NEW PROGRAMS

XELE—a polypeptide model-building program for a graphics workstation

Isao Fujii,* Yukio Morimoto, Yoshiki Higuchi and Noritake Yasuoka

Department of Life Science, Faculty of Science, Himeji Institute of Technology, Kanaji 1479-1, Kamigori, Akou-gun, Hyogo 678-12, Japan

A model-building program, XELE, for use in protein crystallography has been written in C under UNIX on a graphics workstation. This program makes full use of the X *Window system to display the electron density distribution and to manipulate the polypeptide model, and therefore is named XELE. It utilizes a fast three-dimensional rendering package, Dorè, and is portable to other types of graphics workstations. A part of the program for the man–machine interface uses the library of* X *Window and* X *Toolkit, and therefore is highly interactive. The structure analysis program package, PROTEIN, is also implemented in an interactive mode using* X *Window, and has been interfaced with XELE.*

Keywords: protein crystallography, electron density, polypeptide model-building, isosurface model

INTRODUCTION

Recent advances in semiconductor technology, as well as computer architecture, have made it possible to produce a workstation of rather low cost that has the ability to perform several tens of millions of instructions per second. Some workstations are equipped also with special CPU and memory boards and can display several tens of thousands of polygons per second. A graphics workstation seems especially suitable for use in a protein crystallography laboratory where the graphical display of three-dimensional (3D) structure and electron density distribution is inevitable. Most graphics display terminals have been rather high in cost.[1–3] We recently introduced a TITAN graphics workstation into our laboratory, and developed a graphics display program XELE. We also implemented a protein structure analysis program system PROTEIN[4] in a highly interactive mode using X Window.[5] An X-ray diffraction apparatus that uses an imaging plate also is connected to TITAN via Ethernet, which permits the transfer of the diffraction image data to the workstation where they are processed to give structure factors.

FRODO,[6] which was developed by T.A. Jones, is a famous model-building program that is widely used. XELE has the essential features of FRODO. It was developed using a rendering system called Dorè (dynamic object rendering environment), and is portable to other graphics workstations. We have developed the program from the following points of view: high transparency among different types of machines, dependence on a good choice of rendering systems, no necessity to convert data output from the analysis programs, and easy operation in model building.

X Window and X Toolkit also have made the program highly interactive. We propose a new type of display mode, an isosurface model of an electron density map formed by a polygonal mesh, to perceive the three dimensionality and also to reduce the data specifying the objects. In this paper we describe the program XELE and the graphics workstation dedicated to protein crystallography.

DESCRIPTION OF THE SYSTEM

A TITAN graphics workstation in our laboratory consists of two processors using RISC architecture. Each of the processors comprises an integer unit and a vector unit. The peak performance of one processor is 16 MIPS for the integer unit and 16 MFLOPS for the vector unit. The computer is equipped with 64 MB main memory and 380 MB × 2 hard disks. A multiprocessor architecture enables vector or parallel programming in FORTRAN and C. The operating system installed in the computer is UNIX System V Release 3.0 (4.3 BSD expand). The user interface module consists of a high-resolution color display monitor, a junction box, a keyboard, a mouse and an optional 2-column by 4-row knob box. The monitor has a 19-in screen, and provides 1280 × 1024 pixels resolution. The frame buffer, with 24 color planes, 4 overlay planes (for cursor, pop-up menu, etc.) and 3 control planes (for color-mode and double-buffering selection), mounts on the base board, and TITAN graphics supplies a color palette of over 16 million colors. The expansion board adds 24 additional color planes to supply the double-buffered mode. Some personal computers and

Color Plates for this article are on page 165.

*Present address: Department of Biological Science and Technology, Tokai University School of High-Technology for Human Welfare, Nishino 317, Numazu, Shizuoka 410-03, Japan.
Address reprint and/or source program requests to I. Fujii or N. Yasuoka.

Received 9 November 1990; accepted 7 August 1991

an X terminal are connected to the workstation via Ethernet. An X-ray diffraction apparatus using an imaging plate is also connected to the computer.

PERFORMANCE AND FUNCTION OF XELE

XELE, a program for displaying an electron density distribution and for fitting a polypeptide model to it has been written with ANSI C language. It is based on the graphics package Dorè. A part of the coded program is shown in Figure 1. Dorè enables all images to be displayed as objects, including primitives, shapes, lights, motion and viewing directions. The mode for expressing a primitive is chosen among dot, wire and surface models. Furthermore, it is easy to specify or to change the attributes of the viewed objects, that is, the color, the mode of the view (perspective or orthogonal) and the type of light sources. The interactive feasibility of the program is achieved with the use of Xlib and X Toolkit of the X Window System V. 11. A summary of the characteristics of the program follows.

