BUTTERWORTH HEINEMANN

PReDS: A graphics program to aid in the solution of difficult structures

Matthew D. Stock,[1] Andrew Beers,[1] and G. David Smith[1,2]

[1]Hauptman–Woodward Medical Research Institute, Buffalo, New York, [2]Roswell Park Cancer Institute, Buffalo, New York

INTRODUCTION

Although most small molecule structures can be solved routinely, there are always those structures that resist structure solution for a variety of reasons. These include data of marginal quality, the presence of noncrystallographic symmetry, structures in excess of 150 independent nonhydrogen atoms, or a variety of other reasons. Obviously, there is little that can be done to correct inaccurate *E*-magnitudes short of remeasuring the data. However, it is often the case for difficult structures that important structural information is contained in the electron density maps if one takes the time to examine them with a graphics program rather than rely on the connectivity of the largest peaks in the map. For example, in the structure determination of the gramicidin A dimer,[1] it was noted after the fact that a large percentage of the larger peaks tended to approximate the tracing of the peptide backbone, but attempts to connect peaks based on reasonable geometry proved to be a futile exercise.

In the past, we have used CHAIN[2] and FRODO[3] to display contoured *E*-maps and have been able to build correct models. However, CHAIN was designed to be used for macromolecular structures, and features that we felt to be important for the interpretation of E-maps are absent. These include but are not limited to the generation of symmetry-related regions of electron density, the construction of a "pseudo-map" about a series of peaks or trial model, and the ability to rotate or translate this "pseudo-map" independently with respect to the observed *E*-map, and finally the real-time rotation or translation of a symmetry-related molecule or group of atoms.

PReDS (Pattern Recognition for Difficult Structures) has been written to aid the small molecule crystallographer in extracting a direct method solution from difficult X-ray crystal structures. The program contains many features that are at present not available in existing graphics programs. The following section describes the hardware and software requirements for PReDS, the operations available in PReDS, how to obtain the software, and the future directions for PReDS.

HARDWARE AND SOFTWARE REQUIREMENTS

PReDS is written in the C programming language, using the X window system, the Silicon Graphics Graphics Library (GL), and the Motif widget set. Owing to the extensive use of GL graphics primitives, it requires an SGI workstation. The IRIS Development Option (IDO) is also needed for compilation of the software. A dialbox is optional, as a virtual dialbox can be invoked. The use of stereo glasses (CrystalEyes) is also supported and the "distance" between the eyes may be changed to suit the user. The perspective distance may be adjusted, allowing the user to make the effects of the *Z*-depth cueing more or less pronounced. The *Z*-clipping boundaries of the display can be repositioned at any time. A small subwindow gives a bird's eye view of the viewing area, illustrating the relationship of the clipping planes to the objects being displayed. Additional rendering options, such as *Z* buffering, depth cueing, and the display of maps as wireframes or isosurfaces, may be enabled or disabled at any time. PReDS has been tested with IRIX Versions 4.0.5, 5.2, and 5.3 on the Elan and Entry model Indigo workstations.

PReDS uses pulldown menus, illustrated in Color Plate 1, to allow the user to select most operations. A menu bar listing the major groups of operations available in PReDS is displayed at the top of the PReDS window. These topics may be clicked on with the mouse pointer to display the pulldown menu appropriate for that category. Some options, such as rotation, translation, and scaling (explained below), are input using the dialbox, if available. If no dialbox exists, input is obtained from buttons in a special "virtual dialbox" window. "Hotkeys," keys defined by the user as a shortcut to common operations, are planned but not yet implemented.

ATOMS

The format to input a potential molecular model is a subset of the PDB format. This format was chosen because of its

Color Plates for this article are on pages 303 and 304.

Address reprint requests to Dr. Smith at the Molecular Biophysics Department, Hauptman–Woodward Medical Research Institute, Inc., 73 High Street, Buffalo, New York 14203.

Received 4 May 1995; revised 2 August 1995; accepted 3 August 1995.

