

A short guide to the Visualization of Membrane Protein systems using *Visual Molecular Dynamics* (VMD)

VMD (Humphrey et al., *J. Molec. Graphics*, 1996) is a molecular graphics program developed by the Theoretical Biophysics group at the University of Illinois and the Beckman Institute. VMD was specially designed for the visualization and analysis of large biomolecular systems.

This guide contains two tutorial exercises whose goal is to introduce you to the use of VMD for the visualization of membrane protein systems. The VMD User's Guide provides a full description of VMD's capabilities for visualization and analysis of model molecular systems. Once you have gone through these exercises, you are encouraged to read the VMD User's Guide. All the information about VMD can be found at: <http://www.ks.uiuc.edu/Research/vmd/>

0. What do you need?

In order to view a single molecule or a full complex system in VMD you need to provide three kinds of information:

- Identifiers: Atom, molecule, and residue names, and their corresponding numerical indices.
- Atoms Coordinates.
- Bond Connectivity.

For a single molecule, or for an ordered system such as a protein crystal structure, you may allow VMD to figure out the connectivity. You will only need a file containing identifiers and coordinates. Usually this will be in the PDB format.

For a complex system, it is recommended that you provide VMD with all the connectivity information. In this guide such information will be in a file in the PSF format.

Molecules in motion. The main product of a molecular dynamics simulation is a trajectory, a set of system configurations taken at regular intervals over a specific period of time. For trajectories we will use the DCD binary format.

Summarizing: For a single molecule we will use a single PDB file.

For a complex system we will use a PSF file to specify topology and a PDB file for coordinates information. In the case of a trajectory, we will use a DCD binary file instead of the PDB file.

1. Visualizing a Membrane Protein Crystalline Structure:

Our first example consists on the visualization of the high-resolution crystalline structure of the CIC Chloride Channel (Dutzier et al., *Science* 2003. PDB ID 1OTS). The summary information for 1OTS, indicates that 3 molecules are present: the CIC itself and two distinct Fab complexes, each one of these is composed of two polypeptide chains.

Start VMD. On the VMD Main panel select Display -> Orthographic.

1.1. Load Molecule.

On the VMD Main panel select File -> New Molecule. The Molecule File Browser panel will appear. Make sure that Load files for: is set to New Molecule click on **Browse** and select the 1OTS.pdb file. At this point the **Filename** field should contain the correct filename (1OTS.pdb) with its path. Click on **Load**.

1.2. Mouse motion modes.

By default, after loading a new molecule, VMD renders all the atoms as lines colored by chemical element. Also by default, a set of Cartesian axes is represented in the

lower left corner of the display, oriented with the z-axis outward and the y-axis upward. Notice that this set of axes corresponds to the atoms coordinates.

Select the **OpenGL Display** window and type `r` . This will set the mouse in rotation mode (this is the default setting). Rotate around axes parallel to the screen by holding the left button down and moving the mouse. Rotate around the normal to the screen by holding the right button down and moving the mouse left and right.

Select the **OpenGL Display** window and type `t` . This will set the mouse in translation mode. Translate on the plane of the screen by holding the left button down. Translate along the normal direction to the screen by holding the right button down and moving the mouse left and right.

Select the **OpenGL Display** window and type `s` . This will set the mouse in scaling mode. Moving the mouse to the left while holding down either button will shrink the molecule, moving to the mouse to the right while enlarge the molecule. Using right button scales faster than using the left button.

Here is a summary of mouse motion modes.

mouse mode	left button	right button
rotation	axes parallel to screen	axis normal to screen
translation	axes parallel to screen	axis normal to screen
scaling	slower	faster

1.3. Changing Representations.

It is not easy to identify macromolecules in the default rendering. Usually, the first representation of interest for proteins is either according to secondary structure or rendering of the backbone only. We will explore both cases.

On the VMD Main panel select **Graphics -> Representations**. The **Graphical Representations** panel will appear. We will describe this panel in detail as it contains the main operations and commands for visualization.

The top field **Selected Molecule** should indicate the name of the loaded file and the Molecule ID. This information coincides with the one displayed on the VMD Main panel window.

Notice that only one representation is listed and that all the atoms are selected. The main attributes of a representation are **Style** and **Color**. As indicated before, by default VMD will render atoms as lines (**Lines in Drawing Method**) and will color them by chemical element (**Name in Coloring Method**). Can you guess what are the colors that VMD uses for carbon, oxygen and nitrogen?

Change the **Drawing Method** to **Cartoon** and the **Coloring Method** to **Chain**. Rotate around until you reproduce Figure 1.

