Visualizing Molecules with VMD

Christoph Weber 3/7/2003

Aim of the course

After completing this course, participants will be comfortable and proficient with VMD. Participants will learn to display and manipulate molecules, change display styles for all or parts of the molecule and add annotations with the overall goal to create an image of their molecule ready for publication. Some prior knowledge of molecular modeling is helpful

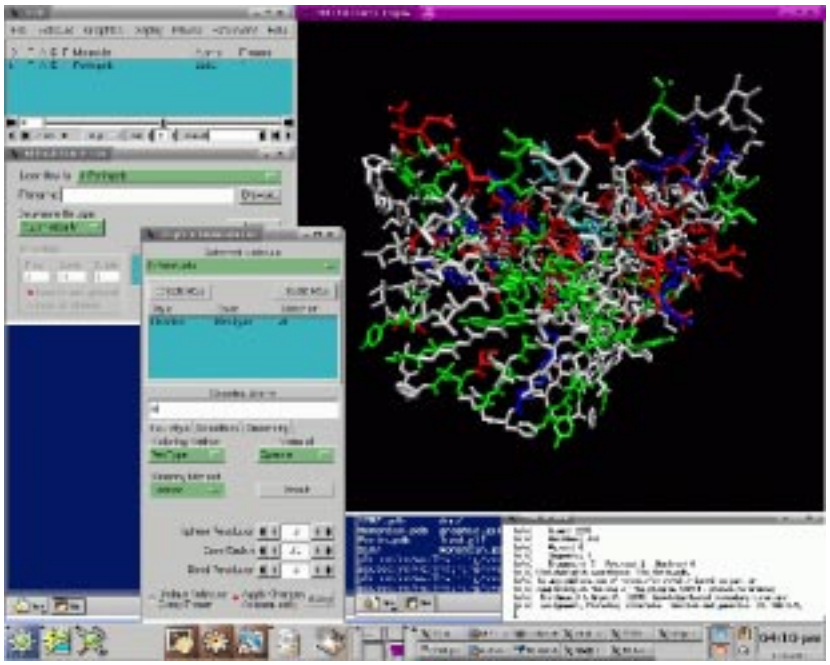
1.0 Introduction to VMD

1.1 Some Specifications

- No limits on the number of molecules, atoms, residues or number of animation frames, except available memory.
- Many (many!) molecular rendering and coloring methods.
- Stereo display capability.
- Extensive atom selection syntax for choosing subsets of atoms for display (includes boolean operators, regular expressions, and more).
- Reads PDB, Charmm and X-Plor style PSF topology, Charmm and X-Plor style DCD trajectory, Amber structure and trajectory (PARM and CRD) and Gromacs structure and trajectory files. In addition, through coupling to Babel and translation to PDB, VMD can import most molecular file formats.
- Ability to export displayed graphics to files which may be processed by a number of popular ray tracing and image rendering packages, including POV-Ray, Rayshade, Raster3D, and Radiance. VMD also outputs RGB image files and PostScript files for printing.
- Can by scripted with Tcl/Tk or Python. A selection of donated scripts for VMD is available at http://www.ks.uiuc.edu/Research/vmd/script_library/
- VMD runs on most Unix variants, as well as on Windows and MacOS X.
- Available for free at http://www.ks.uiuc.edu/Research/vmd/.

1.2 Layout and Organization of VMD

VMD opens multiple windows on your desktop. Initially, you get the main graphics window, the VMD console and the main form. VMD refers to its menus and dialog boxes as *forms*, so we'll use this term also.



In the image above, you can see the graphics window on the top right, the VMD console below it, and the main form on the top left. Also open are the Molecule File Browser and the Graphical Representations forms, along the left edge. The rest are various icons and windows from the Linux KDE desktop and are not related to VMD.

From the main form you can access all graphical interface features of VMD, i.e. open other forms which will let you drive VMD.
From the console window, you can issue VMD text commands and/or run scripts (as well as save them).

1.3 VMD Philosophy

It is important to note that while VMD can display any type of molecule, it is especially geared towards display of macromolecules. For smallish molecules it lacks good support to display double and triple bonds, largely a product of its reliance on PDB coordinates as input.

An important design feature of VMD is that you work with multiple graphical representations of your molecule. For example, you may display the whole molecule as lines (wireframe) and then display selected residues as spacefilling balls.These rep-

resentations are independent of each other, but hierarchical, i.e. the last one overdraws previous ones, if applicable.

1.4 Availability of VMD

VMD can be downloaded for free from http://www.ks.uiuc.edu/Research/vmd/
It is available for most Unix flavors, Windows and MacOS X.

At TSRI we have Unix versions installed under /tsri/$archosv/bin, where $archosv is sgi4DIRIX6, sun4SunOS5, i686Linux2, etc., and should be set in your .cshrc file.

