Return to Main Documents Page

How to Use Visual Molecular Dynamics (VMD)

This is a very basic introduction to VMD.

So, what is VMD? In the words of its developers:

> VMD (Visual Molecular Dynamics) is a molecular visualization and analysis program designed for biological systems such as proteins, nucleic acids, lipid bilayer assemblies, etc. It is developed by the Theoretical and Computational Biophysics Group at the University of Illinois at Urbana-Champaign. Among molecular graphics programs, VMD is unique in its ability to efficiently operate on multi-gigabyte molecular dynamics trajectories, its interoperability with a large number of molecular dynamics simulation packages, and its integration of structure and sequence information.

The aim of this tutorial is to very quickly get you familiar enough with VMD to be able to view individual protein structures and the sorts of trajectories containing many structures that are produced by molecular dynamics and other simulation techniques. This document is deliberately designed to cover only the most basic features of VMD. Excellent tutorials teaching the full range of the functionality provided by the program can be found at the VMD website, here.

Prerequisites

This tutorial assumes three things:

1. Access to the following two files (download and save them in a location of your choice):
   - a PDB single structure file, dbd.pdb
   - a DCD multiple structure trajectory file, dbd-short-traj.dcd
2. VMD is installed on your computer.
3. You have started VMD.

If you don't currently meet the last two criteria then follow the instructions in the next two sections.

Installing VMD

If VMD is not already available on your computer then you will, obviously, need to install it. Fortunately, VMD is free and pre-compiled ready to go versions are available for Windows, OS X and Linux. You just need to register with your email address and then download the appropriate package from here.

Full installation instructions are available here, but the Windows and OS X installations are pretty self explanatory. The small extra effort involved in Linux installation is also documented in the README file distributed with the program.

Running VMD

To start the program:

- OS X: Double click on the VMD icon in the Applications directory.
- Linux: Type vmd in a terminal window.
- Windows: Select *Start -> Programs -> VMD*.

Upon opening VMD opens three windows (Figure 1): the Main, OpenGL Display and Console (or Terminal on OS X) windows. To end a VMD session, go to the Main window, and choose *File -> Quit*. You can also quit VMD by closing the Console or Main window.

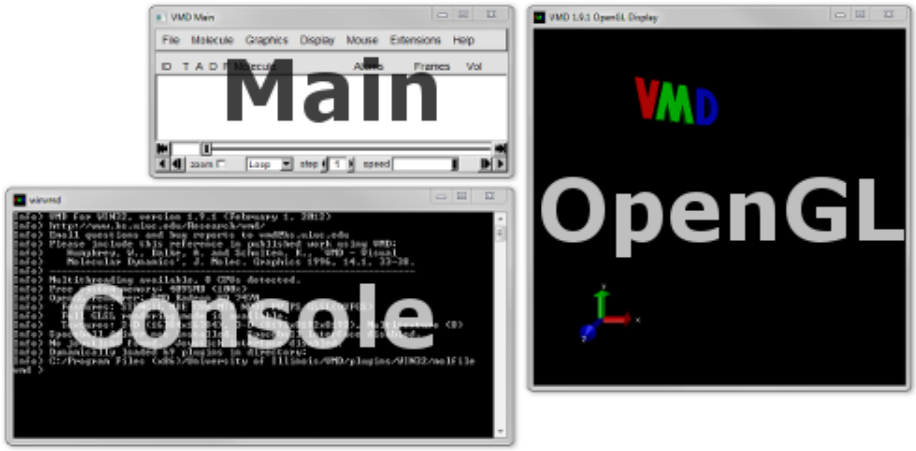


Figure 1: Figure 1. The three initial windows opened by VMD.

Viewing Single Molecules

In this section we will load a single structure from a PDB and learn how to view it from different angles and to alter the way it is rendered on screen.

Loading a Molecule

1. From the menubar in the Main window select *File -> New Molecule...* (Figure 2).

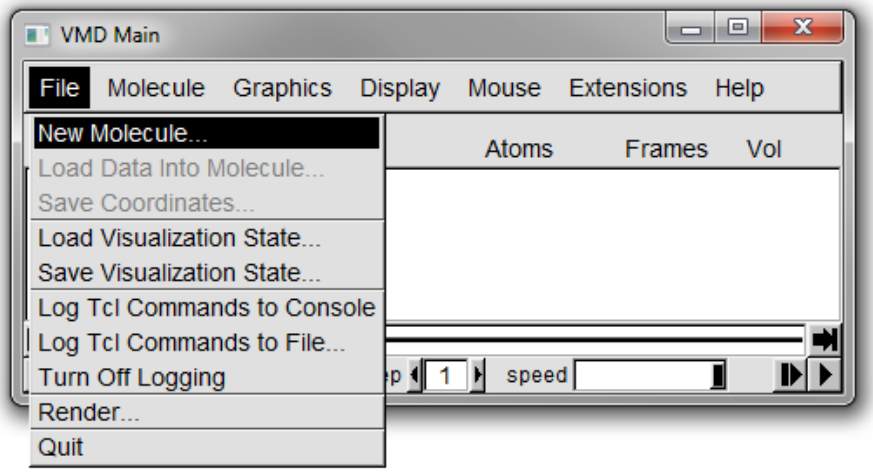


Figure 2: Figure 2. Selecting the *New Molecule...* menu option.