It is customary, for structures deposited in the PDB, to identify each polypeptide chain with a specific letter. By coloring by polypeptide chain we are able to identify the six components indicated above. The CIC chloride channel is a homodimer, corresponding in this case to the polypeptide chains colored blue and red.

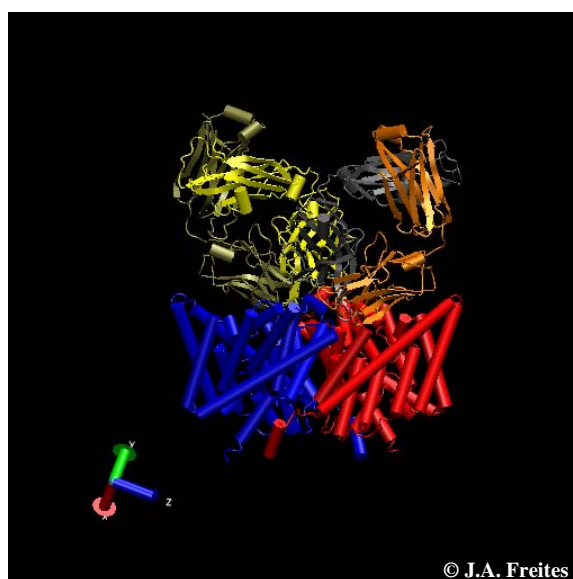


Figure 1

The **Cartoon** representation corresponds to secondary structure as identified with STRIDE (Frishman and Argos, *Proteins*, 1995). α -Helices are represented as cylinders, beta sheets as flat ribbons, and the other conformations as thin tubes.

We will now focus on the chloride channel. Delete `a11` from the **Selected Atoms** text field, type `chain A B` and hit Enter. This should leave only the chloride channel displayed. Notice that this and any other changes performed in the **Graphical Representations** panel assume that **Apply Changes Automatically** is selected otherwise you need to click on the **Apply** button every time.

Magnify and center the channel. Try different representations of the backbone by changing the **Drawing Method** to **Ribbons**, **NewRibbons**, and **Tube**.

Go back to **NewRibbons** representation. We will now represent some protein residues explicitly.

Basic residues are less likely to be located in the hydrophobic region of the membrane, let us test that assumption for the CIC channel. Make sure that you have only one representation and that it is highlighted in the window of the **Graphical Representation** panel, the parameters should read **NewRibbons / Chain / chain A B**. Click on the **Create Rep** button. This copies the highlighted representation. The active (highlighted in the window) representation is now the second one. You can make any representation active by clicking on its entry line in the panel window.

Edit the **Selected Atoms** text field by typing and `basic` right after `chain A B`. The backbone representations (**Ribbons**, **NewRibbons**, and **Tube**) thread the locations of the alpha carbons, so adding the `basic` keyword did not produce any effect. Change the **Drawing method** to **VDW** (for Van der Waals or space filling spheres representation) and the **Coloring Method** to **ColorID**. Select **yellow** in the middle pull-down menu. As expected the basic residues are distributed towards the connecting loops and away from the transmembrane α -helical domains. Located in this fashion they, presumably, will be in contact with lipid headgroups and waters.

To produce more contrast between the two representations, let us change some parameters of the **NewRibbons** representation. Make the **NewRibbons** representation active. Increase the ribbon thickness to 0.60 by moving the **Thickness** sliding bar to the right. Change the ribbon **Resolution** to 16 by clicking twice on the double right-arrow. Compare your result to Figure 2.

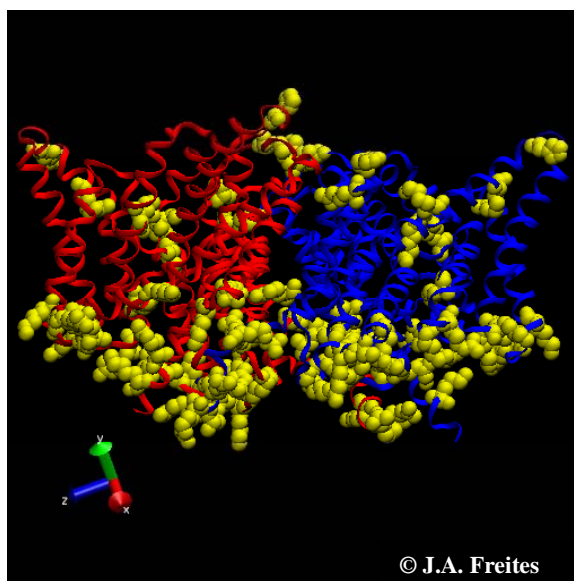


Figure 2

In the previous example we have made use of one keyword (basic) and one Boolean operator (and). All the possible keywords and some operators are listed under the **Selections** tab. Click on the **Selections** tab and verify that the keyword basic is among the choices in the **Singlewords** window. Keywords that require specific values are listed on the **Keyword** and **Value** windows. You can left-click on a specific keyword name to get the allowed values. Verify that, as observed before, there are six polypeptide chains as possible values for the **chain** keyword (A to F, the X chain identifies the so-called “heteroatoms” see the PDB website for more details). Verify that the possible values for residue names (resname) are the usual three-letter codes for amino acids plus CL for the chloride atoms presented in the channel. By double left-click you could transfer any word from the **Singlewords**, **Keyword** and **Value** windows to the Selected Atoms text field. There are also three buttons for the **and**, **or**, and **not** Boolean operators.