Command architecture

After the pointer cursor is moved to the menu displayed on the subwindow, the operator inputs some parameters to click the command widget with the mouse or to type letters in the dialog widget. The following operations are in the command menu: file input, symmetry operation of a molecule, motion of a partial model, rotation of a moiety of the model around a chemical bond, change of viewpoint, stereo view switching on divided windows, model clipping, pick-up of specified coordinates, pick-up of water molecules, and label switching. A displayed picture showing a polypeptide model with the menu is shown in Color Plate 1. Thus a beginner easily can build a model, without special skills or reference to the manual. He can input the rotation angle of models or the translation distance with the dials in the knob box. The functions of dials may be switched according to operations during a model building. Therefore, the model can be manipulated easily in spite of a limited number of dials.

Description of a dataset

The coordinates and connectivities of atoms are supplied to the program in formatted form. The name of a data set is input from the keyboard through a dialog widget into the menu. The format of the coordinate data is the same as that used in the Brookhaven Protein Data Bank[7] and in the PROTEIN system. The atom data set is stored in the main memory, and the wire model can be built by recursive calling. Thus the generated wire primitive can be drawn in a single stroke of a pen and be rendered faster. If a part of the peptide is rotated around a specified bond, the connections also are constructed in a similar way, and the coordinates of atoms are expressed by the following

```
make_objects()
{
  /*  environment definitions */
  DsSetObjName(DoGroup(DcTrue), DcNameString, "cameragroup", DcTrue);
      DgAddObj(DoLookAtFrom(camera_origin, camera_from, up));
      DgAddObj(cameral=DoCamera());
    DgAddObj(DoPopMatrix());
    DgAddObj(DoPushMatrix());
      DgAddObj(DoLookAtFrom(origin, light0, up));
      DgAddObj(DoLightIntens(.7));
      DgAddObj(DoLight());
    DgAddObj(DoPopMatrix());
  DgClose();
  /*  primitive definitions */
  DsSetObjName(DoGroup(DcTrue), DcNameString, "objectgroup", DcTrue);
    DgAddObj(clip);
    DgAddObj(translate_ingrp);
    DgAddObj(DoPushMatrix());
      DgAddObj(rotate_ingrp);
      all_objects  =DoGroup(DcTrue);
        mesh_gp  =DoGroup(DcTrue);DgClose();DsHoldObj(mesh_gp);
        atom_gp  =DoGroup(DcTrue);DgClose();DsHoldObj(atom_gp);
        symm_gp  =DoGroup(DcTrue);DgClose();DsHoldObj(symm_gp);
      DgClose();
      DgAddObj(all_objects);
    DgAddObj(DoPopMatrix());
  DgClose();
}
```

Figure 1. Part of the program code for displaying moleculer and electron density distributions.

equations:

$$(X\ Y\ Z\ H) = (x\ y\ z\ l)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -h & -k & -l & 1 \end{pmatrix} \quad (\mathrm{R}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ h & k & l & 1 \end{pmatrix}$$

$$(R) = \begin{pmatrix} n_1^2 + (1 - n_1^2)\cos\theta & n_1 n_2(1 - \cos\theta) + n_3\sin\theta & n_1 n_3(1 - \cos\theta) - n_2\sin\theta & 0 \\ n_1 n_2(1 - \cos\theta) - n_3\sin\theta & n_2^2 + (1 - n_2^2)\cos\theta & n_2 n_3(1 - \cos\theta) + n_1\sin\theta & 0 \\ n_1 n_3(1 - \cos\theta) + n_2\sin\theta & n_2 n_3(1 - \cos\theta) - n_1\sin\theta & n_3^2 + (1 - n_3^2)\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

