MacMolecular: A program for visualization of molecular structures on the Macintosh

Timothy D. Weaver

Hinman 4328, Dartmouth College, Hanover, NH, USA

Suhail A. Islam

Biomolecular Modelling Laboratory, Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London, UK

David L. Weaver

Department of Physics, Tufts University, Medford, MA, USA

MacMolecular displays small- to medium-sized biomolecules, with particular emphasis on peptides. It has been developed to run on color Macintosh computers. The display can be stick, ball and stick, depth cued by thickness stick, or several types of space-filling representations. The program takes input from standard PDB files, simple Cartesian coordinate files, and, in addition, from Kinemage files in which atom information has been included. The program allows color changes of various types as well as the normal functions of translation, rotation, and zooming. In addition, animation files may be produced for subsequent display. Bonding of atoms is done by a distance algorithm (standard) or sequentially to properly display Cα traces and traces of peptides containing simplified representations of amino acids. Stereo viewing is available, and manipulated structures which were drawn from PDB files can be written out to new PDB files. In addition, PICT files of the drawing window can be generated.

Keywords: molecular graphics, Macintosh computer, peptide structure, PDB file, Kinemage file

Color Plates for this article are on page 200.

Address reprint requests to Professor Weaver at the Department of Physics, Tufts University, Medford, MA 02155, USA.

Received 15 July 1993; accepted 11 January 1994.

INTRODUCTION

The display and manipulation of molecular structures is of great importance in biochemistry, molecular biophysics and molecular biology, both for research purposes and for education. Because of the rapid advances in computer hardware, most laboratories have at their disposal one or more graphics workstations for displays and calculations involving polypeptides and nucleic acids. However, most professionals in these fields use personal computers for their word processing and most other day-to-day computational needs, and usually have a Macintosh or IBM PC (or clone) on their desk or at home. Thus, there exists a need for inexpensive software packages which can be used for realtime display of and interaction with biomolecules on personal computers. In addition, the commencement of a new journal, *Protein Science,* with its companion software package Kinemage[1] containing tools for the display of information relating to articles in the journal, is clearly just the beginning of efforts in this direction. The criteria for these types of software packages are ease of use; availability; the ability to do most molecular manipulations in (near) real time; multicoloring capability for highlighting aspects of structure; file input and output in standard form, i.e., the PDB[2] and Kinemage format,[1] as well as simple Cartesian coordinate input format; multiple display representations so that different aspects of structures can be examined; some form of depth cueing for three-dimensional perception of stick representations; some space-filling representations; animation capability to allow dynamical viewing of a particular structure; and general flexibility of the program.

Ideally, the software package should run on both PCs and Macs, but since programs are not instantly portable in most programming languages, especially when graphics is used, emphasis must be placed on one platform to start. In our case, the platform is a Macintosh with 8-bit color (the program will run with a BW monitor, e.g., a powerbook 170, but color is important for mole-

cule display) and a floating-point coprocessor.

In the next section, the program, called MacMolecular is described in detail, with respect to hardware required, programming language and algorithms, user interface data input and output, and molecular size capabilities. A final section contains some conclusions, and information about future developments and availability.

PROGRAM OVERVIEW

Hardware

MacMolecular runs on Mac IIs, powerbooks and Quadras in various memory and color configurations and a floating point coprocessor. The development system was a Mac II with 5 MB of RAM, a MC68020 CPU, a MC68881 floating-point coprocessor, and 8-bit color. The program runs under Mac OS 6.07 through System 7, the latter requiring more memory for the increased size of the operating system. The program has been tested on a Macintosh computer with MC68020, MC68030 and MC68040 CPUs. From these tests, we expect it to run fully on any Mac with 8-bit color and a MC68020 CPU or higher. The graphics resolution is 640 × 480. A subset of the functionality runs on systems with BW monitors (e.g., a powerbook 170).

Programming

The program is written in Think Pascal 4.0. This language implementation[3] was used because of its excellent user interface and debugging capabilities. As well as being easy to use, the Pascal language interfaces extremely well with the Macintosh Toolbox and Quickdraw routines. The compiled application has a size of 104K.

