VIrus Particle ExploreR Database (VIPERdb)

Yaroslav Daniel Bodnar
University of Illinois at Urbana-Champaign
May 2009

VIPERdb (http://viperdb.scripps.edu/index.php) is a comprehensive database and on-line tool suite for structural virology. VIPERdb is integrated with handy visualization and bioinformatics tools, which makes it possible to quickly get a handle on these large, complex structures. This document provides an overview of the information available on VIPERdb and instructions on using several of the utilities. VIPERdb is an extensive, attractive, and well-designed site, so have fun surfing.

Searching the Database

A nice front-end for exploring VIPERdb is the Virosphere2003 utility (located in the Utilities drop-down on the top menu). This page brings up a nice graphic illustrating the biodiversity of viruses and principles of viral taxonomy. Representative viruses of many viral families can be reached by clicking on one of the yellow pentagons (which brings up database entries in the selected virus family) and then clicking on the PDB ID of the example (shown in blue). This takes the user to the front-page of the selected virus, from which data about the virus can be accessed.

VIPERdb also makes it easy to search the database by virus family and T-number. This can be done by clicking on the Family Index and T-Number Index options of the Data drop-down on the top menu. To access the data, just choose what you're looking for in the drop-down search box and click on the name of the virus that you want (first data column from the left).

If you know the name of the virus you want, you can search for it by clicking on Find a Virus on the top menu (on the right in orange). Do not hit enter to search; instead click on the name of the virus you want, which appears in the drop-down menu. Some things that you may try here are Poliovirus, Rhinovirus, Hepatitis, and Papilloma.

Navigating the Database

The data available for each virus can be navigated from the front-page for each virus by switching the tab located at the left (Biodata, Illustrations, 3D IAU, and so forth).

On the front-page (Biodata tab) for each virus you'll find some useful physical properties of the virus, including the T-number, net surface charge, and the average radius.

The 3D IAU tab brings up a simple molecular visualization tool (implemented with Jmol) that shows one asymmetric subunit of the virus and its location relative to neighboring subunits. To the right of the Jmol window, you'll find some options used to change the shown graphic representations. SHOW & HIDE allows you to view only selected parts of the structure (such as polypeptide chains, crystallographic waters, and so forth). Below the structure segment check-boxes there are also options that allow you to visualize the icosahedral lattice, change the background color to white (which may help visualize the structure), the x, y and z-axes, the molecular surface (SAS), and to spin the structure. COLOR BY allows you to color the structure either by secondary structure (yellow = beta; red = alpha) or protein chain. REPRESENTATION allows you to represent the molecules as ribbons (Cartoon), van der Waals spheres (Spacefill), or a back-bone trace. NEIGHBORS allows you to also visualize the neighboring asymmetric subunits around the 5 and 3-fold symmetry axes. VIEW has some useful options, including resetting the view (green button), popping the Jmol out as a separate window (white/blue button with arrow; this is very useful because you can make the graphic window full-screen), and launch STRAP with the structure of the full virus loaded (this will be discussed when describing STRAP). One Jmol hint is that you can access other features (such as Measurement, which allows the user to measure distances and angles) by right-clicking on the visualized structure.

The Annotations tab gives access to some useful computed data for the virus structure. Clicking on Accessible Surface Profiles brings up a page that shows the Solvent Accessible Surface Area (SASA) for each residue (note that there is a graph for each polypeptide chain). Clicking on one of the blue triangles on a graph will zoom-in around this point. Clicking on the Contact Tables brings up a large data table that shows which residues of an asymmetric subunit are involved in making contacts. The first column (Residue1-Residue2) shows the residues involved in the interaction and the second column (Type1-Type2) shows the type of interaction (scroll all the way to the bottom of the page to see a summary and key). All of the other columns tell which polypeptide chains are involved in the interaction.