2. A Molecule File Browser window similar to that shown in Figure 3 should now open. Select the *Browse..* button (circled in red in Figure 3). This will open a file selection dialog. Navigate to the PDB file, dbd.pdb, which you downloaded earlier, select it and click *Open*.

3. You should now have been returned to the Molecule File Browser window (the structure will not yet have been loaded). To load the file you need to click the *Load* button (circled in blue in Figure 3). The structure should now be loaded and the OpenGL Display window look something like Figure 4.

Notice that once the molecule is loaded basic information including the name of the file and number of atoms appear in the Main window.

Once the structure is loaded you can close the Molecule File Browser at any time.

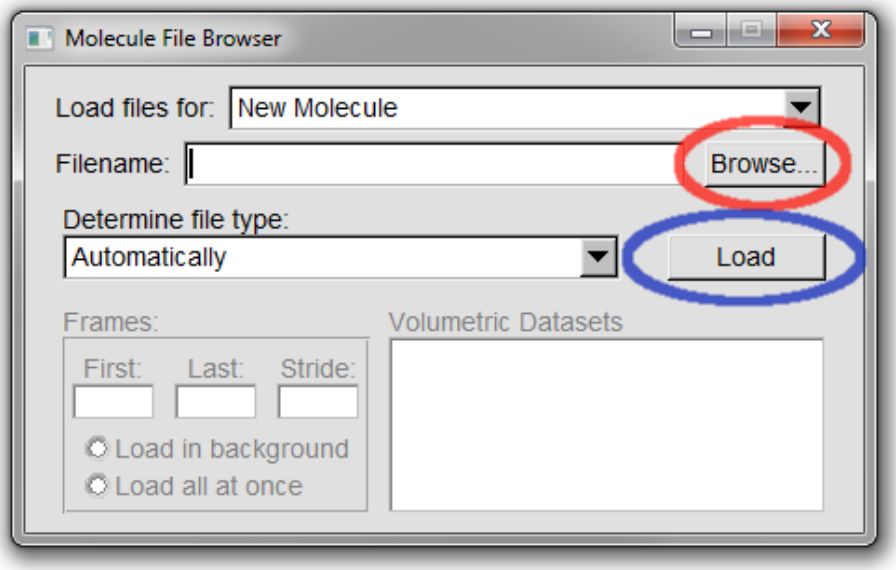


Figure 3: Figure 3. The Molecule File Browser Window used to load new molecules.

Altering Your Viewpoint

From the menubar in the Main window select *Mouse* (Figure 5).

A larger array of options for how the mouse interacts with the molecule shown in the OpenGL Display. We will concentrate on the first four options: rotate mode, translate mode, scale mode and center. The keyboard shortcut for each option is shown next to the choice in the menu (for example pressing 't' when the OpenGL windows is selected will change into translate mode). By default VMD uses the rotate mode.

Generally, to change the viewpoint you need select the OpenGL Display windows, hold the left mouse button and move the mouse (the right mouse button can be used in rotate mode, see below). Have a play with all of the modes below until you feel comfortable changing the view and switching between the various modes. You can easily tell which mode you are in because each one has a distinctive cursor (Figure 6).

Rotate mode: With the left mouse button held moving the mouse horizontally rotates the molecule around the vertical axis, up and down around a horizontal one. When the right mouse button is held then the rotation is performed around an axis running into and out of (i.e. perpendicular to) the screen.

Translate mode: When you hold down the left mouse button you can now move the molecule up, down, left or right in the viewing plane.