User interface

MacMolecular opens with three windows and a traditional Macintosh menu bar. All of these windows can be moved anywhere on the screen to fit the user's preference. In addition, the main drawing window can be resized like a normal Macintosh window. The main window displays the structure or structures on a black background (see Color Plate 1). There is a window labeled *Controls* which has buttons to control rotation about the *x*-, *y*- and *z*-axes, and translation along the axes. There are also two zoom buttons for changing the displayed size of the structures. The magnitude of the change is determined by the distance of the mouse from the center of the window. In this window, there is also a button to turn all labels on and off, and buttons labeled *all molecules* and *one molecule*, with one being highlighted in red, to choose the structures to which to apply the movement buttons. The third window, labeled *Hide/Show*, contains a column of structure names and check boxes which are used to turn the display of a particular structure on and off.

The details of the pull-down menus on the menu bar are shown in Table 1. The File menu contains the menu items for introducing a structure via the *Open PDB File, Open Kinemage File,* and *Open DATA File* menu items. The first allows the visualization of structure files in the normal PDB format,[2] the second allows viewing and manipulation of most structure files from the Kinemage[1] disks accompanying the new journal *Protein Science,* namely, those containing atom information (e.g., Color Plate 1 shows a "homemade" Kinemage file of alanine), and the third is a simple Cartesian file format. Since the Kinemage structures come in sets, all the members of the set are read in when a particular set is chosen, and the labels of the members of the set with associated check boxes are displayed in the *Hide/Show* window. The Help menu contains entries for the items in each of the menus.

Data input and output

As mentioned above, PDB, DATA, and Kinemage files are allowed for input into the program. As long as the original file was a PDB file, output can be to another PDB file incorporating any modifications of the orientation. An additional possibility is to save a picture of the structure in a PICT file format for later incorporation in drawing and painting programs or documents (Color Plates 2–5 are slides made from PICT files).

Animations in any of the representations can be made and later played back. This is a useful feature for examining a structure in one of the *z*-buffered representations where drawing can be slow, but once the animation is read in, playing the frames is relatively fast. The animation is created by generating a sequence of PICT files and a corresponding animation file that contains the number of frames in the animation. When the animation file is opened, the sequence of PICT files is displayed in order like the frames in a movie. The utility of this method of generating animations is that each of the individual frames within the animation are separate PICT files that can be printed or manipulated by many word-processing and drawing or painting software packages available for the Macintosh.

The program has the capability of displaying structures in which the bonding is sequential, e.g., Cα files or simplified representation files in which amino acid residues are replaced by a single center. For example, Color Plate 2 shows a depth-cued stick representation of a snapshot from a Brownian dynamics simulation of an artificial four-helix bundle containing 73 residues, each of which is represented by a single sphere.[4] Such displays require choosing a parameter file with appropriate radii using the *Choose Parameter File* item in the File menu and checking the item *In sequential order* in the dialog box that pops up when the first structure is read in. The parameter file is just a list of the names of the different types of atoms, their sizes, and a corresponding color. Keeping this information in a parameter file allows any labels, sizes, or colors to be used for each atom, making the input very flexible. Addition parameter files may be created with a word processor and saved as text files to address particular display needs.

Structure representation

There are many combinations of types of structures and display options in the Representation menu, as shown in Table 1. In addition, the Color menu allows many possibilities for highlighting particular features of a structure (see Table 1). For example, Color Plate 3 shows two different-colored orientations of the backbone of the A helix of myoglobin from the PDB file 1MBD,[2] residues 3 to 18, using the depth-cued stick representation. Color Plate 4 shows a close-up space-filling view of the PDB file 1PPT,[2]

Table 1. Representation of the MacMolecular menu bar. At start up, the items *Mono, Show main chain, Show side chains,* and *Show hydrogens* are checked and operative in the Representation menu, and the item *Color by type* is checked and operative in the Color menu.