As a final exercise with the CIC chloride channel, let us build a third representation with a specific amino acid residue. Without giving detailed instructions, here is what is needed to obtain Figure 3.

Change the NewRibbons representation of the channel secondary structure to Cartoon. Change the coloring to lime as a single color. Change the Coloring method for the basic residues to ResType. Create a new representation and edit the Selected Atoms text field to read chain A B and rename TRP. Color by ResType. Compare your result to Figure 3.

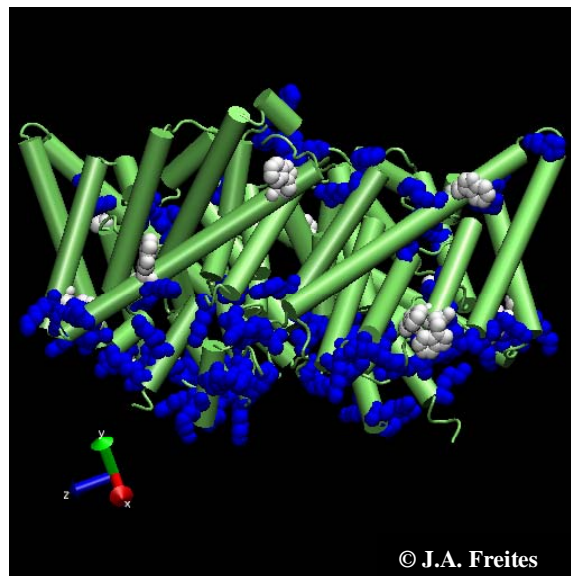


Figure 3

The tryptophan residues are located at the end of the α -helical transmembrane domains, where they will be, most likely, in contact with lipid headgroups.

1.4. Coloring Control.

Coloring by residue type implied a specific color key, here we mention briefly how to verify and change VMD default color keys.

On the VMD Main panel select **Graphics -> Colors**. The **Color Controls** panel will appear. Click on **Restype** in the **Categories** window and verify the possible options for amino acids in the **Names** window (**Basic, Acidic, Polar, and Nonpolar**) these are all VMD single-word keywords. By clicking on the specific names, verify that basic residues are colored blue and that nonpolar residues (including trp) are colored white.

You may change these choices by clicking on a different color in the window located in **Color Definitions** tab. You may also change the actual color with the **RGB** sliding bars.

1.5. Saving Your Work.

A simple way to save a set of specific representations, such as the ones made here, is to save the VMD-state. This will create a text file with a series of instructions that VMD can follow to automatically recreate the current state of your work.

On the VMD Main panel select **File -> Save State** type a meaningful filename and save. Start a new VMD session and select **File -> Load State** on the VMD Main panel. Load your saved state.

2. Visualizing a Membrane Protein Model System.

We will perform the visualization of a molecular dynamics simulation of the CIC chloride channel based on the structure described in the previous example (PDB ID 1OTS). The simulation system consists of the two polypeptide chains of CLC (with hydrogens added), 767 POPC molecules, 40548 water molecules, 4 structural chloride ions, and 8 chloride ions playing the role of system counterions. The total number of atoms in the system is 238022. This simulation system can be considered “large” for today’s standards, but at the same time, it is not too far from the smallest size

necessary to study a full biological unit (i.e. the CLC chloride channel) in a membrane model environment (i.e. a fully-hydrated POPC lipid bilayer).

Start VMD. On the VMD Main panel select **Display -> Orthographic**.

2.1. Loading a trajectory.

As mentioned before (see Section 0), when working with complex systems, it is advisable to provide VMD with all the required information. In this case, we will load two files both named “CLC_simulation”. One is a PSF file, containing identifiers and connectivity information. The other one is a DCD file containing 10 system configurations.

On the VMD Main panel select **File -> New Molecule**. The **Molecule File Browser** panel will appear. Make sure that **Load files for:** is set to **New Molecule**. Make sure that the **Determine file type:** pull-down menu is set to **Automatically**. Click on **Browse** and select the CLC_simulation.psf file. At this point the **Filename:** field should contain the correct filename with its path. Notice that the **Determine file type:** pull-down menu has changed to **psf**. Click on **Load**.

