ELSEVIER

Biochemistry and Molecular Biology Education 28 (2000) 297–300

BIOCHEMISTRY and MOLECULAR BIOLOGY EDUCATION

www.elsevier.com/locate/bambed

Articles

Bio-molecular modelling utilising RasMol and PDB resources: a tutorial with HEW lysozyme

J. Tony Pembroke*

Industrial Biochemistry Programme, University of Limerick, Ireland

Abstract

The availability of an excellent and freely available molecular visualisation package authored by Roger Sayle–RasMol, coupled to the protein data bank (PDB) resource of molecular structure data makes the visualisation and manipulation of molecular data accessible to the Internet community. Tutorials based on almost any protein deposited in the PDB can be utilised as an aid to understand bio-molecular modelling. The utility of these resources is illustrated for modelling and visualisation of HEW lysozyme. © 2001 IUBMB. Published by Elsevier Science Ltd. All rights reserved.

Keywords: RasMo; Tutorial; Lysozyme

1. Introduction

The detailed molecular diagrams and illustrations available in some of the excellent biochemistry textbooks currently available [1–3] can be accessible to anyone with internet capability through the use of molecular information and good visualisation and manipulation software packages which are freely available over the web. Individual students can utilise the visualisation programmes to investigate the 3D structure of proteins, zoom in on the active site of enzymes, observe groups at the active site or involved in binding or look at the molecular dimensions between groups on a protein. Primary data on 3D structure of proteins are available at the protein data bank (PDB), the single international repository for the processing and distribution of 3D macromolecular structure data, mostly obtained by X-ray crystallography or solution nuclear magnetic resonance (NMR). The empirical results of these experimental methods accurately describe the 3D structure of a protein molecule in the state in which measurements were made.

To illustrate the modelling capabilities of available programmes data on the 3D structure of the enzyme lysozyme will be used. Lysozyme is a peptidoglycan hydrolysing enzyme found in human tears and in hens egg white (HEW lysozyme). Even rather small proteins such as lysozyme contain very complex 3D structures which are difficult to imagine from 2D representations. To determine the 3D structure of proteins such as lysozyme the location of each atom in the macromolecule must be positioned [4], a process generally carried out via X-ray diffraction or solution NMR analysis. The data representing the 3D structure in its raw form is a set of co-ordinates representing the atomic co-ordinates of each atom of the protein with respect to a set of axes. These co-ordinates are given with respect to the Cartesian axes X, Y, and Z. Generally, because there are atoms of nitrogen, carbon, hydrogen, oxygen and sulphur in proteins the structure can become quite complicated even for small proteins. It is often the practice to draw only the α-carbon atom of each amino acid with the position of each carbon atom linked together giving a representation of the protein backbone. The PDB database contains files of such atomic co-ordinates which can be represented and manipulated via molecular visualisation software. For this tutorial some background on the composition, structure and function of the macromolecule is useful as a lead in. Such details on HEW lysozyme can be obtained in a range of good biochemistry textbooks.

2. Methods

2.1. Obtaining molecular data from the PDB

The PDB resource can be accessed at http://www.rcsb.org/pdb/. At the home page one can

* Tel.: + 353-61-202491; fax: + 353-61-202568.
E-mail address: tony.pembroke@ul.ie (J.T. Pembroke).

PII: S1470-8175(00)00050-3

interrogate the database using the PDB ID of the molecule (if known) or utilising the *Search Lite* facility using simple keywords. In this example click on Search Lite. This opens a window which requires search input. One can type in HEW lysozyme, click search and the search returns with five possibilities (as of May 2000). One of these is 1AZF. If one clicks on explore one reaches the structure explore window for 1AZF-HEW lysozyme. This gives details on the structure, the authors, the unit cell, number of residues, the atoms and the resolution, in this case 1.8 Å. One can view the structure at the PDB utilising the View menus but for this exercise, we wish to download the file. Click on download/display. Here one has an option of looking at the display, the structure file or clicking on text under the PDB heading. This will display a text edition of the data in the PDB file which contains the atomic co-ordinates. In the section, download the structure file choose the uncompressed pdb file by selecting it with the mouse, right click and select save target as. It should default to 1AZF.pdb any name can be chosen and indeed any location, however, it must be saved as a .pdb file. RasMol only reads .pdb files and therefore it is necessary to save all files from the PDB as .pdb files.

Alternatively at http://www.umass.edu/microbio/rasmol/pdblite.htm there is another search tool PDB lite authored by Eric Martz and Jaime Prilusky for non-specialists who wish to search the PDB database. There are mirror sites in the UK, USA, Israel and Australia which can be accessed from this site. If one types in 1AZF into the search window, the data on lysozyme is presented. If one then clicks on the section data retrieval, full instructions are given for downloading and saving the data files in Windows or Macintosh formats. There is a hypertext link to Open and 'Save as ..' dialog in the instructions. Clicking on this will allow the user to save to a specific disk location.