File	Representation	Zoom	Rotation	Translation	Calculation	Color	Help
Open PDB File ...	Stick	Zoom + ...	Rotate about the *x*-axis	Move origin to center	Distance between atoms	White	
Open Kinemage File ...	Depth Cued Stick	Zoom − ...	Rotate about the *y*-axis ...	Translate-X ...	Angle between three atoms	Red	
Open DATA File	Ball and Stick	Reset clip plane	Rotate about the *z*-axis ...	Translate-Y...	Torsion angle	Green	
Remove Molecule ...	Space-Fill		Wobble			Blue	
Remove All Molecules	Flat Disks		Spin			Yellow	
Record Animation ...	Z-buffered 10 Shades		Set Origin			Magenta	
Play Animation ...	Z-buffered 32 Shades		Set origin COM			Cyan	
Create PICT File ...	Stereo					Color by type	
Create PDB File ...	Mono					Change Molecule Color	
Cover Full Screen	Show C-alpha					Change C-Alpha Color	
Choose Parameter File	Show main chain					Change Main Chain Color	
Set Movement Keys	Show side chains					Change Side Chain Color	
Quit	Show hydrogens					Change Hydrogen Color	
						Change Residue Color	
						Change Atom Color	
						Color All Residues Sequentially	

using *z*-buffering with 32 shades, and Color Plate 5 shows a flat disk representation of cysteine. The Zoom, Rotation, and Translation menus contain items for resizing and moving the structures, and for display purposes, there are *Wobble* and *Spin* items in the Rotation menu, as well as the items *Set Origin* and *Set origin COM* (center of mass) for choosing the point about which to perform rotations.

In addition, individual atoms can be labeled and unlabeled by picks with the mouse, and all labels can be turned on or off with a button in the *Controls* window. The Calculation menu has items for distance (two atoms), angle (three atoms) and torsion angle (four atoms) measurements. The Colors menu allows the user to change the colors of individual atoms or structures to aid in the visualization of the molecule or molecules displayed.

Structure size

The maximum number of atoms that can be viewed at one time is limited by memory. At the moment, all of the functions of the program can be performed with a limit of 800 atoms (atoms are no longer read in after this limit is reached), provided that the system has 2.5 Mb of free RAM with which to run MacMolecular. With more memory, more atoms could be viewed, but due to the limited size of the usual Mac color screen (13–14 inches), 800 atoms seem like a good maximum to avoid cluttering of the screen.

Kinemage manipulation

Choosing *Open Kinemage file* from the File menu item allows the reading in of most of the *Protein Science* structure files, in particular, those that have atom information included. Once read in, these structures may be translated, rotated, have the representation changed among the various stick and space-filling possibilities, have the colors changed and have true animations prepared involving rotation about any combination of the three cartesian axes. None of these capabilities are available using the MAGE program to display the Kinemage files. However, there is additional information (particularly textual descriptions) available in the Kinemage files which is not displayed by MacMolecular, so the programs are complementary.

DISCUSSION

MacMolecular has the features required for displaying and manipulating moderate-sized peptides, nucleic acid

fragments, and any molecule or crystal for which a PDB, DATA, or Kinemage file is available or can be produced. An appropriate parameter file is created if the standard one is not adequate. Additional features to be incorporated in the next major update are the capability to produce color postscript files of the displayed structures, and a build function to design peptides and nucleic acid fragments. Building structures will necessitate the introduction of an energy minimizer.

MacMolecular is copyrighted. Information on availability can be obtained from D.L. Weaver, Department of Physics, Tufts University, Medford, MA 02155; USA; email address: dweaver@jade.tufts.edu.

REFERENCES

1 Richardson, D.C. and Richardson, J.S. *Protein Science* 1992, **1,** 3–9

2 Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., and Tasumi, M. *J. Mol. Biol.* 1977, **112,** 535–542

3 *Think Pascal User Manual.* Symantec Corporation, Cupertino, CA, 1991

4 Yapa, K. and Weaver, D.L. *Biophys. J.* 1992, **63,** 296–299