Go to the **VMD Main** panel and verify the information displayed in the panel window. The **Molecule** field should read **CLC_simulation.psf**. **Atoms** should read **238022**. **Frames** should read **0** since no configurations have been loaded yet. Select the **OpenGL Display** window. No coordinates have been loaded yet, consequently, there is nothing to render.

Go back to the **Molecule File Browser** panel. Make sure that **Load files for:** is set to **#: CLC_simulation.psf**, where the **#** corresponds to the ID number indicated in the **VMD Main** panel window for this molecule. Make sure that the **Determine file type:** pull-down menu is set to **Automatically**. Click on **Browse** and select the CLC_simulation.dcd file. Notice that VMD has detected that your file is a DCD file, and the **Frames:** section of the **Molecule File Browser** panel has now been

activated. You have the choice of loading frames in the background, or dedicate all the resources to loading frames by selecting **Load all at once**. In the present case we will choose the default parameters. VMD will render each new frame as it gets loaded. Click on the **Load** button. Select the **OpenGL Display** window. Observe the new configurations as they are loaded. Go to the **VMD Main** panel the **Frames** entry should increase until it reaches **9**. VMD starts the frame count at zero.

2.2. Overall system configuration.

You are now observing the whole system from the top. The system is contained in a imaginary parallelepiped (not shown) and it posses periodic boundary conditions. Then, it is a single cell that is being rendered. Rotate the system 90° about the x-axis so that the z-axis is pointing now upward and the y-axis inward.

The system appears to be constituted by three layers. Where is the protein? Let us identify the main system components.

On the **VMD Main** panel select **Graphics -> Representations**. As before, on the **Graphical Representations** panel window there should be only one representation and the **Selection** field should read **all**. In the **Selected Atoms** text field delete the keyword **all** and type **protein**. Since you are already familiar with the CIC channel tertiary structure, you can identify it even in the current line representation. The view is similar to those presented in Figs. 1 to 3.

Create a new representation. Activate it and change the **Selected Atoms** text field by deleting **protein** and typing **water**. Now you can see that the top and bottom layers are water. What is the color that VMD uses for hydrogen?

Create a new representation. Activate it and change **Selected Atoms** text field by deleting **water** and typing **lipid**. This will get you back the “whole” view of the system. The protein is embedded in the lipid bilayer, as expected.

2.3. Protein-in-membrane visualizations.

The first observation that we will like to make is about the location of the membrane protein. Let us produce a cut-away view of the bilayer and water components while leaving the protein intact. Make sure that the system is oriented as indicated above, with the z-axis pointing upward and the y-axis pointing inward. On the **Graphical Representations** panel, activate the lipid selection and edit the **Selected Atoms** text field by adding and $y > 0$. Verify that the Selection field in the panel window reads now lipid and $y > 0$. This phrase is interpreted by VMD as: “display only those atoms belonging to lipid molecules and with a y-coordinate greater than zero”. VMD evaluates the logical sentence automatically, provided that **Apply Changes Automatically** is selected; otherwise you need to click on the **Apply** button every time. Proceeding in a similar way, make VMD display only those atoms belonging to water molecules and with a y-coordinate greater than zero.

Since all three representations are being rendered in the same way, it is still difficult to identify the protein. Activate the protein selection and change **Drawing Method** to **Cartoon**. The location of the protein is now obvious. Magnify the view, focusing on the protein. The Cartoon representation depends on the position of the α -carbons, so that by default when coloring by name in the Cartoon representation, the result is a single cyan coloring. There is very little contrast between the protein secondary structure and the bilayer hydrocarbon region. Let us change the coloring method of both water and lipid components. Activate the water selection and change the **Coloring Method** to **ColorID** and select **white** in the middle pull-down menu. Activate the lipid selection and change the **Coloring Method** to **ColorID** and select **orange** in the middle pull-down menu.

We could enhance the contrast even more by rendering waters and lipids with different representation methods. Activate the water selection and change **Drawing Method** to **Licorice**. Activate the lipid selection and change **Drawing Method** to **VDW**. Compare your result with Figure 4.

