SYMM a1, a2, a3, a4, a5, a6, a7, a8, a9, b1, b2, b3
If x, y, z represent the fractional unit cell coordinates, then symmetry related
sym x= a1*x + a2*y + a3*z + b1
sym y= a4*x + a5*y + a6*z + b2
sym z= a7*x + a8*y + a9*z + b3
There is a requirement that the operations represent space group symmetry.

An atomic record shall contain the following information.
1. IA, 2 byte integer, used as a status flag. At present has following values:
1 = non-dummy atom
2 = dummy atom
3 = fixed non-dummy atom
4 = fixed dummy atom
A non-dummy atom is one having a non-zero weight in structure factor calculations. A dummy atom has zero weight in structure factor calculations. This status is also meaningful in regularisation using the 'ADD' option.
A fixed atom is not allowed to move in regularisation.

2. IB, a 2 byte integer pointing to a record of another atom in the data set. It is only meaningful in regularisation; then if it is non-zero, the indexed atom is treated as if it were in the residue but has the character 'Z' added to the atom name.
e.g. a disulphide has non-zero IB, the IB for each CA points to the CA record ofthe other residue and likewise for CB and SG. The regularisation treats a CYS residue as one containing the atoms, N, C, O, CA, CB, SG, CAZ, CBZ, SGZ, where the 'Z' atoms are from the other residue name.

3. Atom name, 4 bytes, left justified.

4. Orthogonal atomic coordinates, 3 floating point numbers.

5. 2 byte integers for atom class (normally C=1, N=2, O=3, S=4, P=5), temperature factor scaled by 100 (BMAX = 327.67), structure factor weight scaled by 100 when the data set was made.

SAM INTERFACE

SAM exists as a standalone program and as an exit from the main graphics routine in FRODO. Its function is to manipulate the DSN2 dataset in various ways. It is used, among other things, to create an initial empty dataset, to fill it with data from various external formats and to list the atoms or sequence in the dataset.

When running the stand alone program, the user is prompted for the name of the FRODO random access coordinate dataset to be used, and then for the name of the 'names' dataset (this must be defined if one uses REPLACE or INSERT). When running FRODO, these prompts are not issued since the dataset names are already known (DSN2 and DSN9).

Terminal input is free format (but limited to fit on one line) unless otherwise explicitly stated.

The present commands are:

EMPTY

This is to allocate a fixed amount of space for the controller record, the sequence records and the extra records. Once made they cannot be expanded. The program allocates one record for the controller, and then prompts for how many sequence records are to be allocated and then for how many extra records are to be allocated.

The number of sequence records must be greater then the number of residues in the structure (say, number + 50). Every time the user does an INSERT, a sequence record gets used up. When one does a DELETE, the free sequence record will not be used later by an INSERT; REPLACE uses the existing record. If there are n sequence records, then the user should define at least (n+4)/5 extra records. If the user wants to define SYMMETRY, there must also be space for (nsymop+2) more records, where nsymop is the number of symmetry operations.

ABINITIO

This allows one to generate from scratch a file with atomic records corresponding to a given sequence. The program prompts as for EMPTY and then asks for the dataset containing the sequence. This dataset must contain the residue types, left justified with reference to this type in the 'names' dataset. The format is (18A4). If a residue type is not defined in the 'names' dataset, an error message is printed, but the program continues checking through the full

sequence. However the file is of no use. One record is generated for each atom; the coordinate is set as 1500., 1500., 1500., and it is of type dummy (IA=2). If an atom has this magic number in its coordinates field, it means that its coordinates are still undefined.

There are limits of 1000 residues in the sequence, 50 residues in the 'names' file, and 1000 atoms in total in the 'names' file.

READDIAM

To read in the modified form of Diamond's Real Space Refinement coordinates. The format of the Diamond file is :

1. First record contains the angular cell constants (30X,3F10.0).

2. A title record.

3. A heading record.

4. Atomic records, one atom per record. Diamond's non-atomic records are not used. Each record contains the coordinates in A along the cell edge, a temperature factor, an atom type (IDW in Diamond's nomenclature - the record is skipped if this values is less than or equal to 0), an internal residue count (NO in Diamonds nomenclature - not stored), a weight, a residue type, a residue name and an atomic name. The format is (4F10.0,3I5,2A4,T76,A4). Input will be terminated by an IDW greater than 0 and a NO less than 0.