1.5 Getting Help

VMD documentation is on-line at http://www.ks.uiuc.edu/Research/vmd/current/docs.html

In addition to web pages, the manual also comes in pdf format so that you can download and print it. Also available on-line are quick help topics and a FAQ. In particular, you can select Help on the main form, and this should point your web browser straight to the VMD Quick Help section of the VMD web site.

2.0 Some Sources of 3D Molecular Structures

You will need molecular coordinates to load into VMD. Unless you have your own experimental or modeled coordinates, here are a few resources:

2.1 Small Molecules

The Cambridge Structural Database (CSD) holds structures of organic and metallo-organic molecules. It also contains software tools - the main one being ConQuest - to search for molecules by name, substructure, author, etc. TSRI maintains a subscription for the CSD, which is updated twice per year. For more information and details on how to access the CSD see http://www.scripps.edu/rc/unix/sw/csd.html

The MDL Available Chemicals Directory (ACD) contains 3D structures of most commercially available chemicals. These aren't experimental, however, but rather derived from 2D sketches through software. Even so, for many purposes these 3D coordinates might be sufficiently good. Contact the library for details about access to MDL ACD.

See also http://molvis.sdsc.edu/visres/pdbs/titles.jsp

2.2 Proteins and Other Biomacromolecules

The Protein Databank (PDB) is the main resource for 3D coordinates of biological macromolecules. The PDB web site is http://www.rcsb.org/pdb/
There are extensive search functions at this site, as well as a host of other information.

TSRI keep its own mirror of the PDB at /tsri/pdb, accessible from all Unix hosts. If you know the PDB code of your protein, you can FTP to any TSRI Unix host using your email account, change directory to /tsri/pdb/struct and download what you need.

See also http://molvis.sdsc.edu/visres/pdbs/titles.jsp

3.0 Viewing and Analyzing Molecules and Preparing an Image for Presentation

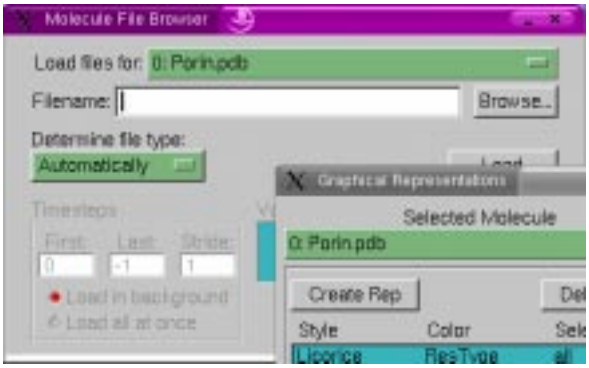
3.1 A Small Molecule - Monensin

Tasks in this section:
Loading a molecule, rotation, translation, scaling, atom picking, changing and adding graphical representations, saving application state, rendering output.

On Windows: Start VMD from the Windows *Start* menu. Select *Programs/VMD*.
On Linux or Unix: Start VMD by typing *vmd* in a terminal window.

You will get three windows: The main graphics window, the VMD console and the main form. Reposition the windows such that you can see all of the graphics window, all of the main form and most of the console.

Then select *File/Load Molecule* in the main form to open the Molecule File browser. In the Molecule File browser click on the *Load From Files* button. Another form opens.



Click the *Browse* button. In the resulting file chooser find and select the file *Monensin.pdb* and click OK. Now click *Load* in the Molecule File Browser. The structure of monensin is displayed in the graphics window. Also, the main form now has a line in its main area indicating various things about the loaded molecule. As is apparent, you can load more than one molecule into VMD. The limit is set by the sum of real and virtual memory of your computer.

Now move the mouse pointer into the graphics window. When you hold down the left mouse button and move the mouse you rotate the molecule about the x and y axis. When you hold down the right mouse button and move the mouse you rotate the molecule about the z axis. (If the mouse is in rotate mode, that is.)

Next, select *Mouse* in the main form and survey the options. This menu controls all mouse functions and you'll see that the mouse is in Rotate mode right now. If you select another mode and go back to the graphics window, you will find that the mouse buttons have different functions now. When you start VMD, the mouse is in rotate mode and picking is off as depicted by the absence of a highlighted radio but-

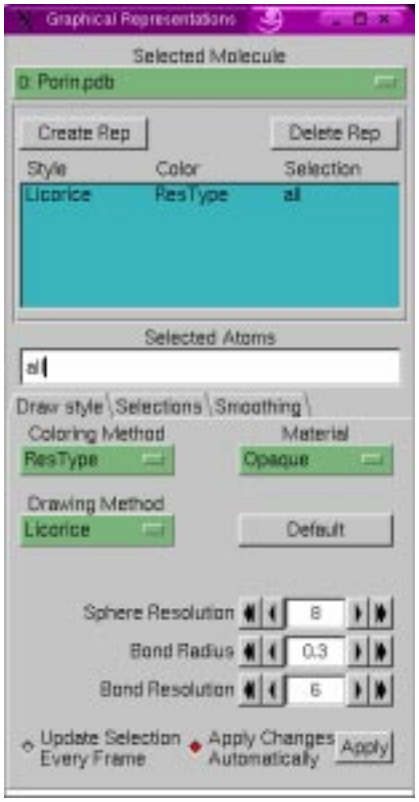
ton in the pick section of the menu. It is a good idea to select *Atoms* for Pick mode instead of nothing.

