BUTTERWORTH HEINEMANN

MolView: A program for analyzing and displaying atomic structures on the Macintosh personal computer

Thomas James Smith

Department of Biological Sciences, Purdue University, West Lafayette, Indiana

A program is described that allows the user to analyze and display atomic structures on any Macintosh personal computer. The program reads ASCII format structure files including PDB, plot files from the graphics programs O and FRODO, and Cartesian coordinates from ChemDraw 3D. The program has a graphical interface that features floating button palettes for objects and tools. The structures may be displayed using stick, ball-and-stick, space-filling, and ribbon models. Each type of drawing can be colored according to a variety of schemes to accentuate various structural aspects. The figures can be rotated, displayed in stereo, and exported using the Clipboard, PICT files, or QuickTime movies. The structure can be further analyzed by displaying hydrogen bonds, making Ramachandran plots, labeling atoms, measuring distances, and finding neighboring atoms. By using the Macintosh computer and emphasizing a graphical interface, this program helps to bring structural analysis to students and researchers that may not have access to, or experience with, large graphics workstations. In addition, the object-oriented output PICT images are ideal for creating publication-ready diagrams that can be easily modified or inserted into other documents (e.g., see Refs. 1–3).

Keywords: Macintosh, graphics, atomic structure

INTRODUCTION

It is often difficult to convey three-dimensional aspects of atomic structures without a "hands on" demonstration. Because most of the more powerful graphics software is written exclusively for the expensive workstations, it is impractical to incorporate structural analysis into curricula for larger classes. In addition, many instructors only occasionally create structural diagrams and therefore need an application that is easy to use and readily available on inexpensive machines.

MolView represents an attempt to bring a number of the display tools found in other graphics programs (e.g., FRODO[4,5] and O[6]) to the desktop. Emphasis was placed on the interface so that it can be used in undergraduate teaching or by researchers that are not experienced with molecular graphics. By being able to create QuickTime movies, instructors can demonstrate three-dimensional aspects of complex models with interactive, real-time motion on even the most inexpensive of Macintoshes without the aid of stereo viewers.

MolView was written using the Macintosh Programmers' Workshop (MPW) environment and the associated C compiler. In the current distribution, there are versions of the program compiled for the 68K Macintoshes (with and without FPU), and a version compiled (native) for the new PowerMac line. Because all arrays are dynamically dimensioned, the only limitation to the size of protein under examination is the memory allocated to the application. MolView is a greatly enhanced and rewritten version of a previous program called MacInPlot.[7] The program is currently ~25 000 lines of code (excluding the resource fork), is ~1 MB in size, and requires at least 4 MB of RAM and that QuickTime be installed. When working on larger proteins, the program requires about 8 MB of RAM. Clearly, there are practical considerations when trying to view very large proteins on the older, slower Macintoshes.

Because there are a number of options within the program, major aspects are outlined below.

USER INTERFACE

As shown in Color Plate 1, MolView has several floating windows that control the creation and aspects of the objects. The Tool window contains buttons that allow the user to perform nearly all functions without having to use the menu bar. The images on the buttons convey the functions that they perform: rotate the model, label the atoms using the mouse or a scrolling list of atoms, create or modify the

Color Plate for this article is on page 115.

Address reprint requests to Dr. Smith at the Department of Biological Sciences, Purdue University, West Lafayette, Indiana 49707.

Received 25 October 1994; revised 10 December 1994; accepted 13 December 1994

655 Avenue of the Americas, New York, NY 10010

0263-7855/95/$10.00
SSDI 0263-7855(94)00019-O

various types of models, measure distances, find neighbors, make and play QuickTime movies, scale the image, and define the center of rotation for the model.

The Object window allows the user to control the display and attributes of the various objects. For each object, there are a series of up to four buttons. The top button, containing either a logo or an abbreviation representing the object, toggles the object off and on. The button directly beneath displays the color of the object. By first selecting one of these buttons, then clicking on the color palette, the user can interactively change the color of the objects. Below this button is a box displaying a line of the same thickness as that used in drawing the object. When selected, the user can change this pen thickness. Finally, the bottom box displays and allows the user to change the pen pattern used in the object. In the case of ribbon, ball-and-stick, and *B*-value models, these three buttons are replaced with a preference button that allows the user to change many attributes all at once. Up to 10 molecular objects (MOL; a name originally used in FRODO[4,5]) can be created. As each MOL object is created, this window expands and displays a new set of buttons.

The Ribbon window controls the coloring and display of the ribbon diagram. In this window, each polypeptide is preceded by two buttons. The top button toggles the entire polypeptide off and on, and the bottom button allows the user to change the color of the entire polypeptide (when coloring by polypeptide). For each element, a button is displayed that has a picture of that particular type of structure (e.g., a loop, α helix, or β strand). This button allows the user to toggle the element off and on. Beneath each of these buttons is a colored button that contains a number. The number designates the first residue of that element and the color of the button represents the color of the element (when coloring by segments). By selecting this button, the user can change the color of the element. In addition, if the user holds down the Apple key while selecting one of these buttons, another window opens containing a button for each residue in that structural element. Using these buttons and the color palette, the user can color individual amino acids in any secondary structural element.