This option will also read Diamond's standard list if the first, normally blank line, is replaced by the unit cell angles (format (30X,3F10.0)).

This option overwrites the current contents of the dataset DSN2. Before issuing this option, an empty dataset must be created either by EMPTY or ABINITIO. Once the command has been issued, any number of Diamond datasets can be combined into one new DSN2 dataset. The first prompt issued by the program is for the name of the Diamond dataset; typing END terminates the option. The program lists the angles found in the dataset and the title. It then prompts for 2 entry numbers, and only atoms between these 2 entries are added to the DSN2 dataset. The program lists on the terminal each residue name, its type, the disc address for the coords for that residue and the pseudo sphere centre and radius for the residue (this is a standard WRITE statement in subroutine PUTSEQ and is called when any program updates the DSN2 sequence records). The program returns to the first prompt when it is finished.

READWH

This option allows one to read in any formatted list of coordinates. The I/O list specifies the residue type, residue name, atom name, X, Y, Z and temperature factor. If this is not the order in the file, remember the use of T format which allows one to set the position on the record from which to read (e.g. if we have a formatted file X, Y, Z, B, residue type, residue name, atom name with format (4F10.3,3A4), we can define a new format (T31,3A4,T1,4F10.0)).

The program first prompts for the file name containing the coordinates. It then prompts for the unit cell constants to be used on the input coords (typing just return defaults to 1., 1., 1., 90., 90., 90. i.e. to orthogonal A). The next prompt asks if the input format is the default. If the user types NO, a new format will be requested; otherwise the atoms will be read with Wayne Hendrickson's PROLSQ format. Residue names can be supplied in either the single or triple letter codes. The standard format is (3X,A3,1X,A1,A3,A4,4F10.5), where the 'A1' is the position for the single letter code. The program returns to the first prompt to define another input dataset. This can be terminated by typing END in answer to this prompt.

PDB

This option allows one to read in coordinates which are in the standard Protein Data Bank format. The program prompts for the name of the PDB dataset and then asks whether these coordinates are to make the DSN2 dataset, or to update certain atoms in it (cf. UPDATE). All atoms are set free. In the former case, the user gets the PUTSEQ listing on the terminal.

READATOM

This option is to read in free format a list of atoms. It does not make a fully operational dataset since it does not read in the residue type. The free format coordinates must have one line for each atom with X, Y, Z, residue name, atom name, and an IA flag. The program prompts for the dataset of coordinates, and then for the unit cell of the coordinates in the prepared dataset (typing just a return defaults to 1.0, 1.0, 1.0, 90., 90., 90.). In a later version, this will be updated to include the residue type.

If you need to read in coords in some other format, list the subroutine called READRD and use this to see how to add a new option.

SSBRIDGE

This option is used to define disulphide bridges in the DSN2 data set. The option updates the IB flag so that it points to the record of the equivalent atom in the disulphide, i.e. if residues 100 and 200 are have a disulphide, CA of 100 points to CA of 200, and CA of 200 points to CA of 100, and likewise the CB and SG atoms. The program checks that the residues are of type CYS or CYH. Having defined these connections, the bridge atoms will be used in regularisation.

The program prompts for 2 residue names to define the pair. Typing END causes an exit. The program issues error messages if the residues are of the wrong type and if the expected atoms are not in the residues.

SYMMETRY

Crystallographic symmetry can be encoded in the DSN2 dataset. The crystal symmetry is generated in SPHERE mode and in the sphere in MIX mode (not ZONE). It prompts for the 6 unit cell parameters (A and degrees). The program then prompts for the lattice symmetry (the user must type one of P, I, R, F, A, B, C). It then prompts for symmetry operations until the user types END. The symop must fit on one line and is of the form shown in International Tables vol. 1, i.e. X,Y,Z or -X,1/2+Y,-Z etc. One must include the identity operation.

Example (P3-1); n.b. extra spaces are ignored, but the commas are essential.

1. X, Y, Z
2. -Y, X-Y, 1/3 + Z
3. Y-Z, -X, 2/3 + Z
4. END

Each of the inputs to SYMMETRY uses up one record in the extra area in the DSN2 dataset. If this area gets filled an error message is issued. The controller record does not get updated until the user types the END. Nonsense symops will be rejected with an error message. The symop is encoded as 12 integers, 9 defining a rotation and 3 the translation operation in 1/24'ths of the unit cell.