If a pick mode is on and you click a mouse button while hovering over an appropriate part of your molecule, the VMD console will display information about this part of the molecule and a label will be added in the graphics window. You might in fact want to glance at the console regularly to catch up on what VMD is doing.

Important: You can cycle between rotate, translate and scale mode with the *r, t* and *s* keys instead of going through the Mouse form. This saves a lot of time and will soon become second nature. Experiment with the different mouse functions now so that you get to know them. Hint: You can delete atom labels when they become too numerous by selecting *Graphics/Labels...* and taking the appropriate action. The Labels form also contains info about the picked/labelled part of the molecule.

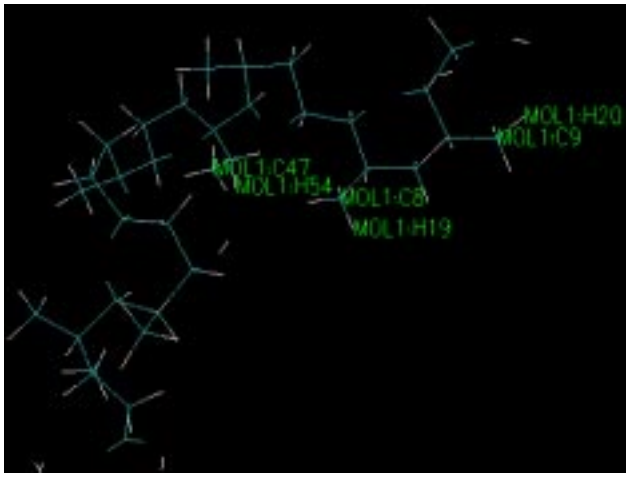
Finally, before the desktop gets more crowded, close all forms except the main form by clicking on their *close* button in the window title bar (usually an X near the top right).

Changing the display style (graphic representation)



Select *Graphics/Representations...* on the main form. The Graphical Representations form appears. This will most likely be your most used form. It lets you build up sophisticated displays in successive ‘layers’. Right now there is only one representation where all atoms are selected which are drawn as lines and colored as dictated by the atom names.

First, change the Drawing Method to *Licorice*. This draws bonds as round tubes and caps them with half spheres. This is a very nice and informative representation for smallish molecules. The Bonds method would be similar, except for the fact that terminal bonds are not capped and vertices are not closed, which makes Bonds virtually useless (in my opinion, anyway).



In the following we are going to highlight two of the methyl substituents and the ethyl substituent, pretending that these are important to make a point in a paper or presentation. VMD will require us to input atom names, so the first thing we need to do is find out about atom names in this structure. We will do this by picking them and noting the resulting names. Orient the molecule such that you can comfortably see the atoms and pick them as shown below. The full list of atom names is C8, H17, H18, H19, C9, H20 H21, H22, C46, H53, H54, C47, H55, H56, H57. Did you get them all?

Now click *Create Rep* in the Graphical Representation form. Replace the atom selection with *name C8 H17 H18 H19 C9 H20 H21 H22 C46 H53 H54 C47 H55 H56 H57* and hit the return key (Important! Otherwise the selection gets lost again.) The Coloring Method should be *ColorID* with the actual color being *4* (yellow). The Drawing Method should be *CPK*. Initially, not much will be apparent, because the yellow CPK representation is smaller than the licorice which is also still in effect for these atoms. Therefore, increase the *sphere radius* to 1.6 and *bond radius* to 1.2, at which point the yellow balls and sticks will have taken over the licorice. Note that VMD uses CPK for what is commonly referred to as ball-and-stick in other molecular viewers. The spacefilling balls for atoms with full van der Waals radii are called VDW in VMD.

Now we have a pleasing and informative display of monensin, ready for output. We can now do two things: 1) save the application state so that we can reload the molecule and regenerate the exact view and graphics representation, and 2) save an input file for a raytracing application to get a very high quality rendering of the molecule which can be printed and published (or shown at a seminar).

Saving the application state

In the VMD console window, type

```
save_state monensin.vmd
```

The file name, including extension are completely arbitrary.
To recreate the same view of monensin in your next VMD session, you will have to type

```
play monensin.vmd
```

You can also log all or part of your session with the command `logfile <filename>` typed in the vmd console. You can replay the logfile with `play <filename>`.