Up to two QuickTime movies can be displayed at a time. As is common with most QuickTime movie players, there are controls at the bottom of these windows that allow the user to play or step through the movies.

The main window can be resized and the image position can be controlled by scroll bars. All windows can be repositioned on the screen and a menu item forces the selected window to the front layer. The positions and sizes of the windows are monitored during the execution and then written to the preference file on exiting the program. This allows the user to personalize the window interface and have it appear that way each time the program is used.

As with other Macintosh drawing applications, the colors of the objects are controlled by a floating color palette. One hundred and forty-four colors are displayed in this window. Individual elements can be changed by holding down the Apple key while using the mouse to select a color button. This allows the user to utilize the full 32-bit color depth while allowing object colors to be changed rapidly. This color palette is stored in a preference file when the program is exited so that customization of the palette need only be done once. The user can revert back to the default palette by holding down the Control key while using the mouse to click on the color window.

TYPES OF DRAWINGS CREATED BY MolView

Stick figures

Several different types of stick figures can be created to highlight various aspects of the atomic model. (1) The C_α backbone can be traced for regions of the model; (2) all the atoms in a region of the model can be drawn and accented with different colors; (3) stick figures can be drawn in which each atom type is assigned a different color; (4) atoms of particular residue types within a region can be drawn; (5) balls of various radii can be drawn at desired C_α positions; (6) lines representing potential hydrogen bonds (between side chain atoms, main chain atoms, or both) within desired regions can be drawn; and (7) regions of the model can be drawn in which the line color is related to the *B* value (thermal factor) of the atom. The color, line thickness, and pen pattern of each object can be interactively altered using buttons in the Object window.

Ribbon drawings

As with other ribbon-drawing programs, this routine requires a TEXT file containing a list of secondary structural elements. However, because it was assumed that the user might be naive as to where these structures might exist in the protein, the program can generate this file. The program first calculates the ψ, ϕ values for each residue. Subsequently, each residue is examined for potential backbone hydrogen bonding between it and another residue. These ψ, ϕ values are then combined with the hydrogen-bonding information to determine if the residue is part of a helix or a sheet. For example, a residue in a helix should be hydrogen bonded to $n + 3$ or $n + 4$ residue, but a residue within a β sheet is probably not. This information is then used to find "runs" of secondary elements and the positions of these elements are written out to a TEXT file. In the default mode, this file is automatically created and used to generate the ribbon diagram. If available, this information can also be read from the header of standard PDB files.

There are a number of ways to color the ribbon diagram: by assigning a spectrum of colors (red to blue) to the ribbon as the chain proceeds from the first to the last residue, by assigning different colors to different polypeptide chains in oligomeric proteins, by assigning an overall color to the entire protein, or by color according to individual secondary structural elements.

The ribbon diagram can be toggled off and on using a button in the Object Palette. Beneath this toggle is a preference button that allows the user to change color schemes, apply a light source to the diagram, enhance the depth cueing by fading the colors to the background color, give the β-strand thickness, highlight (with color) individual amino acids, and alter the width of the ribbons.

CPK (space-filling) models

Space-filling models can be created using a variety of coloring schemes. Regions of the model can be colored by

atom type, residue range, residue type, or hydrophobicity. All of these color schemes are set within a dialog box activated by a button in the Tool window. To speed the drawing rate, spheres are not made to be intersecting but rather overlaid. Errors caused by this approximation are noticeable only at high magnification. The atomic radii for the various atom types are set within an editable TEXT file (names.pdb) that is read during application startup.

Ball-and-stick models

Ball-and-stick models can be created for regions of the molecule. The radius of each ball can be proportional to the atomic radius (set in TEXT file names.pdb), or set to an overall radius. The stick thickness is also adjustable. There are three coloring schemes: overall color, color by segment, and color by atom type.

TYPES OF STRUCTURAL ANALYSIS PERFORMED BY MolView

Ramachandran plots

The main-chain ψ, ϕ values for regions of the atomic model can be plotted onto a Ramachandran plot. This plot can be displayed in color or black and white, and can be written out to a PICT file. In addition, the user can use the mouse to pick points on the Ramachandran plot, following which the identity of the residue is displayed. The various regions of the plot are outlined. Proline and glycine residues are designated by special symbols in the plot.

Neighbors

The Neighbors option allows the user to identify neighboring atoms. The mouse is used to select an atom in the model, and lines are drawn between it and potentially neighboring atoms. The distances between the atom and the neighbors are displayed along the lines. The distance criteria for neighboring atoms, colors, pen thicknesses, and labeling options can all be set by the user.