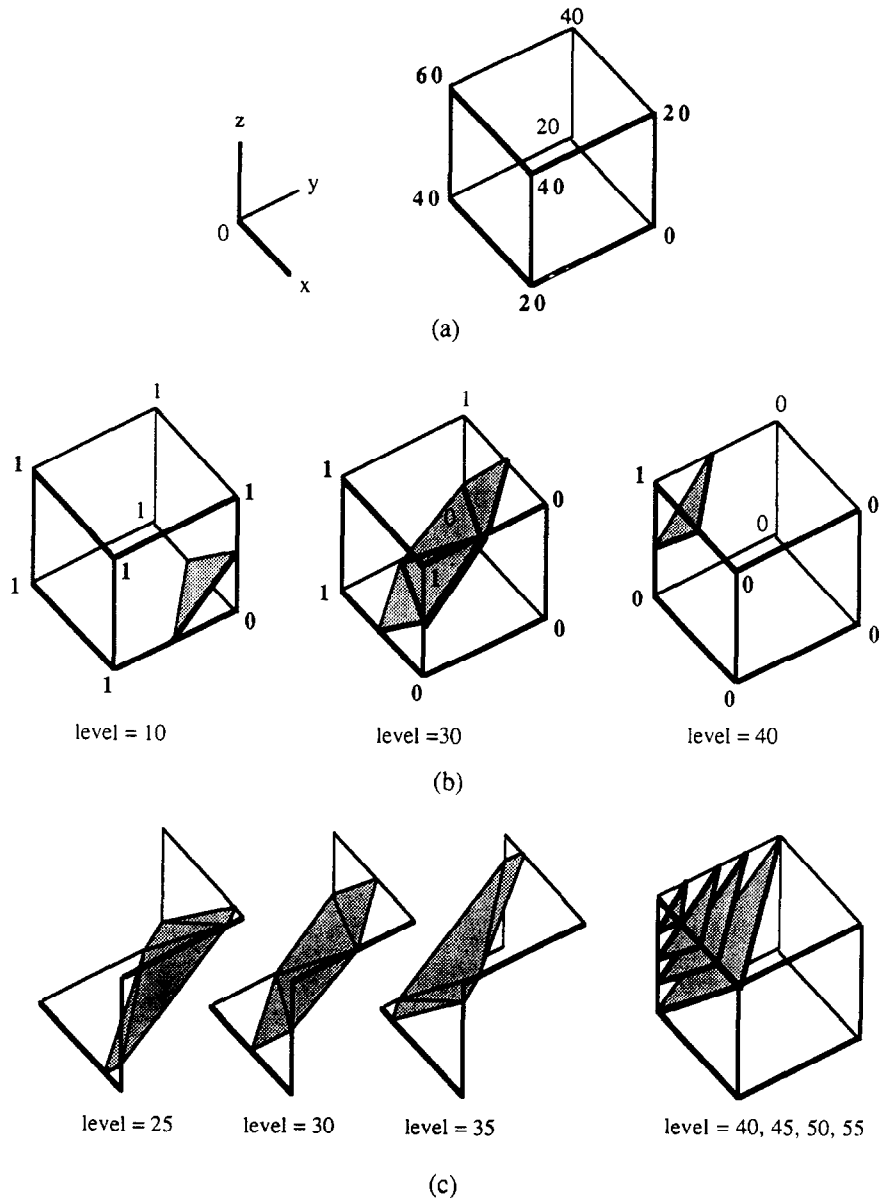

Figure 2. Construction of isosurface models from three-dimensional electron density distribution data: a, electron density distribution at grid points; b, a flag is set to 1 when the electron density at a grid point is larger than the specified level; and c, the intersecting point can be displaced according to the difference in these values.

where the point (h, k, l) is the origin of rotation; the axis of the rotation is designated by the direction (n_1, n_2, n_3) (directional cosines), and θ specifies the angle of the rotation.

Electron density map

A wire model usually is employed to display an electron density map on a graphics terminal: contours corresponding to a specified level are stacked along three orthogonal directions to give a basket-like model. It is the simplest way to represent an electron density distribution graphically on a vector scan-type graphics display, and is still used on raster scan-type graphics displays. The latter has the ability to draw surface primitives rather than line primitives. To make better use of this characteristic, we have developed an isosurface model to represent the electron density distribution. We assume a curved surface to represent a specified level of the electron density, and call it an isosurface. This surface is drawn as a composite of small planar triangles. The steps to form an isosurface model are illustrated in Figure 2.

(1) The values of the electron density at each grid point are input into the 3D array.
(2) If a specified level is smaller than the value of a grid point, a flag is set to be 1 on this point; otherwise 0. If we assume a parallelpiped consisting of 8 grid points, the pattern of the combination of flags may be one of 256 types that are formed from triangles and stored in advance. The corresponding patched element is assigned to this parallelpiped.
(3) The intersecting point can be displaced according to the difference in electron density between the values at the grid point and the specified level (Figure 2c).
(4) These elements are connected in the whole unit cell and are patched to form an isosurface model.