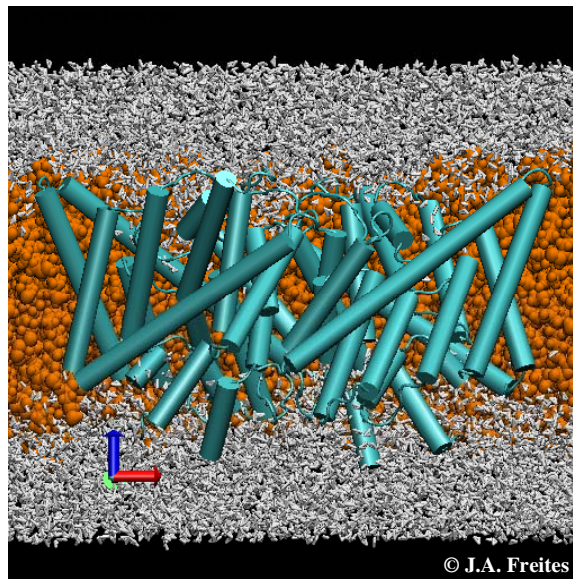


Figure 4

For a large system such as the present one, you may have noticed that combining complex representations, such as those of Figure 4, could slow down the program. To expedite your work with VMD, keep the number of *displayed* representations to a minimum or choose the **Lines** drawing method whenever possible. To hide a representation, double-click on its entry in the **Graphical Representations** panel window. You can test this by attempting a rotation while displaying the system as in Figure 4 and comparing the same operation while displaying only the protein.

As expected, the CIC channel spans the lipid bilayer. Except for the four end-terminal domains, all the α -helical domains are located within the lipid bilayer. Each polypeptide chain of the CIC Channel contains an independently gated anion conduction pore. Figure 4 shows water molecules penetrating the vestibules of both polypeptide chains.

2.4.Distance-based selections.

Given these observations, it would be of interest to visualize only those waters interacting directly with the protein. The rigorous way to do this would be computing the pair distribution function for specific protein atoms and identify water molecules in the first coordination shell. How could we obtain a similar result with VMD?

Start by hiding or deleting the lipid selection. Activate the water selection and edit the **Selected Atoms** text field by replacing `and y > 0` with `within 4 of protein`. Verify that the **Selection** field in the panel window reads now **water and within 4 of protein**. This phrase is interpreted by VMD as: “display only those atoms belonging to water molecules and whose distance to *any* protein atom is less or equal than 4 Å”. Change **Coloring Method** to **Name** and **Drawing Method** to **VDW**. Compare your result with Figure 5.

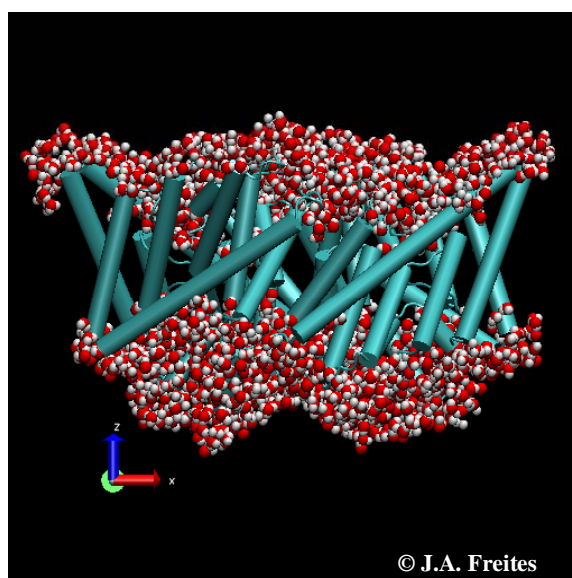


Figure 5

The representation contains isolated hydrogen and oxygen atoms. This is due to the ambiguity of the keyword **water** (i.e. *any* atom belonging to a water molecule). There is also a slight inaccuracy on the distance criterion, since it does not distinguish between pairs of hydrogen-no hydrogen, hydrogen-hydrogen, and no hydrogen-no hydrogen. Modify the sentence so that the **Selection** field in the panel window reads now: **water and same residue as name OH2 and within 4 of protein and noh**. VMD interprets this phrase as: “display all the water molecules whose oxygen atoms

are within 4 Å of any protein atom except hydrogen”. We will now deconstruct that phrase.

In this system, each water molecule is identified as a residue, just in the same way that is done for each aminoacid of a polypeptide. That accounts for the use of the **same as** selection method with the keywords **water** and **residue**. We are excluding from the selection all the water-protein atom pairs containing hydrogen atoms. This is being done in two ways: for the waters, by specifying the **name** of the water oxygen (name **OH2**), and for the protein by adding the keyword **noh** which is short for no-hydrogen. The result is shown in Figure 6.