Distance lines

The user can measure the distance between any two atoms and place a label at the midpoint of the line. These lines and labels become objects that can be modified and toggled on and off.

Labeling

For each atom, the residue number, residue type, and atom type are stored in memory. One can use any or all of this information (in any order) when labeling an atom. Atoms can be selected or deselected for labeling by picking the atoms from a list of radio buttons in a dialog box, or by using the Hand Pick Atom option in the Toolbox window. The font, size, and style of the lettering is controlled by menu items.

GENERAL OPTIONS

Background

The user can choose any background color.

Depth cueing

All of the various models and labels can be depth cued by fading the objects at the back of the image to the background color.

Stereo

The user can display either convergent or divergent stereo pairs. The separation between these images can also be adjusted.

Sorting

To speed the drawing rate, the user can toggle off/on a sorting function. In this way, all of the objects can be drawn back to front in order to maintain the proper three-dimensional (3-D) perspective. On the older, slower machines, it is advantageous to keep this option turned off until the diagram is ready for output to a PICT file.

Slabbing

The user can trim away the objects in all three dimensions in order to center on a portion of the model. The slabbing of each of the six sides of the viewing cube can be controlled independently.

Output size

The user can output images equal to the size of the window or equal to the size of the actual image (except when CPK models are displayed). When writing out an image that has an actual size greater than the viewing window, the user can effectively increase the resolution of the output file.

OUTPUT

PICT files

When a CPK model is displayed, the output PICT image is a 32-bit color copy of what is being displayed on the Main window. When a CPK model is not being displayed, then the PICT file is object oriented. When this PICT file is read by object-oriented drawing programs, each line and label can be independently manipulated.

Clipboard

The Macintosh clipboard can be used to interactively transfer the image in the Main window directly to any other Macintosh application.

Macmolecule files

Macmolecule is another freeware program used to display line, ball-and-stick, and CPK models. This option in MolView allows the user to export a file to Macmolecule without having to create the rather cumbersome input file manually.

CAD DXF files

DXF files are standard computer-aided design TEXT files. These files can be input to some of the rendering programs on the Macintosh for creating special surfaces, extrusions, and shading.

QuickTime movies

QuickTime movies are composed of a series of compressed images that can be played on any Macintosh that has QuickTime installed. These movies can be played within MolView or inserted into nearly every commercial word-processing and drawing program.

DISTRIBUTION

The program is distributed free of charge and is available through anonymous FTP sites at the University of Wisconsin (rhino.bocklabs.wisc.edu; this site is most frequently updated and contains several sample files) and the University of Michigan (mac.archive.umich.edu). It will also be distributed on a series of compact disks (CDs) called *Molecules of Life* (W.C. Brown Publishers, Dubuque, IA). This set of CDs will contain the program, ~200 protein structures, and many interactive animations that teach physical chemistry and structural biology. In addition, an interactive tutorial is available that is designed to teach students the main aspects of the program, using movies and examples. Additional inquiries and bug reports can be addressed to the author (e-mail, tom@bragg.bio.purdue.edu; FAX, 317-496-1189).

REFERENCES

1 Smith, T.J., Olson, N.H., Cheng, R.H., Chase, E.S., and Baker, T.S. Structure of a human rhinovirus-bivalent antibody complex: Implications for virus neutralization and antibody flexibility. *Proc. Natl. Acad. Sci. U.S.A.* 1993, **90,** 7015–7018

2 Liu, H., Smith, T.J., Lee, W.M., Leippe, D., Mosser, A., and Rueckert, R.R. The purification, crystallization, and structure determination of an Fab fragment that neutralizes human rhinovirus 14. *J. Mol. Biol.* 1994, **240,** 127–137

3 Zlotnick, A., Reddy, V.S., Dasgupta, R., Schneemann, A., Ray, W.J., Rueckert, R.R., and Johnson, J.E. Capsid assembly in a family of animal viruses primes an autoproteolytic maturation that depends on a single aspartic acid residue. *J. Biol. Chem.* 1994, **269,** 13680–13684

4 Jones, T.A. A graphics model building and refinement system for macromolecules. *J. Appl. Crystalogr.* 1978, **11,** 268–272

5 Jones, T.A. Frodo: A graphics fitting program for macromolecules. In: *Computational Crystallography* (Sayre, D., Ed.). Clarendon Press, Oxford, England, 1982, pp. 303–317

6 Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjelgaard, M. Improved methods for building protein models in electron density and location of errors in these models. *Acta Crystallogr.* 1991, **A47,** 110–119

7 Smith, T.J. MacInPlot II—an updated program to display electron density and atomic models on the Macintosh personal computer. *J. Appl. Crystalogr.* 1993, **26,** 496–498

8 Carey, J. and Lawson, C.L. Tandem binding in crystals of a TRP repressor/operator half-site complex. *Nature (London)* 1993, **366,** 178–182