SYMERROR

This option can be used to correct any errors introduced by the user in defining the SYMMETRY operation. For each parameter the program prints the present value and asks if it is correct. Typing NO allows one to input the correct data.

UPDATE

This option allows one to update the coordinates and IA flag of an existing atom in the DSN2 dataset. The new coordinates are read in free format with data for each atom on one line. This data is X, Y, Z, residue name (usually the sequence number), atom name, and IA (optional). The first prompt is for the file containing the new coordinates. The coordinates must be orthogonal. The atoms can take on new IA values, and the program prompts for a number to define the desired option. Typing 0 means use the flag in the update file, 1 makes all the UPDATE atoms real, 2 makes them dummy, 3 makes them fixed real, 4 makes them fixed dummy.

This option is normally used when using the 'explosion' method of making an initial model. The UPDATE coordinates are then guide coords read from a mini-map and should include data for the CA and at least one side chain atom. One would make these atoms fixed and real, i.e. type 3. The user must have previously made a dataset using ABINITIO, READDIAM, PDB or READATOM.

PSEUDO

This options updates the pseudo sphere centre and radius for any part of the DSN2 dataset. The program prompts for a residue range; typing ALL forces the update of all residues in the dataset. If the sphere centre or radius is greater then 0.5A from its expected position, a message is issued.

ATOMTYPE

If the user wishes to use the MAKEDIAM option and has used SAM options ABINITIO, REPLACE or INSERT, then the atom type must be updated for the affected atoms. This is needed because the 'names' dataset does not store the atom type and not every one likes to use Diamond's data structure. The program prompts for a list of characters and values (format (A1,I1)) i.e. one character and one integer per line. The user terminates with escape/return. All atoms with a name starting with the character take on the atom type defined by the integer value. In Uppsala, this atom type later gets associated with a form factor in the PROTEIN system.

REINIT

This reinitializes the atomic records for a range of residues, i.e. the coordinates become 1500., 1500., 1500., and the atom types dummy (IA=2). The program prompts for the range of residues to be affected.

DELETE

This deletes a residue from the dataset, by removing it from the sequence area of the DSN2 dataset. The program prompts for the residue name, and is re-issued until the user types END.

REPLACE

This replaces a residue of one type by a residue of another type. If the 2 residues have atoms having the same name, then the information for those atoms is kept in the new residue. Otherwise, new atoms are placed at (1500., 1500., 1500.), are dummy, and of atom type 0. Note therefore that, e.g. LEU and PHE share CD1 and CD2 and must be regularised after REPLACE.

The program first prompts for the name of the residue to be replaced and then for the new residue type. A new residue can contain a maximum of 100 atoms. If successful, the program returns to the first prompt. The user terminates by typing END.

INSERT

This allows one to add a new residue to the DSN2 dataset. The first prompt asks AFTER WHAT RESIDUE DO YOU WANT AN INSERT ? If you wish to insert a residue at the start of the dataset, ???? should be entered, which is the name of a non existent residue always in the dataset. The second prompt is to define the name of this new residue (e.g. 100A), and the third prompt is to define the type of residue (e.g. ALA). The user then returns to the first prompt and can terminate input with END.

LISTXYZ

This produces a listing on the terminal of all the data held concerning the atom. It lists the dataset record number, IA (the atomic status flag), IB, internal residue number, X, Y, Z, temperature factor, weight, atom type, atom name, residue name (sequence number). It then lists all the data concerning each of the residues: name, residue type, 2 record pointers to the coordinate data, pseudo sphere centre and radius.

The program prompts to define the zone. Typing just ALL lists everything in the dataset.

LISTSEQU

This produces a sequence listing as described above. It has the same prompt as LISTXYZ.

ISITMADE

This checks for the occurrence of coordinates at (1500., 1500., 1500.) and reports it on the terminal.

MAKEDIAM

This makes the version of Diamond's dataset from the DSN2 dataset. The first prompt asks for the name of the dataset to be made. The second prompt is DEFAULT IS ALL ATOMS, O.K.?; typing NO allows one to select specific atoms from the DSN2 dataset. The prompt then asks for the atom names one per line until one types END, up to a maximum of 20. Continuing the program asks if one wants an overall temperature factor. Typing NO forces the use of the stored temperature factor, otherwise the program prompts for the value. The next prompt is for the unit cell angular constants to be applied; typing just return defaults to (90., 90., 90.). The program always forces A= B= C= 1.0A. The program then prompts for a residue range; typing ALL causes everything in the DSN2 dataset to be copied out.

