NEW PROGRAMS

MOLPACK: Molecular graphics for studying the packing of protein molecules in the crystallographic unit cell

Dawei Wang,*† Huub P. C. Driessen† and Ian J. Tickle*

*†Imperial Cancer Research Fund Unit of Structural Molecular Biology and *Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, London, UK*

A graphics program, MOLPACK, has been developed on the Silicon Graphics IRIS-4D computer system for displaying the packing of proteins in the crystallographic unit cell. In addition to the normal viewing operations of rotation, translation and scaling, the program has the ability to translate molecules along the cell axes while maintaining their crystallographic equivalent positions within the unit cell. This allows the user to observe the packing of protein molecules generated by molecular replacement, to create a new packing model or to locate an unknown molecule. A special feature of the program is that up to four independent molecules can be manipulated in the asymmetric unit.

Keywords: crystallographic unit cell, crystal packing, molecular graphics, molecular replacement

Address reprint requests to Dr. Driessen at the Imperial Cancer Research Fund Unit of Structural Molecular Biology, Birkbeck College, Malet Street, London WC1E 7HX, UK.
Mr. D. Wang's permanent address is Laboratory of Enzyme Engineering, Jilin University, Changchun, P.R. China.
Received 21 August 1990; accepted 4 September 1990

INTRODUCTION

The way in which the molecules of a protein pack together often has a significant influence on its biological properties. Packing may be studied in detail during single-crystal X-ray structure analysis.[1] However, more global studies are needed for molecular replacement, where often a series of peaks of the translation function need to be checked for the existence of a plausible solution.[2,3] Although there are many excellent molecular graphics software packages available for the former, e.g., FRODO[4] and QUANTA,[5] none have features that make the latter especially easy. The program FITZ,[6] designed for the Evans and Sutherland Picture System 2, features this packing option. With this graphics system becoming obsolete and the proliferation of molecular replacement studies,[3] a new program has become desirable.

This paper describes a new molecular graphics program, MOLPACK, designed specifically for studying the packing of protein molecules in the unit cell. The program is written in FORTRAN and can be run on the IRIS-4D, a high-performance computer graphics workstation.[7] It reads in atomic coordinates in the standard Brookhaven Protein Data Bank format and displays the three-dimensional structure of molecules, each in a different color, with a representation of the edges of the unit cell. In addition to the usual viewing operations, such as rotation, translation, clipping and scaling, the program also allows the molecules to be translated separately along the *a*-, *b*- and *c*-axes while maintaining their crystallographic symmetry within the unit cell. An important new feature is that it allows up to four independent molecules per asymmetric unit, although this number can be increased easily. The molecules can be displayed simultaneously and manipulated individually, thus enabling the user to study and experiment with new packing models in more complicated cases.

PROGRAM DESCRIPTION

The program can be divided into two parts: The first part reads in the crystal information and atomic coordinates and sets up transformation matrices to perform space-group symmetry operations, fractionalization and orthogonalization, and molecular replacement rotations and translations. It optionally outputs new coordinates to a data file. The second part displays the unit cell and its contents, performs viewing transformations and generates new packing models. In this part the program operates in a graphics mode, as supported by the IRIS graphics hardware and graphics library.

USER INTERFACE

MOLPACK is run interactively and supports two of the input devices available for the IRIS-4D system, the keyboard and mouse. One uses the keyboard to input commands, file names, control data and parameters, and the mouse to select options from the various menus or to control the motions of molecules and objects on the screen.

The main menu allows the user to manipulate the object; that is, the user may perform rotations, translations or scalings and may move molecules within the unit cell or reset them to their original positions. Control over manipulations is effected in a natural and straightforward manner with the mouse. A fixed box is used for clipping. Other less frequently used options are available through submenus. The user may select atoms for the display of the molecules; these can be C_α atoms, backbone atoms or entire molecules. A second option allows one to switch labels, text, the axes cross, z-buffering and so forth. A third option allows the user to generate static layers or spheres of molecules around the central one. In this mode one may perform the same kinds of object manipulations as in the main menu; however, manipulations in one mode are independent of manipulations in the other.

INPUT/OUTPUT

MOLPACK supports the Brookhaven Protein Data Bank file format. The program first asks how many molecules are present in the asymmetric unit. It then reads in the atomic coordinates and unit cell parameters from the coordinate file(s) entered and determines the fractionalization and orthogonalization matrices where possible. If no cell parameters or matrices are available, the program will prompt the user for them. The user can then enter the space-group number or space-group symbol as defined in *International Tables*.[8] The program extracts the appropriate symmetry operations from a library file and decodes the symmetry operations into matrices. Alternatively, the user can input individual symmetry operations in *International Tables* format using the keyboard. For molecular replacement studies, the program can apply rotations and translations. A rotation may be input as a matrix or as eulerian angles according to the convention used in Crowther's fast rotation function.[9]

MOLPACK uses the symmetry operations to generate all of the molecules in the unit cell from the coordinates in the coordinate file. For each molecule of the asymmetric unit only one set of coordinates is held in memory, with the remainder being generated by the symmetry operations. This minimizes memory demand and means that only the matrices (rather than the coordinates) need be updated, giving maximum speed of operation. The matrices are updated each time a new model is generated by translations in the unit cell. The user can save the "current" coordinates in an output file at any time, or can reset the matrices.