2.2. Obtaining RasMol

RasMol is a powerful educational tool for showing the structure of DNA, proteins and smaller molecules. The program reads in molecular co-ordinate files and interactively displays the molecule on the screen in a variety of representations and colour schemes. The programme was initially written by Roger Sayle and E. J. Milnerwhite [5] and developed at the University of Edinburgh's Biocomputing Research Unit and the Biomolecular Structure Department at Glaxo Research and Development, Greenford, UK. Information for getting and installing RasMol can be obtained at http://www.umass.edu/microbio/rasmol/getras.htm.

Various versions of RasMol can be downloaded for various operating systems including PC/windows, MAC, VAX, Unix, Acorn RISC OS. For example if downloading version 2.6 for PC or Windows one has the option of downloading a 32 bit version for Windows 95 or NT or higher or indeed a 16 bit version for windows 3.1 for those with less powerful machines. If one clicks on the download hypertext link one will be asked which option one requires. Upon clicking *Get 32 bit RAS Win* one will be asked whether one wishes to save to disk. The 32 bit version can be saved to a 3.5 in. floppy as it only occupies 341 kbytes of memory.

There are other molecular visualisation programs available, some free such as CHIME http://www.umass.edu/microbio/chime/index.html which was derived in part from RasMol, and other commercial programmes which can do more powerful representations. However, RasMol itself is considered quite powerful.

2.3. Working with RasMol

Quick start tutorials on how to use RasMol can be viewed at http://www.umass.edu/microbio/rasmol/rasquick.htm

To start RasMol under Microsoft Windows, double click on the RasMol icon in the program manager (or form the saved floppy). When RasMol first starts, the program displays a single main window (the display window) with a black background on the screen and provides the command line window minimised as a small icon at the bottom of the screen. The command line or terminal window may be opened by double clicking on this RasMol icon. The Command Line allows the user to define various parts of the bio-molecular structure one wants to interact with and will be used in some of the examples described later. To open a .pdb file within RasMol go to the file menu, click open. One can now open any of the PDB files one wishes having saved them from the PDB database. For the purpose of this example this will be the 1AZF.pdb file we saved earlier. The molecular structure should now be displayed within RasMol. You will notice, if you use the PC version, that the two RasMol icons are at the bottom one of these is the display, the other is the RasMol command line.

3. Results and discussion

3.1. Using and exploring the display options

Many different kinds of display options are available to model the .pdb file, some examples are outlined below:

1. Holding down the left mouse button rotate the molecule in any and all directions to visualise different parts of the molecule.
2. In the Display Menu alter the molecular display by selecting Space fill, Ribbons, Cartoons, Wireframe or Ball and Stick.

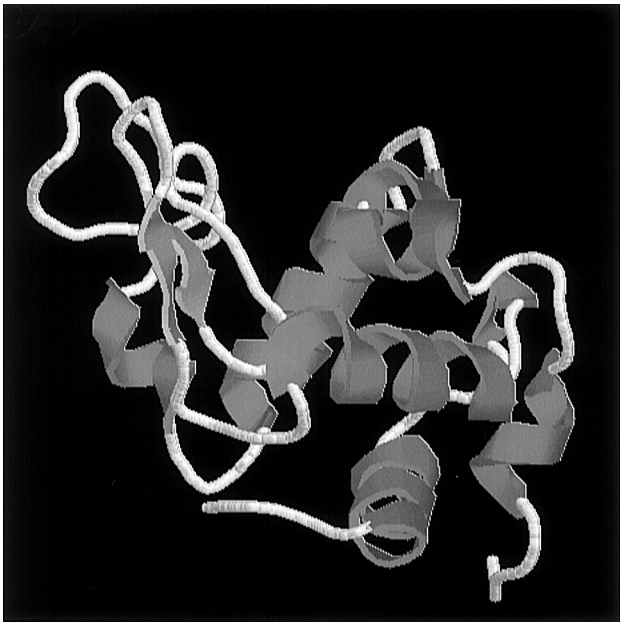

Fig. 1. A RasMol generated ribbon diagram of the 1AZF.pdb file illustrating the α-helical and β-sheet regions of the HEW lysozyme molecule.

3. Selecting the Display, Cartoons options select different colouring modes such as structure to view an outline of the macromolecule with α-helical regions and β-sheet regions (see Fig. 1).
4. Using the options menu select stereo to get a stereo view of the molecule.
5. Holding down the right mouse bottom allows one to move the molecule to different parts of the screen.
6. Holding down the shift key and moving the mouse forward miniaturises the display while moving it back blows up the display allowing the viewer to zoom in on the molecule. This can be particularly useful if one selects wire-frame within the Display option and setting Labels On within the Options menu. The positions of particular amino acids are clearly labelled and one can zoom into the molecule at these positions.

Any of the molecular representations can be saved and exported using the export menu and selecting the option that best suits, such as a gif or graphic interface file for future display in a presentation format such as Microsoft's Power Point. In the file menu if one selects information one gets a brief description of the molecule which can be useful if looking at lots of structures. From the file menu one can also print the display.

3.2. Using the command line options

Now for some of the really interesting things RasMol can do. Upon activating the RasMol command Line icon at the bottom of the desktop one enters the command line interface. At the RasMol icon (RasMol >) one can enter specific commands RasMol > ZAP clears display of a molecule if one wishes to load another. There are a number of commands which will allow better visualisation of the bio-molecule.