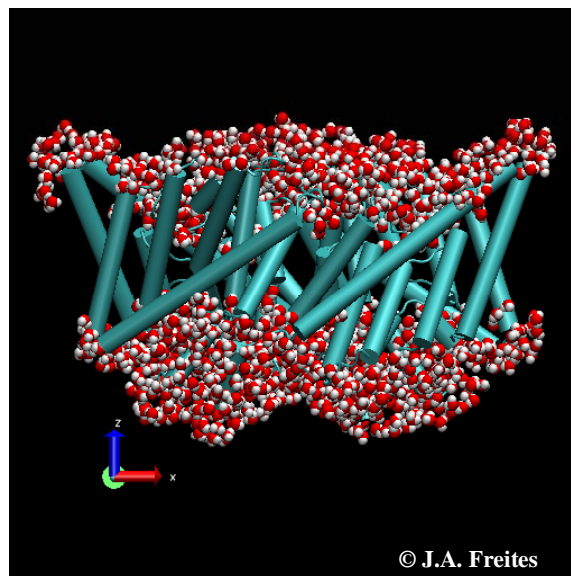


Figure 6

2.6. Atom querying.

The last selection included the name for the water oxygen, a specific identifier. It seems obvious that, in the VDW representation and colored by Name, the water oxygen is the red sphere. But what are its identifiers?. How to obtain all the identifiers associated to a specific atom on the visualization?.

On the VMD Main panel select **Graphics -> Labels**. The **Labels** panel will appear. Make sure that the pull-down menu is set to **Atoms** and that the **Picked Atom** tab is selected. Select the **OpenGL Display** window and type 0 (zero). This will set the mouse in query mode. Obtain the information about a single water oxygen by left-clicking on any red sphere. An example is shown in Figure 7.

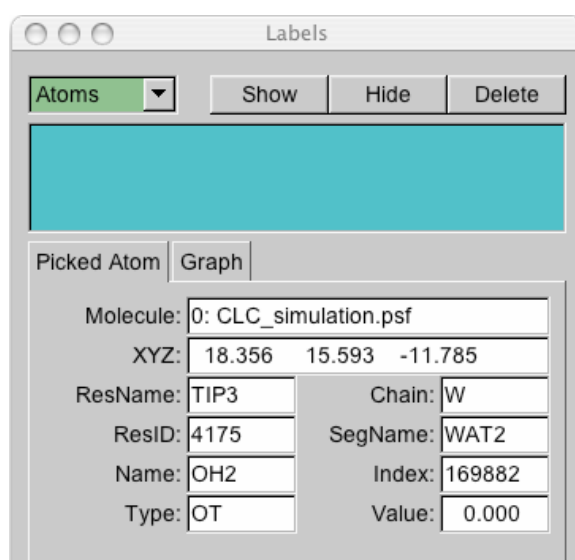


Figure 7

The description of each field is as follows:

- Molecule: provides ID number and name, as it appears on the VMD Main panel window.
- XYZ: the atom Cartesian coordinates.
- ResName: the atom belongs to the residue of this name. In this case, provided in the PSF file. The particular name TIP3, corresponds to the model of water employed in the simulation.
- ResID is an internal index (i.e. provided by VMD) that runs over all the residues belonging to a single segment.
- Name: the atom name as provided, in this case, in the PSF file.
- Type: the atom type, in this case, according to the CHARMM force field.
- Chain: the same identifier employed in PDB files to identify polypeptide chains or specific group or type of molecule.
- SegName: the atom belongs to the segment of this name. Similar to the Chain identifier, a segment usually identifies a specific type of molecule or a single polypeptide chain. Provided, in this case, in the PSF file.
- Index: is an internal index (i.e. provided by VMD) that runs as a single sequence over all the atoms in the system. Usually, this sequence corresponds to the order in which the atom names were read by VMD. The sequence starts on zero.
- Value: the value of some simple geometric measurements that VMD can perform. Not used in atom query.

2.8. More on selections.

As a last exercise on atom selection, let us contrast the protein regions that are in contact with the lipid bilayer and with chloride ions located at specific binding in the selectivity filter. Dutzier et al. (*Science*, 2003) located three different ion-binding sites in their high-resolution structure of the ClC channel. Chloride ions at the so-called S_{in} and S_{cen} binding sites were included in the simulation and were kept highly constrained about the crystalline positions. The S_{cen} sites are located next to the channel gate, whereas the S_{in} are located at the entrance of the selectivity filter right after the vestibule.

Delete all your representations by clicking on the **Delete Rep** button. In the Selected Atoms field type `segname CHLO`. This will select the four chloride ions. Choose **VDW** representation. To render a full coordination shell of no-hydrogen protein atoms for these ions, create a new representation and type `protein and noh and within 6 of segname CHLO`. Choose **VDW** representation. Create a new representation. A reasonably continuous region of lipid-protein contacts between no-hydrogen atoms can be obtained by typing `protein and noh and within 5 of lipid and noh`. Again, choose also **VDW** representation. Color the second selection green and the third selection orange. Let us also create continuous smooth surfaces for those two representations, select **Surf** as drawing method for both representations. VMD will employ the program SURF (Varshney et al., *IEEE Comp. Graphics and Applications*, 1994) to render a so-called solvent-accessible surface. Compare your result with the top-view shown in Figure 8. This image suggests that the somewhat complicated tertiary structure of the ClC channel is capable of providing the right environment for the selectivity filter while at the same time shielding it from the hydrophobic environment.