Atoms with IA=2 or 4 are written out as dummy atoms with a weight of 0.0. There is a warning for atoms of type 0.

MAKEWH

This makes a list of coords in Wayne Hendrickson's format. This option has the same prompts as MAKEDIAM. Note that PROLSQ expects integer residue names. The temperature factor is written out in this version.

MAKEPDB

This makes a list of coords in standard Protein Data Bank format. This option has the same prompts as MAKEDIAM. HETATM records are dealt with correctly.

STOP

In the stand alone version of SAM this halts the program. In FRODO, control is passed back to the display.

FRODO MENU DESCRIPTION

The User interacts via a menu on the screen and via a terminal. The menu items are activated by using the mouse to move the cursor to the desired command and then pressing the middle button. The menu item is hi-lited when the cursor is over the item. If the hit is carried out correctly, the . next to the command will become a *. If an atom is hit by the cursor, its label (atom name + residue number) appears to the right of the atom position. A # before the residue number indicates that the atom was generated by symmetry from the original input position. Such an atom can be moved, but its new position cannot be SAVEd. A minus sign before the residue number indicates that the atom is dummy. If the operation is unsuccessful the system beeps and ****NO**** is displayed on the screen scratch text area. To reactivate an active function, hit it again - the asterisk will flash; it is not necessary to clear it first.

The pseudo A-D's are three rows of lines along the bottom of the screen. The cursor must be located on the desired A-D and the middle button pressed. The square of the distance of the cursor from the centre mark determines the rate of change; the direction of change is negative to the left, positive to the right. Positive X is to the right; positive Y is up; positive Z is out of the screen. All rotations and translations are with respect to the screen axes - not the crystal axes shown in the top right corner.
VIEWX, VIEWY, VIEWZ control the overall rotations.
INTA, INTB, ZOOM control the intensities of the molecule and the contours, and the picture size.
SLAB, Z, ORTHO control the window thickness in the Z direction, the window position in the Z direction, and 90 degree rotation about Y.
MOVEX, MOVEY, MOVEZ are for use with MOVE and FBRT and control translations of an atom or molecular fragment.
ROTZ, ROTY, ROTZ are for use with FBRT and control the rotations of a molecular fragment.
TOR1 to TOR6 are for use with TOR and control up to 6 active torsion angles.

The number keys on the keypad can be used to toggle on and off display of the various contour levels in the background object. Keys 0-4 correspond to the five available contour levels of the normal background object; if a composite background object is being displayed (see the CHAT command COBA), keys 5-9 correspond to the five levels from the secondary object.

Pressing the PF1 key will toggle on and off display of the screen menu, pseudo A-D's and text scratch area; any title (see the CHAT command TITL) moves to the bottom of the screen. This can be useful when taking slides from the screen. Subsequent pressing of the PF2 key will toggle on and off display of the pseudo A-D's remain displayed, so facilitating adjustments to the model.

Any number of commands can be active at the same time. However, this is an event driven system and certain commmands cause an exit (possibly just temporary) from the main program. However, some combinations must not be made - in particular the user should not try .TOR and .FBRT since both commands make use of the same array.

It is an advantage for the user to have some idea how the menu structure works. Each menu item has a flag associated with it. Each flag takes on values 0,1,2,3..... where 0 signifies that the item is not active. The value can then be used to signify what has been done and what is to be done next. Each command is checked one after the other and so information gathered by one command can be made known to any other command, e.g. .LSQP is used to determine least squares planes, and when activated the flag has the initial value 1. At each refresh cycle .LSQP is on, so if there has been a successful atomic identification, then the atom is kept for use in defining the plane. This carries on until the flag .YES is set. .LSQP checks for this and then sets the .YES flag off and increments its own flag. If enough atoms have been defined, it then calculates a least squares plane, and returns to wait for more atom hits. When a hit is made, the distance of the atom from the plane is calculated and displayed in the scratch text area. The command stays active until cleared by .WAIT.

If you want to know more about the menu, get a listing of inter.f. The menu commands are arranged in a (more or less) sensible order, defined in the data statement filling up the array MENU. The array ICC is the important flag array, and JCC contains pointers to go from menu position to flag index e.g. the I'th menu item corresponds to the flag ICC(JCC(I)).