TRANSLATION OF THE MOLECULES

One of MOLPACK's important functions is its ability to translate molecules along the a-, b- or c-axes relative to a certain symmetry position and to restrict movement within the unit cell. According to the standard definitions,[7] the unit cell edges are defined as vectors **a** in the (100) direction, **b** in the (010) direction and **c** in the (001) direction; the interaxial angles are α, β and γ. The transformation is based on the orthogonalization matrix acting on the input coordinates. To move molecules individually while maintaining their symmetry positions, the program builds up a graphics object for each molecule in the asymmetric unit before display. It then gets the translation value from the mouse and applies the translation vector to each symmetry operation matrix to move the molecule in accordance with its symmetry relation. The unit cell edge remains fixed. To restrict the motion of the molecules within the unit cell framework, the program updates the coordinates of the center of mass of each molecule when it is translated. If the center goes beyond one of the unit cell edges, a fractional translation of 1.0 is added or subtracted to the relevant matrix element so that the molecule is reintroduced into the unit cell from the opposite edge.

TRANSFORMATION MATRIX

The total transformation matrix $SYMO(ji)$ for molecular object i (in orthogonal Ångstrom coordinates) inside asymmetric unit j is

$$SYMO(ji) = O\ Tin(ji)\ SYME(j)\ Tv(i)\ Tmr(i)\ O^{-1}\ Rmr(i)$$

All operations, including translations, are represented by homogeneous 4×4 matrices that premultiply the homogeneous coordinate vector. Matrices $Rmr(i)$ and $Tmr(i)$ describe rotations and translations derived from molecular replacement for molecule i; O is the orthogonalization matrix for the input coordinates; $Tv(i)$ is a translation vector for examining different packings; $SYME(j)$ is the equivalent position; and $Tin(ji)$ is the translation vector needed to confine the molecule to the unit cell.

RESULTS

For test purposes we have used the program on structures in triclinic,[10] monoclinic,[11] orthorhombic[1,12] and tetragonal[13,14] space groups. Atomic coordinates other than our own were obtained from the Brookhaven Protein Data Bank. The program has been used successfully to examine the molecular replacement solution of the rat eye-lens protein γE-crystallin, which has two molecules in the asymmetric unit of space group $P2_1$.[15] The packing of the C_α atoms of the search molecule in the correct orientation and position is shown in Color Plate 1. There are no packing conflicts and molecules are at the right distance for interactions. A slightly different packing model, with one of two molecules translated in the asymmetric unit, is shown in Color Plate 2.

NOTE

The MOLPACK software can be obtained from the Department of Crystallography, Birkbeck College.

ACKNOWLEDGMENTS

D. Wang wishes to thank the staff at the Department of Crystallography, Birkbeck College for help and hospitality. Particular thanks are due to Mr. Z. Zhu, Mr. R.A. Laskowski and Mr. R. Westlake. We thank Professor T.L. Blundell for his support and useful discussions. Thanks are due to Collaborative Computing Project 4 (CCP4), Science and Engineering Research Council, Daresbury Laboratory for the use of the input/output subroutines for atomic coordinates.

REFERENCES

1 White, H.E., Driessen, H.P.C., Slingsby, C., Moss, D.S. and Lindley, P.F. Packing interactions in the eye-lens. Structural analysis, internal symmetry and lattice interactions of bovine γIV*a*-crystallin. *J. Mol. Biol.* 1989, **207**, 217–35

2 *The Molecular Replacement Method* (Rossmann, M.G., Ed.) Gordon and Breach, New York (1972)

3 *Molecular Replacement: Proceedings of the Daresbury Study Weekend* (P. Machin, Ed.) Science and Engineering Research Council, Daresbury, UK (1985)

4 Jones, T.A. A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* 1978, **11**, 268–72

5 *QUANTA Reference Manual.* Polygen Corporation (1989)

6 Taylor, G. FITZ-Interactive graphics tool for investigating molecular symmetry and homology. *J. Mol. Graphics* 1983, **1**, 5–8

7 *IRIS Reference Manual.* Silicon Graphics Inc. (1988)

8 *International Tables for Crystallography. Vol. A: Space-Group Symmetry* (T. Hahn, Ed.) 2nd edition, Kluwer, Boston (1989)

9 Crowther, R.A. The fast rotation function. In *The Molecular Replacement Method* (M.G. Rossmann, Ed.) Gordon and Breach, New York (1972) 173–8

10 Lederer, F., Glatigny, A., Bethge, P.H., Bellamy, H.D. and Mathews, F.S. Improvement of the 2.5-Å resolution model of cytochrome b_{562} by redetermining the primary structure and using molecular graphics. *J. Mol. Biol.* 1981, **148**, 427–48

11 Liljas, A., Kannan, K.K., Bergsten, P.-C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, U., Jarup, L., Lovgren, S. and Petef, M. Crystal structure of human carbonic anhydrase *C*. *Nature New Biol.* 1972, **235**, 131–7

12 Timkovich, R. and Dickerson, R.E. The structure of *paracoccus denitrificans* cytochrome C_{550}. *J. Biol. Chem.* 1976, **251**, 4033–46

13 Remington, S., Wiegand, G. and Huber, R. Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7- and 1.7-Å resolution. *J. Mol. Biol.* 1982, **158**, 111–52

14 Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. and Wistow, G. The molecular structure and stability of the eye lens: X-ray analysis of γ-crystallin II. *Nature* 1981, **289**, 771–7

15 Driessen, H.P.C. and Slingsby, C., The three-dimensional structure of rat eye-lens γE-crystallin, unpublished