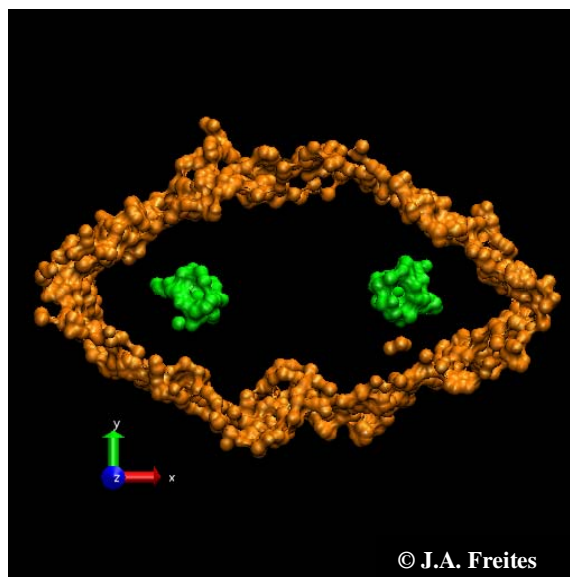


Figure 8

2.8. Animation.

Let us now work with all the ten configurations that were loaded at the beginning of the exercise. These configurations were taken from the simulation every 500 ps, i.e. the ten frames span a total of 5 ns. The speed with which VMD can playback a series of frames depends on the complexity of the image being rendered, in terms of the drawing method employed, number of atoms, and the kind of selection performed.

Let us first observe the motion of the protein. Create a single representation of the protein using the Cartoon drawing method. The playback controls are at the bottom of the VMD Main panel, they are similar to those found in any media player. Play forward by clicking on the button with a single right-arrow at the bottom right of the panel. Stop by clicking the button again. The button with the bar-arrow icon will advance step-by-step. Notice the difference on fluctuations of the end-terminal domains compared to the rest of the protein.

We will now contrast the lipid molecule fluctuations by observing two single lipid molecules at the same time. Create a new representation and type `resname POPC` and `resid 291 760`. The lipid molecule with resid 291 is the one next to the protein, the lipid molecule with resid 760 is as far from the protein as possible in this

simulation system (see Figure 9). The average separation between these two lipids phosphorus atoms along the trajectory is 143 Å. Notice in particular the larger headgroup fluctuations for the lipid far from the protein. Observe also, for both lipids, the great variety of conformations that the acyl chains assume, as is characteristic of the liquid crystalline phase of lipid bilayers.

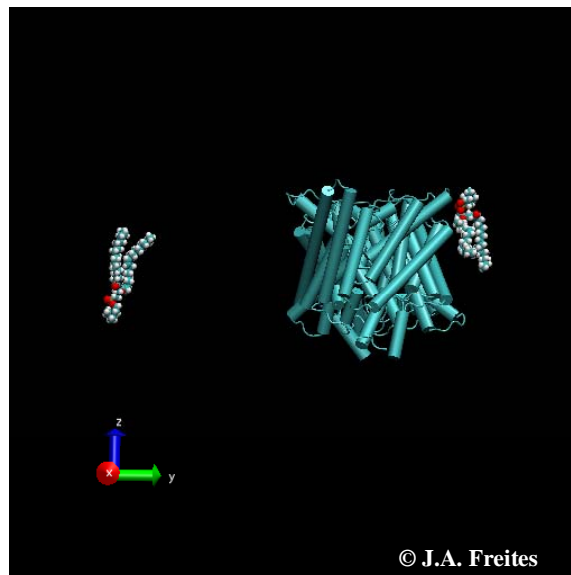


Figure 9

Stop the animation. Hide or delete the protein representation. Create a new representation with drawing method VDW and type water and within 4 of resname POPC and resid 291. This should give you what appears to be the hydration shell of the lipid molecule. Start the animation, what do you observe? You should observe the waters flying-away and converging back towards the lipid headgroup. By default, VMD will perform the selection for the initial frame only, what you are observing is the motion of those specific waters that were near the lipid at the beginning of the animation. To instruct VMD to perform the selection every frame, stop the animation, go to the **Graphical Representations** panel select the **Trajectory** tab, and choose **Update Selection Every Frame**. You should observe now a hydration shell on every frame. Since VMD needs to perform the distance calculation every frame, the playback is slower.