The present menu is as follows:

.WAIT

All flags are switched off. Since the commands .MOVE, .FBRT, .TOR update the atomic coordinates in real time, this is equivalent to accepting the present positions.

.NOID

Clears the text scratch area, and the text on the atomic positions.

.NODA

Clears "continuous" functions (.LSQP, .XYZ, .DIST, .ANGL, .NAYB).

.SET

This resets the active torsion angles to their initial values.

.YES , .NO

These are used to cause "events", and to pick options for a number of commands. Hitting one switches off the other. This is useful in relation ta the status functions .STAT, .ALL & .FIX.

.CONT

Contours are picked out of the current DSN3 data set, centering on the last identified atom, using the current CRAD/CCOM parameters.

.POP

The user must be in ZONE mode, otherwise the operation is ignored. The action is to pop the display zone one unit along the linear sequence, writing the coordinates of what was the first residue in the zone to the DSN2 data set. If the last residue in the display zone is also the last residue in the coordinate data set, then the display zone is shortened by one unit. For proteins, the action is to move along the sequence from the N-terminal to the C-terminal. If .NO is not set, the coordinates of the new last residue in the display zone are read directly from the DSN2 data set. If .NO is set, then a stereochemically correct connection is made for polypeptides, provided the added residue already existed on DSN2 with correct bond lengths and angles. The command is normally used in the early stages of fitting a model to a map.

.SAVE

Has 3 different functions depending on .YES/.NO .

If neither .YES/.NO flags are set, the coordinates of non-symmetry generated atoms are written to disk. The user is free to make any nonsense on the screen, until .SAVE is set (see also .POP) at which time the disk coordinates get overwritten.

The .YES/.NO flags provides a backup and restore feature. With YES set, the prompt BACKUP: ARE YOU SURE? appears on the terminal. If the reply is YES, the coordinate dataset (DSN2) is copied to its backup (DSN7), which must already exist. With NO set, the prompt is RESTORE: ARE YOU SURE? With a YES reply, the reverse takes place, ie, DSN7 is copied to DSN2. The screen coordinates are not affected.
WARNING: a SAVE following a NO/SAVE will corrupt the coordinate dataset if the structures of DSN2 and DSN7 differ.
The current screen coordinates are not affected, since the operation is a straight file to file copy. If more backup capacity is required, use must be made of system utilities to make copies of the coordinates.

.BACK

This is an on/off switch for displaying the background contour object, and is useful when there are so many vectors on the screen that one cannot make out one's coordinates any more. It should not be confused with the BACK/NOBA(ck) options in .CHAT, which determine whether the contour object gets calculated or not.

.SIZE

This option is redundant on the Iris 2400 system.

.LSQP

Least squares plane calculation, using the method of Blow (Acta Cryst. (1960) 13,168). Subsequently hit atoms will be used in the calculation, up to a maximum of 100, until .YES is hit. Terminal output then gives information on the plane, and deviations of the atoms from the plane. More atoms can be identified, and their distance from the plane will appear in the text scratch area. At present all other commands are forced off by .LSQP.

.STAT

This is used to change the atomic status of subsequently hit atoms. Each atom carries a flag in its record in DSN2 which allows the atom to have 4 states - real, dummy, fixed real, fixed dummy (the flag is called IA and has values 1,2,3,4). If the atom is hit with .NO set,then the atom is forced to be dummy; if .YES (or not .NO) is set, the atom is forced to be real. The command can be used to:
1. Reset incorrectly .FIX(ed) atoms.
2. In connection with the .CHAT ADD option.
3. As a flag to show which atoms should be used in later structure factor calculations - see .SAM MAKEDIAM option. Note that once set,

the .YES/.NO flags can be changed since they are mutually exclusive.

.ALL

Defines the status of all the atoms in the residue of the subsequently hit atoms. .YES/.NO act as with .STAT.

.FIX

Subsequently hit atoms are not allowed to move when they are next subject to regularisation. After regularisation, the atoms are set free. With .NO set, fixed atoms are set free (this is not the same result as .STAT, since .STAT (or .YES/.STAT) makes the atom real, .NO/.STAT makes it dummy whereas .NO/.FIX keeps the atomic status whilst freeing it).