Using STRAP: An On-Line Bioinformatics Tool

STRAP is a program that performs sequence and 3D structure alignments, as well as visualization. STRAP is easy to use, has nice tools, and is very well integrated with VIPERdb. STRAP also has nice tutorials that can be accessed by hitting F10 (which enables full STRAP functionality) and clicking on the Tutorials option in the top main menu.

To launch STRAP, click on a STRAP button (Strap written in red and blue) somewhere in VIPERdb. There are STRAP buttons located on the front-page of each virus data entry at the bottom of the blue box on the right and on the 3D IAU tab under the VIEW options. After clicking on the button, the web browser will begin to load a Java Web Start application. Run the application and accept all default settings.

When STRAP comes up, by default the amino acid sequence for the selected virus is loaded and a back-bone trace of the full virus structure is shown in the graphics window.

At the bottom of the STRAP graphical interface is located the sequence information. The name of the loaded structure file is shown, to the right of which is the sequence. To select a sequence (which makes it a target of any selected options, such as sequence coloring, alignments and analyses), simply click on the protein name (selected residues are colored pink). Below this is a scroll bar that allows you to move around in the sequence. In the lower left-hand corner is a drop down menu that allows you to choose the property by which to color the sequence. Chemical, for example, colors the sequences by amino acid residue. The scroll bar at the bottom center allows only residues that are highly conserved to be colored. The buttons on the bottom right, allow you to change options (in the options menu it may be useful to turn on the “color shade horizontal scroll-bar” and turn off the “colorize only selected sequences” options), show the 3D structure of a selection, and perform quick alignments.

On the sequence display, it is possible to select particular residues of a protein by dragging across them with the mouse. The selected residues will be highlighted in blue on the sequence display, and will also be highlighted on the structure shown in the visualization panel.

The visualization panel (all utilities that perform operations on the sequences are also displayed here) located above the sequence information contains some important options. To the left of the graphic display in the gray box, there will be a buttons that show a utility that you are currently running). On the graphic display, there are some important buttons that become visible when you hover your cursor over the panel. The drop-down

menu at the top left allows you to visualize only one asymmetric subunit or the full virus (although less impressive, it is usually more useful to look at one asymmetric subunit). In the top right corner, the leftmost button (which looks like a graph) allows you to pop the visualization panel into a new window, which can be maximized for easier viewing), the camera allows you to take a snapshot, and the X closes the display (which is useful for cleaning up the visualization panel).

To perform sequence/structure alignments and sequence analysis, you will need to open the full functionality of STRAP. This is done by pressing the F10 key.

A sequence alignment is performed by selecting Align in the Align option on the top menu bar. This will bring up a new panel that asks you to select the alignment algorithm on the left side (choose MultipleAlignerClustalW for a multiple sequence alignment) and which sequences to align on the right side (select a sequence by clicking on it; multiple sequences can be selected by holding shift). Start the alignment by clicking the GO button on the left panel. Once the calculation is done, the resulting alignment is shown in the visualization panel. Click the approve alignment button on the bottom of the visualization panel to apply the alignment to the loaded sequences.

A 3D structural alignment is performed by choosing the Superimpose Structures in the Align option on the top menu bar. This brings up a new panel. On the left, you must specify the reference protein to which the other structures will be aligned. On the right, select the structures that you want to align to the reference sequence. Use the default alignment algorithm (can be changed via the drop-down). Start the alignment by clicking GO. This will bring up a series of tabs. Each tab shows the alignment of a selected structure to the reference.

Several analyses can also be performed in STRAP, the most useful one of which is the phylogenic tree generator. The Phylogenic Tree generator is accessed in the Analyze option on the top menu bar. This brings up a panel that allows you to select the phylogenic tree generator algorithm. Select the first option, which provides much better options, and hit GO. This will bring up a new window with the tree. The root of the tree is in the lower left-hand corner. Several options, located on the left can also be changed (a useful option is the branch length values, which gives the distance that each sequence is from the root of the tree).