An example of the surface representation of the isosurface model is shown in Color Plate 3.

A sphere 1 Å in diameter is generated anywhere and may be moved through the space. As shown in Color Plate 2, a closed shell of electron density includes this sphere. It is most probable that this part corresponds to a water molecule, and it is useful when one wishes to pick up an isolated water molecule. A label is easily attached to the specified atom to show the kind of elements, the name of the residue, etc. One stroke-writing mode is used as much as possible to reduce the number of objects, which enables a fast rotation of the model. A rotation or translation of the model is executed in the coordinate system of the observer. The area to be shown on a display is easily specified by clipping the model.

A copy of the source program will be supplied upon request by one of the authors.

IMPLEMENTATION

Image processing

FILME,[8] originally written to obtain structure factors from oscillation photographs, have been modified to process the data taken on a DIP100 system. A displayed picture taken on a DIP100 of a hydrogenase crystal[9] is shown in Color Plate 4.

Program WEIS[10] also was implemented, to reduce structure factors for larger molecules from data from the Weissenberg camera[11] for macromolecular crystallography installed in The Photon Factory at the Institute for High Energy Physics.

Structure analysis

The program package PROTEIN, which was developed in the Max Planck Institute für Biochemie, was implemented for structure analysis. A derived electron density distribution can be displayed on a monitor and then output to a plotter. PROLSQ,[12] which performs the restrained parameters least-squares refinement, was also implemented. Many programs used for structural analysis of small molecules, such as MULTAN78,[13] also are implemented.

DISCUSSION

XELE and other programs have been used in the structural refinement of cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough.[14] A detailed description of the structural analysis is given elsewhere.

We have developed our system under UNIX, which provides a comfortable environment. The program package, including the graphics programs, is highly portable. In conclusion, a graphics workstation of this level is quite suitable for the protein crystallography laboratory.

REFERENCES

1 Hubbard, R.E. Trends in Workstation Hardware and Graphics. *Acta. Crystallogr.* 1990, **A46**, C33
2 Iga, Y. and Yasuoka, N. VENUS—a program to display protein structure using raster color graphics. *J. Mol. Graphics* 1984, **2**, 79–82
3 Nakamura, H., Kusunoki, M. and Yasuoka, N. A new approach to illustrating electrostatic molecular surfaces. *J. Mol. Graphics* 1984, **2**, 14–17
4 Steigemann, W. Ph.D. Thesis, Techniche Universität München, 1974
5 O'Reilly, T. *The Definitive Guides to the X Window System.* O'Reilly & Associates 1989, vol. 1–5
6 Jones, T.A. A graphics model-building and refinement system for macromolecules. *J. Appl. Crystallogr.* 1978, **11**, 268–272
7 Bernstein, F.C., Koetzle, T.F., Williams, G.T.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, T. The Protein Data Bank: a computer-based archival file for macromolecular structure. *J. Mol. Biol.* 1977, **112**, 535–542
8 Schwager, P., Bartels, K. and Jones, A. *J. Appl. Crystallogr.* 1975, **8**, 275–280
9 Higuchi, Y., Yasuoka, N., Kakudo, M., Katsube, Y., Yagi, T. and Inokuchi, H. Single crystal of hydrogenase from *Desulfovibrio vulgaris*. *J. Biol. Chem.* 1987, **262**(6), 2823–2825
10 Higashi, T. The processing of diffraction data taken on a screenless Weissenberg camera for macromolecular crystallography. *J. Appl. Crystallgr.* 1989, **22**, 9
11 Sakabe, N., Nakagawa, A., Sasaki, K., Sakabe, K., Watanabe, N., Kondo, H. and Shimomura, M. *Rev. Sci. Instrum.* 1989, **60**, 2440–2441

12 Hendrickson, W.A. and Konnert, J.H. *Biomolecular Structure, Function, Conformation and Evolution*. Pergamon Press, Oxford 1980, 43
13 Main, P., Hull, S.E., Lessinger, L., Germain, G., Declercq, J.P. & Woolfson, M.M. *MULTAN78*. University of York, England 1978
14 Morimoto, Y., Tani, T., Okumura, H., Higuchi, Y. and Yasuoka, N. Effects of amino acid substitution on three-dimensional structure. An X-ray analysis of the cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough at 2-Å resolution. *J. Biochem.* 1991, **110**, 532–540