655 Avenue of the Americas, New York, NY 10010

0263-7855/95/$10.00
SSDI 0263-7855(95)00069-I

simplicity and standardization. Only the SCALE1, SCALE2, SCALE3, REVDAT, MASTER, COMPND, ATOM, HETATM, and CRYST1 record types are read or modified by PReDS, while all other record formats are stored in memory. Following the end of a session, the user may write a file containing the atoms of the manipulated trial structure in this format, and all record types stored in memory are written as they were in the original file.

While loading a PDB file, PReDS reads the COMPND record for the structure name, the CRYST1 record for the cell constants, and the ATOM and HETATM records for atom information. SCALE and MASTER records are ignored since the information contained in them can be computed. When writing a trial structure to a file, PReDS adds a REVDAT entry, writes the cell constants to the CRYST1 record, saves the computed SCALE records, writes the modified ATOM records, and stores the updated MASTER record. All other records in the original file are written without change. An attempt is made to preserve ordering of the various records in the output PDB file, but it is not guaranteed. Because of this and the additional memory required for storage, it is recommended that all unnecessary records be removed from an input PDB file.

While an ATOM or HETATM record in a PDB file specifies the type of an atom, it does not specify the atom's other properties such as covalent radius or van der Waals radius. Consequently, this information is obtained from an atom database file, which contains information pertinent to each type of atom, including the color in which it will be displayed. Information contained in the database file can be modified by the user to suit their individual preferences. Once the atoms in the trial structure are loaded into PReDS, the covalent radius of each atom is used to determine its bonding pattern within the molecule. Bonds are displayed as lines joining the respective atoms, and the two halves of the line are colored as appropriate for the particular atom type. Unbonded atoms are represented as small stars (lines joining the vertices of an octahedron). The display of unbonded atoms may be toggled on and off as desired. Atom labels may also be displayed according to the atom identifier in the ATOM record from the input file.

MODEL MANIPULATION

Two distinct sets of rotations and translations exist in PReDS. The first of these transforms the entire frame of reference and is called a "view transformation." In this case rotations and translations are performed on the entire contents of the screen about the primary orthogonal frame of reference, with the x axis horizontal, the y axis vertical, and the z axis perpendicular to the screen. The second set of rotations and translations is called "object transformations" and this set of operations is applied to one structure or to a map in relation to the primary frame of reference. While "view transformations" do not alter atomic coordinates, "object transformations" do and any changes are reflected in both the display and the subsequent output file.

The geometry of a trial model (moving, deleting, or adding atoms) may also be changed. To add an atom, one selects an atom to "clone," and then moves the cloned atom to the desired site in the crystal structure. Torsion angles within a molecule can also be altered with the exception of a bond within a cyclic moiety. Following a simple selection of the two atoms that define the bond, the dialbox or virtual dialbox can be used to rotate all atoms that comprise the fragment that is connected to the second selected atom. An arbitrary number of bonds may be rotated in this manner. Any of these operations may be aborted before they are completed. However, once the operations are applied they are permanent.

PReDS contains the symmetry for all 230 space groups. Selection of a space group permits the user to generate all symmetry-related atoms within a user-definable bounding box. Symmetry-related atoms may be toggled off and on to reduce screen clutter. Applying an object transformation to a molecule or molecular fragment is reflected in real time by all symmetry-related atoms and, thus, symmetry-related atoms are moving at the same time that the primary molecule is being rotated or translated. This can be useful in assessing the effects of packing constraints on the orientation and position of a trial structure. Optionally, the user may choose to defer application of the symmetry operations on the transformed molecule until after the operation has been completed, which provides better response from the program. It should be noted that molecular manipulations (bond rotations or atom deletions) cannot be directly applied to symmetry-related atoms.

Atoms can be "selected" with the mouse to allow operations to be done on them. When an atom is selected it and its symmetry-related atoms are "highlighted." Selecting atoms allows the user to apply the move, add, and delete operations, and allows torsion angles to be modified. The user can also request information on a selected set of atoms.