Generating input for raytracing

Click *File/Render* in the main form. From the list of available raytracers/rendering programs, select *Tachyon* and click *Go*. Two files are created: plot.dat and plot.dat.tga. The former contains instructions for tachyon about the scene to be rendered and can be deleted once you are happy with the result. The latter file, plot.dat.tga is a TARGA format image file and can be opened and manipulated with most image editors, e.g. xv.

Note that other than tachyon, none are installed automatically with vmd. You will have to make sure that the one of your choice is installed and in your path. We have render (part of the Raster3D package) available for SGI and Linux computers. Also make sure that you are happy with the suggested render command and options.

3.2 A Protein - Porin

New tasks in this section: Cartoon representation for protein, atom selection by properties, complex atom selection, pick (virtual) bonds

In the main form, select the *monensin* molecule and select *Molecule/Delete Molecule*. Select *File/Load Molecule* in the main form., and load the file *Porin.pdb* by following the same procedure as for monensin. Close the Molecule File browser.

Open the Graphical Representation form and change the representation for all atoms on porin: Make the Drawing Method *Cartoon* and the Coloring Method *Structure*. You get a secondary structure representation such as on the left.

With the change to the cartoon representation we lost the two detergent molecules that are part of this crystal structure. To display them, click *Create Rep* in the Graphical Representation form, type in *resname OTE TRS* for atom selection and hit return. Select *Name* for the Coloring method and *VDW* for Drawing Method. Now you can see the two molecules clearly.

Explorin' porin

Porin is a bacterial outer membrane protein that occurs as a trimer in nature. We are looking at a monomer right now. Let's look at the distribution of hydrophobic and hydrophilic residues in this protein: Create a new representation and type in *hydrophobic or neutral* for atom selection. Make the Drawing method *Line* or *Licorice* and the Coloring Method *ColorID 3* (orange). Create another new representation, type in *not (hydrophobic or neutral)* for atom selection, and change the color to *15* (light purple). The last two atom selections are examples of more complex selections, utilizing logical expressions. There are far more possibilities. Just check the Keyword list under the Selection tab in the Graphical Representation form.

Look down the central pore of porin and note the hydrophilic, charged residues that line the pore. Then check out how the outer wall of the beta barrel is studded with hydrophobic residues. Then look at the role reversal in the domain on top of the barrel and half inside the pore, which is mostly solvent (water) exposed. Can you also guess where the monomers interact to form a trimer?

Select *Mouse/Pick/Bonds*. Find two hydrophobic residues at the top and bottom of the beta barrel and click on one atom in each. A bond will be drawn between them with the distance labelled. The distance will be 19-22 A, about right for the thickness of a lipid bilayer into which this protein is embedded.

4.0 Exercises

4.1 Small Molecules - Estradiol and Testosterone

Load the file sexhormones.pdb into VMD. On the left you have estradiol, a member of the estrogen family of hormones, and on the right you have testosterone.

View and compare the two molecules all around. Figure out the differences between them, highlight them with a suitable display style so that you could easily show and explain the differences to an audience. Note that no double bonds or aromatic rings are drawn. You will have to infer these properties by the number of atoms bound to each carbon.
It might help to display a molecular surface and make that surface transparent to emphasize the steric difference(s) of the two molecules. Find out how to do this.

Save your work to a postscript file and print it. If your black-and-white print does not highlight the differences enough, you may want to go back and change to colors and/or display styles that will stand out in B/W.

4.2 Protein - Insulin

If you want to go all out, start a web browser and go to http://www.rcsb.org. Otherwise skip to the next paragraph. Search for *human insulin*. In the resulting hits, look for an entry at 1.60 Å resolution with the title containing '*crystallographic evidence for dual coordination around zinc*'. Download this file by clicking on the *Explore* link on the right, and then on the *Download* link on the left of the next page. On the actual download page, select the option to download in PDB format without compression.

Load the file 1TRZ.pdb. This is human insulin in complex with two zinc ions. Insulin consists of two chains, linked with several disulfide bonds. However, in this coordinate file there are four chains due to the way it crystallized, and there are also a good number of water molecules which got refined into the electron density.

1. Switch off the two additional chains C and D and the waters (chain X) and display only chains A and B.

2. Find a suitable display to highlight the architecture of the insulin backbone and the disulfide bonds linking them. Note that a bug in VMD traces ribbons and other, similar backbone representations wrong in this structure. I suggest that you display only backbone atoms and use one of the generic representations.

3. Display hydrophobic and hydrophilic residues in different colors and note the obvious concentration of the two groups in the core and on the surface, respectively.

4. If you have arrived at a nice display for either backbone architecture or side chain distribution, save it as a POV3 file for later input to the POV-Ray ray tracer.