Let us conclude with a glimpse on VMD measurement capabilities. We will determine how far two bulk waters can get from each other in 5 ns. To start, make sure that you are working on the first frame (frame 0), drag the sliding rule all the way to the left on the VMD Main panel until you see the 0 appear on the small window next to the sliding rule. Delete all your representations. Create a new representation and type `segname WAT4` and `resid 8178 2417`. You might need to use the VDW drawing method to locate the two waters, once you have located them, change Drawing Method to CPK (balls and sticks). Select the OpenGL Display window and type 2. Click on both oxygen atoms, as you click on each oxygen, you should see a green label appear and then the distance between both atoms as a white label, in this case, 2.86 Å (see Figure 10).

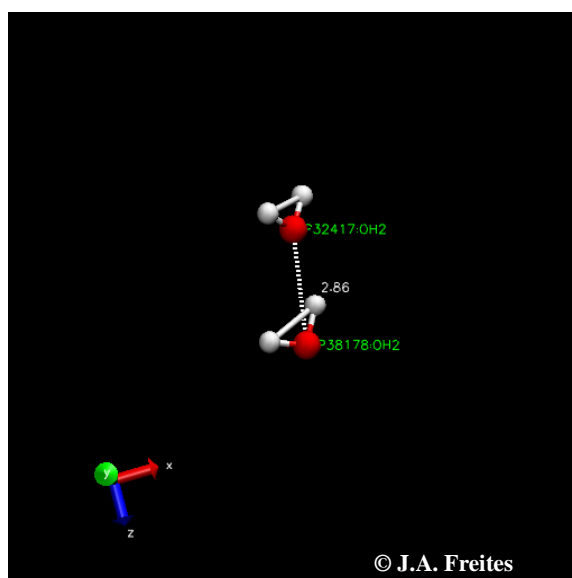


Figure 10

Go to Graphics -> Labels. The Labels panel will appear. If **Atoms** is selected in the pull-down menu, you should have two entries in the panel window, corresponding to the two atom labels. Select **Bonds** in the pull-down menu, you should have one entry in the window for the distance between the two atoms. Select the entry for the distance in the panel window by clicking on it. Look at the **Value** field, it should read 2.861. Select the Graph tab. If you don't see a small graph, click on **Show preview**. Compare with Figure 11.

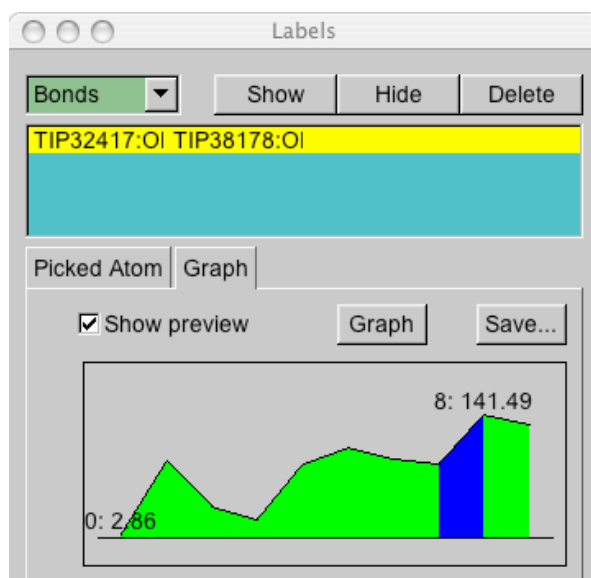


Figure 11

This is a plot of the distance between the two oxygen atoms as a function of time. The minimum (2.86 Å) and the maximum (141.49 Å) values are indicated. Depending on the operating system you are using and the software available, you could obtain a larger plot automatically by clicking on **Graph**. Otherwise, you can save a text file with the plot data by clicking on **Save**. To see the specific values of distance in the **Value** field of the **Picked Atom** tab, change the frame in the **VMD Main** panel and click again on the distance entry in the **Labels** panel window. You can also test this result by playing the animation and tracking the location of the two water molecules.

References.

- Dutzler, R., Campbell, Ernest B., MacKinnon, R., “Gating the Selectivity Filter in CIC Chloride Channels”, *Science*, 300, 108-112, 2003.
- Frishman, D., Argos, P., “Knowledge-based secondary structure assignment”, *Proteins: structure, function and genetics*, 23, 566-579, 1995.
- Humphrey, W., Dalke, A. and Schulten, K., “VMD - Visual Molecular Dynamics” *J. Molec. Graphics*, 14, 33-38, 1996.
- Varshney, F. P. Brooks, W. V. Wright, “Linearly Scalable Computation of Smooth Molecular Surfaces”, *IEEE Comp. Graphics and Applications*, 14, 19-25, 1994.