.XYZ

The current screen coords of a subsequently hit atom are displayed on the screen. A terminal listing is produced giving all the information that exists concerning the atom (this includes atomic temperature factor, weight and atom type if the coordinates were created with .SAM'S READDIAM/ PDB/ READATOM options). Note that the dataset coordinates will be different from the screen coords if the atom is symmetry produced or has been moved without .SAVE(ing). If the hit atom is moved (via .MOVE, .FBRT, .TOR), then the new position is updated on the screen.

.DIST

Display the distance between 2 subsequently hit atoms. The value is kept on the screen and gets updated if one or both atoms move. The function must be re-activated for each pair of atoms; 6 distances can be simultaneously updated.

.ANGL

Displays the angle defined by three subsequently hit atoms. The angle is updated and changes if one of the 3 atoms moves. The function must be re-activated for each triplet of atoms; 6 angles can be simultaneously updated.

.NAYB

Atoms which are within a certain distance (set by .CHAT command NAYB) from the next hit atom are displayed on the screen, together with their separation. Contacts between the hit atom and atoms in the

same residue as the hit atom are not displayed. The command remains active and continuously updates neighbours to the last hit atom. Subsequently hit atoms become new centres for neighbour calculations.

The next three commands allow the user to change the structure viewed on the screen.

.MOVE

The subsequently hit atom is allowed to move. Its position is controlled by the .MOVE A-D converters. The coordinates are updated as the atom moves. Its position is accepted by .YES or .WAIT, or is restored to its starting position by .NO.

Remember that for .MOVE, .FBRT, .TOR, the coordinates are not written to DSN2 until explicitly told via .SAVE, .POP, or the .CHAT command SAVE.

.FBRT

This stands for foreground- background rotate- translate, and is used to move a fragment of the molecule independently of the rest. All the atoms connected to the next hit atom move as a rigid group. The atom hit after activating .FBRT acts as the central pivot point. The maximum allowed size of the fragment is 1000 atoms. The fragment must, if necessary, first be isolated or defined by use of .BOBR and .BOND. Six pseudo A-D converters control the X, Y, Z rotations and translations. The rotations are concatenated and the shifts to apply to the screen axis.

Hitting .YES or .WAIT accepts the new postion, .NO restores the old coordinates.

As the atoms move, their coordinates get updated and can be labelled with .DIST, .ANGL, .NAYB commands.

Once the fragment has been defined, changing the connectivity will not change what atoms are affected by .FBRT.

.TOR

A maximum of 6 consecutive torsion angles can be defined, by defined by hitting a maximum of 9 atoms; i.e. the first torsion angle is for the bond between the 2nd and 3rd atoms, the second between the 3rd and 4th etc. The first atom hit will therefore remain fixed. Defining N angles therefore requires hitting N+3 atoms. Atom

definition can be terminated after defining a minimum of 4 atoms by hitting .YES. The first atom should be connected to the second, the second to the third and so on. The screen connectivity determines which atoms move when one of the torsion A-D converters is changed. All atoms with connections to the first atom other then via the second atom will not move. By applying suitable bond breaks, the user can limit the large long range movements caused by dihedral rotations. Care must be taken within ring structures, e.g. one can change CHI1 and CHI2 of a Phe ring by defining N, CA, CB, CG, CD1, but to vary the angle CB-CG-CD1-CE1 would first require breaking the CG-CD2 bond.

The part of the molecule that will be moved by TOR1 appear in a different colour, but the torsion angle A-D's can be changed in any order. .SET restores the initial conformation, keeping .TOR active. .YES keeps the current conformation; .NO restores the starting conformation, and .TOR is switched off. The dataset coordinates are only updated after a .SAVE command.

The initial screen connectivity is not determined by a bonding dictionary. This means that any molecule can be viewed in the display without having to work out a rigid connectivity list. The actual connectivity depends on the atomic viewing option. For ZONE mode, the atoms in each residue are connected if they are within 1.8A of each other and then a particular link atom in one residue is connected to a particular atom in the next. For SPHER mode, all atoms within the viewing sphere are connected when closer then 1.8A. In MIX mode, the user can define 10 zones plus a sphere, and uses the same connectivity criterion as SPHERE. SMIX is a special option which changes the distance to 1.3A for atoms beginning with H and 2.5A for S-S bridges.

The initial connectivity can be changed with the following two commands:

.BOND

To make a bond between the next two hit atoms.