MAPS

PReDS is capable of displaying two different types of maps. The first of these is a map based on observed amplitudes such as E's, F(obs), and/or F(calc). These observed maps are preprocessed by the stand-alone program *mappreds*, which converts them into a binary file that can be read by PReDS. Necessary information that must be contained within the original map file includes the unit cell constants, the extent of the map, the number of divisions per unit cell edge, and the orientation of the axes. The map is displayed as a wireframe and the user may choose color and contour level. Up to eight maps can be displayed simultaneously. By default, PReDS will display the entire map, but the user may limit the volume of map to be contoured by interactively choosing a volume within a bounding box. These types of maps may be rotated and translated using any of the view transformation devices, but cannot be moved independently of the trial molecule or fragment. As is frequently the case, the molecule or fragment of interest may extend beyond the boundary of the asymmetric unit of electron density that has been externally computed. For this reason, provision has been made to read a map generated in this way and apply a space group symmetry operation to generate the missing regions of density. The extent and grid divisions of the input map must be chosen to be consistent with the symmetry operations of the space group.

A second type of map may be generated internally by

PReDS around a trial structure. These maps are constructed after the method of Barry,[4] in which a pseudo-electron density function, is calculated about each atom or peak. As this function is a Gaussian function, it can be contoured at any level, giving a representation of a van der Waals surface of the molecule or fragment. A typical surface, generated about the backbone atoms of the gramicidin A dimer,[1] is illustrated in Color Plate 2. Unlike *E*- or *F*-maps, the pseudomap is dependent on its parent structure, and thus object transformations can be applied to these maps. If a rotation about a bond is performed, an atom moved or deleted, or the parent structure rotated or translated, the pseudomap will be automatically regenerated to reflect the changes that have been made. This recomputation can be done in real time, or it can be deferred until the operation is complete. The latter option provides better performance in the case of larger structures.

APPLICATIONS

One potential application of PReDS in the determination of a difficult structure can be illustrated using the orthorhombic gramicidin A structure.[1] Shown in Color Plate 3 is a poorly phased *E*-map with an rms phase error of approximately 60°, obtained using the SnB procedure[5] and a 30-atom fragment. The accuracy of these phases is not sufficient to generate a correct structure by tangent refinement. Also illustrated in Color Plate 3 are the interpolated peak positions along with their van der Waals surface generated about the peak positions. While the connectivity of the peaks, based solely on interpeak distances, does not define a chemically reasonable structure, the peak positions do provide a surface that approximates the backbone surface shown in Color Plate 2. Following the manual rotation and translation of the model backbone structure and its associated van der Waals surface onto the surface described by the *E*-map peak positions, shown in Color Plate 4, it becomes apparent that the model does approximate the *E*-map and is sufficient to generate maps that are fully interpretable. Illustrated in Color Plate 5 is the model and the initial *E*-map.

FUTURE PLANS

PReDS is an active research project, under constant development and enhancement. Some features that would have delayed the release of the software were not implemented in the current version. Hardcopy output, which is minimally supported for structures, will be improved to include the printing of maps and will give a broader selection of output formats. In addition, features such as hotkeys, distance measuring, and figures of merit for map comparison are being planned. Mail to preds@hwi.buffalo.edu any suggestions or comments about the software.

ACQUIRING PReDS

PReDS can be obtained by E-mailing the PReDS support team at preds@hwi.buffalo.edu.

ACKNOWLEDGMENT

The authors thank Dr. David A. Langs for helpful discussions and for providing the gramicidin A data, maps, and structure. This work has been supported by NIH Grant GM46733.

REFERENCES

1 Langs, DA. *Science* 1988, **241,** 188–191
2 Sack, J.S. CHAIN—A crystallographic modeling program. *J. Mol. Graphics* 1988, **6,** 244–245
3 Bush, B.L. and Jones, T.A. *PS300 FRODO—Molecular Graphics Program for the PS300* (modified by Sack, J., *et al.*). 1988, Houston, Texas
4 Barry, C.D. Personal communication
5 Miller, R., DeTitta, G.T., Jones, R., Langs, D.A., Weeks, C.M., and Hauptman, H.A. On the application of the minimal principle to solve unknown structures. *Science* 1993, **259,** 1430–1433