RasMol > Colour blue or red or other colours will colour the molecule

RasMol > Set background yellow (or other colours allow one to set coloured backgrounds).

From knowledge of the active site of lysozyme (see texts such as Voet and Voet 1995) it is known that a hexamer unit of the peptidoglycan polymer binds to lysozyme along a cleft within the molecule with the hexamer interacting at Asp 101 Trp 62 and Trp 63, Gln 57, Asn 44, Glu 35, Phe 34, Asn 37 and Arg 114. The catalytic residues involve Glu 35 and Asp 52. This part of the molecule can be visualised using the Command line interface.

RasMol > Restrict 34-55 command restricts the display to just that part of the molecule between the specified amino acids. One can zoom in on this section and use the options to label or vary the display options to visualise this part of the molecule more closely. In the case of lysozyme the active site and catalytic groups are contained here. By selecting wireframe display, label within options, and moving the fragment by holding down the right mouse button and rotating the catalytic groups within lysozyme Glu 35 and Asp 52 (the groups involved in lysozyme catalysis) can be visualised and their orientation relative to each other observed (Fig. 2). Using shift and pulling back the mouse one can zoom in to this part of the molecule.

Setting the crosshairs (the mouse pointer in the shape of a + sign) of the mouse on a particular bond or residue and clicking allows picking and the programme identifies the nearest atom to this point. Information displayed in the command line as a result might be CA 333, Thr 43 indicating one is pointing to the α-carbon of threonine 43 in the backbone. Using the up and down cursor keys one can review the history of the command lines one has used and revert to a previous command.

Within lysozyme the position of particular groups such as Glu 35 involved in catalysis can be located. Using the commands RasMol > select 35 return, RasMol > Colour yellow (or other colour) within the command line, and then from the display menu select space fill one can space fill the particular residue chosen so that its location can be observed within the molecule with ease. One can then restrict this section as above, zoom in or rotate to get a better view.

RasMol > *Set picking distance* allows the distance between two crosshair points chosen to be calculated.

RasMol > *Set picking monitor* and using the mouse to point to two points on the display allows calculation of the atomic distances between these groups or residues. At any stage particular views of a molecule utilising the commands outlined above can be saved and viewed later.

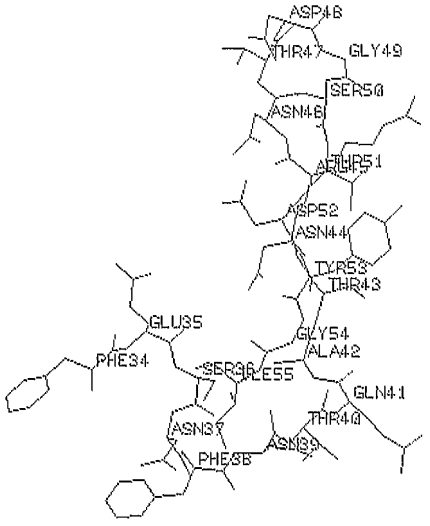

Fig. 2. A RasMol generated wireframe diagram of the active site region of the HEW lysozyme molecule with the region involved in catalysis restricted. The residues involved in catalysis Glu35 and Asp 52 and their position relative to each other can be observed.

3.3. Getting more information and exploring the potential of RasMol

A comprehensive list can be obtained from the RasMol manual. http://www.umass.edu/microbio/rasmol/distrib/rasman.htm and indeed detailed tutorials are available at http://www.umass.edu/microbio/rasmol/rastut.htm. This site contains a wealth of information about RasMol including questions and answers, galleries and tutorials. The site is maintained by Eric Martz emartz@microbio.umass.edu.

RasMol is a powerful visualisation and manipulation package [6] for display of a range of macromolecules including drugs, DNA and proteins. Using a tutorial such as outlined for lysozyme coupled with a detailed knowledge of its structure and catalytic mechanism available in biochemistry textbooks, the details of the groups involved in binding substrate and catalysis can be explored by individual students. In principle any macromolecule in the PDB or other macromolecular structural database could be investigated in a similar manner.

References

[1] A.L. Lehninger, D.L. Nelson, M.M. Cox, Principles of Biochemistry, Worth Publishers, NewYork, 1993.
[2] D. Voet, J.G. Voet, Biochemistry, 2nd Edition, Wiley, New Jersey, 1995.
[3] R.H. Garrett, C.M. Grisham, Biochemistry, 2nd Edition, Saunders College Publishing, London, 1999.
[4] K. Lim, A. Nadarajah, E.L. Forsythe, M.L. Pusey, Locations of bromide ions in tetragonal lysozyme crystals, Acta Crystallogr. D Biol. Crystallogr. 54 (5) (1998) 899–904.
[5] R.A. Sayle, E.J. Milnerwhite, RasMol—Biomolecular graphics for all, Trends Biochem. Sci. 20 (9) (1995) 374–376.
[6] R.M. Horton, R.J.S. Stone, An introduction to molecular visualisation: seeing in stereo with RasMol, Biotechniques 22 (4) (1997) 660.