.BOBR

To break the bond between the next two hit atoms. Sometimes by accident, two residues can interpenetrate and in SPHERE mode this produces many extra non-chemical bonds. If .YES is active, .BOBR will break all bonds between the two residues containing the hit atoms, which must themselves be bonded.

.M1,M2,M3,M4

These are viewport memories holding a set of VIEWX, Y, Z, INTA, INTB, ZOOM, SLAB, Z thickness parameters. The user can store a current set or retrieve the stored set. A loaded memory has a + instead of a . next to the menu item. To clear a memory hit .NO then the memory; .NO will be cleared when the last memory has been cleared. To load a memory it must first be empty and then hit. Note that memories cannot be cleared whilst any other function (e.g. .TOR), which act upon it, is active.

.ROCK

This switches rocking on and off, using the amplitude and speed defined in the .CHAT command ROCK. The rocking is about the screen vertical axis.

.ORIG

This shifts the screen centre to the next hit atom.

.DEL

Deletes the next hit atom from the screen; this is not allowed in ZONE mode. It is not used very often, but was intended as an aid to MULTAN users to pick out the structure from a PEKPIK's E-map. The atom is not removed from the DSN2 data set.

Of the following commands .REFI, .CHAT, .SAM and .MOL are exits from the interactive graphics routine. If the user forgets to SAVE the coords, they are lost when exiting via .MOL and .SAM, but can be written to disk via .CHAT command SAVE if exiting via .CHAT or .REFI.

.REFI

Initiates entry to model building or regularisation options. An initial prompt checks whether to proceed; the reply NO will cause the function to be abandoned and control to be returned to the mouse. The user has access to the CHAT interface to redefine constants such as refinement zones, guide angle options, etc. The model building options allow the user to build exact conformations such as alpha helices, by extension from an existing residue (REFI mode EXTD or from scratch with mode RESI). The regularisation is done by the method of Hermans and McQueen Acta Cryst. (1974) A30, 730. This method is a the molten atom method, since each atom moves by itself to satisfy its local conditions (provided it is not FIXed). It is probably not the most efficient method available but is usually well behaved. All bond lengths,angles and torsion angles are guided to preferred values

without any rigid constraints. This ensures that large errors do not build up at any point in the chain. Prompts are issued on entry to remind the user that he may need to redefine the RZONE and also to define which REFI option is active. During refinement, pressing the mioddle button of the mouse will cause the remaining cycles to be skipped.

.CHAT

This allows the user to redefine constants and/or define a new background from the DSN3 or DSN6 data sets. The user can also do rigid body least squares rotation/translation. See the full CHAT description for details.

.SAM

This is the interface which allows the user to manipulate the DSN2 data set in various ways. It can be used to DELETE, REPLACE or INSERT residues in the DSN2 data set. See the SAM description for full details.

.PLOT

Makes a copy of the screen image (excluding the menu and pseudo A-D's) on a plotting device. If the PF1 key is toggled off, the text scratch area is also excluded. In the export version of the program, the 'feedback buffer' is placed in a file in the /tmp directory with a name derived from plot??????, where ?????? are characters which define a unique name. A program is provided to convert this information into a Unix plot file; further processing may be performed as appropriate for the local plotting device.

.MOL

Allows the user to make a background object from the DSN2 data set. An initial prompt checks whether to proceed; the reply NO will cause the function to be abandoned and control to be returned to the mouse. A flexible set of user functions are available.

.STER

This switches on and off the display of two stereoscopic views.

.RSR

This performs a 6-dimensional rotational and translational real-space refinement of all atoms connected to the next atom hit after

activating .RSR. This region must, if necessary, first be defined by use of .BOBR and .BOND. These real-space refinement routines maximize Sigma(rho(obs)*rho(calc)) for the selected region, using the parameters supplied to the .CHAT option RSR and the scale factor obtained by .RSCL. All the atoms connected to the hit atom move as a rigid group.

.RSFX

This performs a 3-dimensional rotational real-space refinement of all atoms connected to the next atom hit after activating .RSFX; this atom remains fixed (see .RSR).

.RSCL

This calculates the scale factor between the observed (your map) and calculated electron densities for the region defined by the next hit atom after activation of .RSCL (you must do this before any refinement - see .RSR and .RSFX).

.END

Saves the constants dataset and terminates the program.