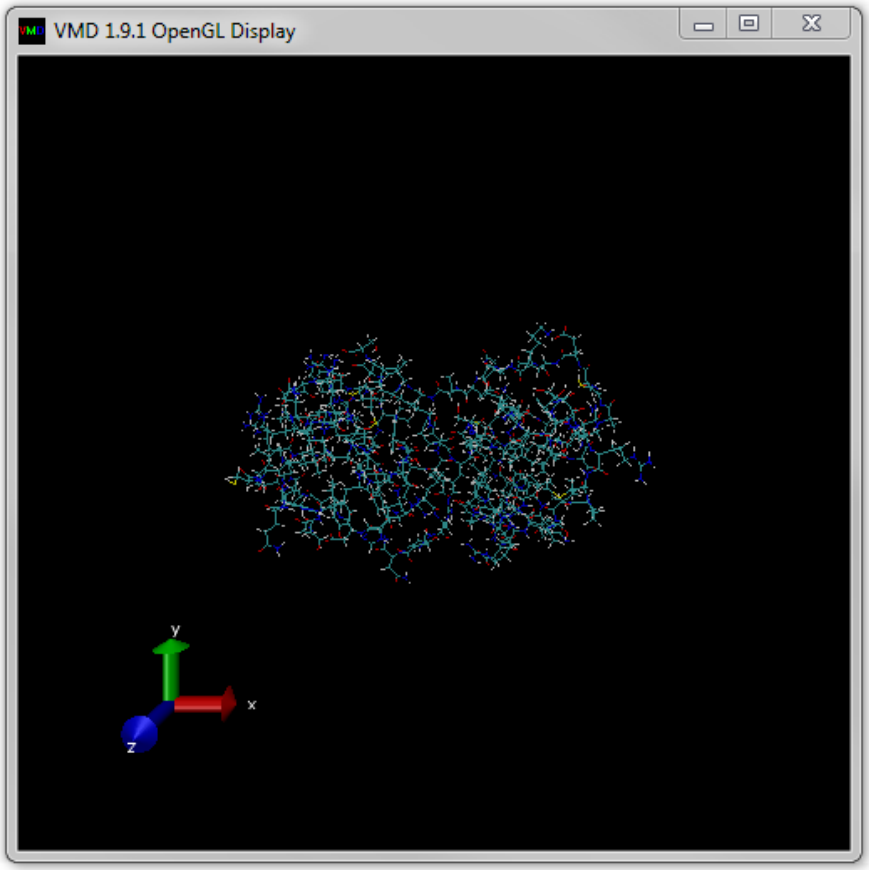


Figure 4: Figure 4. The dbd.pdb structure as displayed in VMD when first loaded.

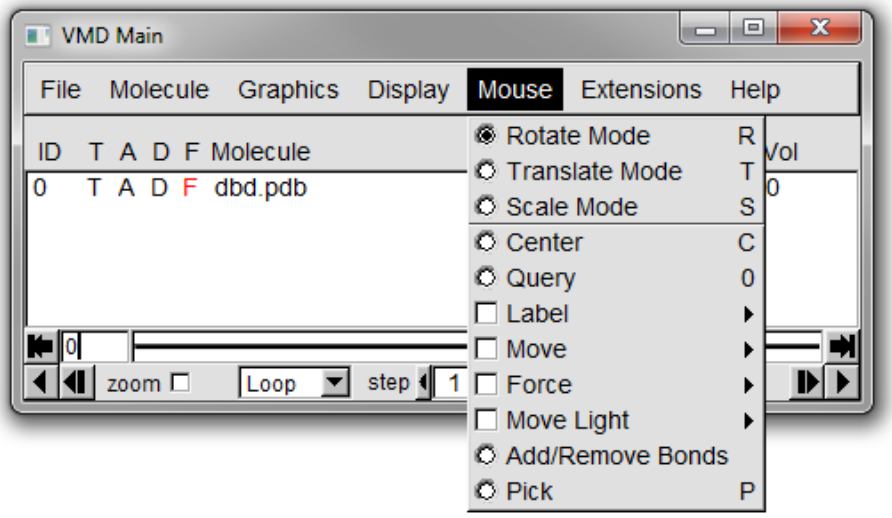


Figure 5: Figure 5. The options available for changing the viewpoint with the mouse.

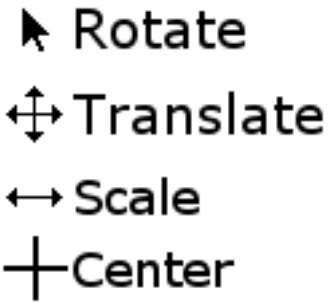


Figure 6: Figure 6. Different viewpoint changing modes are indicated by different cursor styles.

Scale mode: Moving the mouse left or right whilst holding the left mouse button zooms in and out.

Center: The centre option is not a direct method for altering the view point. It is used to pick a point about which to perform a rotation. Once you have selected a center point change to Rotate mode to perform the rotation.

At any point you can reset the view by pressing '=' whilst the OpenGL Display is selected (you can also use the menu option *Display -> Reset View*.

Changing the Appearance of the Molecule

1. Select *Graphics -> Representations...* from the menubar of the Main window. A new Graphical Representations window will open looking like Figure 7. Section Figure 7 (a) shows the current representation being highlighted. It shows the style of the representation (in this case Lines), the colouring method (Name) and the selection of atoms to which this representation is applied. The selection is determined by the contents of the input box shown in Figure 7 (b). The current selection is for all atoms. The language used to select atoms in VMD is documented in the section of the user guide here.

2. Try changing the Coloring Method by choosing different options from the drop down menu shown as Figure 7 (c).
3. Alter the Drawing Method of the graphics used to display the molecule with the drop down menu shown as Figure 7 (d). A good choice is 'New Cartoon' which highlights different elements of protein secondary structure (shown in Figure 8). Notice how the region indicated by Figure 7 (e) changes to give you different customization options for the different Drawing Methods.

Working With Trajectories

When dealing with simulations multiple structures are commonly saved together in files called 'trajectories'. An example of a trajectory format is DCD (used by NAMD and CHARMM). Such files contain only coordinates and no information about which atoms are represented. In order to vizualize the atomic structure represented by a trajectory it has to be combined with a file containing structural information. PDB files contain both structural and coordinate information and can perform this role, other filetypes such as PSF files contain only structural information and no coordinates.

The trajectory from this point assumes that you have already loaded the dbd.pdb PDB files into VMD. If that is not the case go through the instructions for Loading a Molecule.

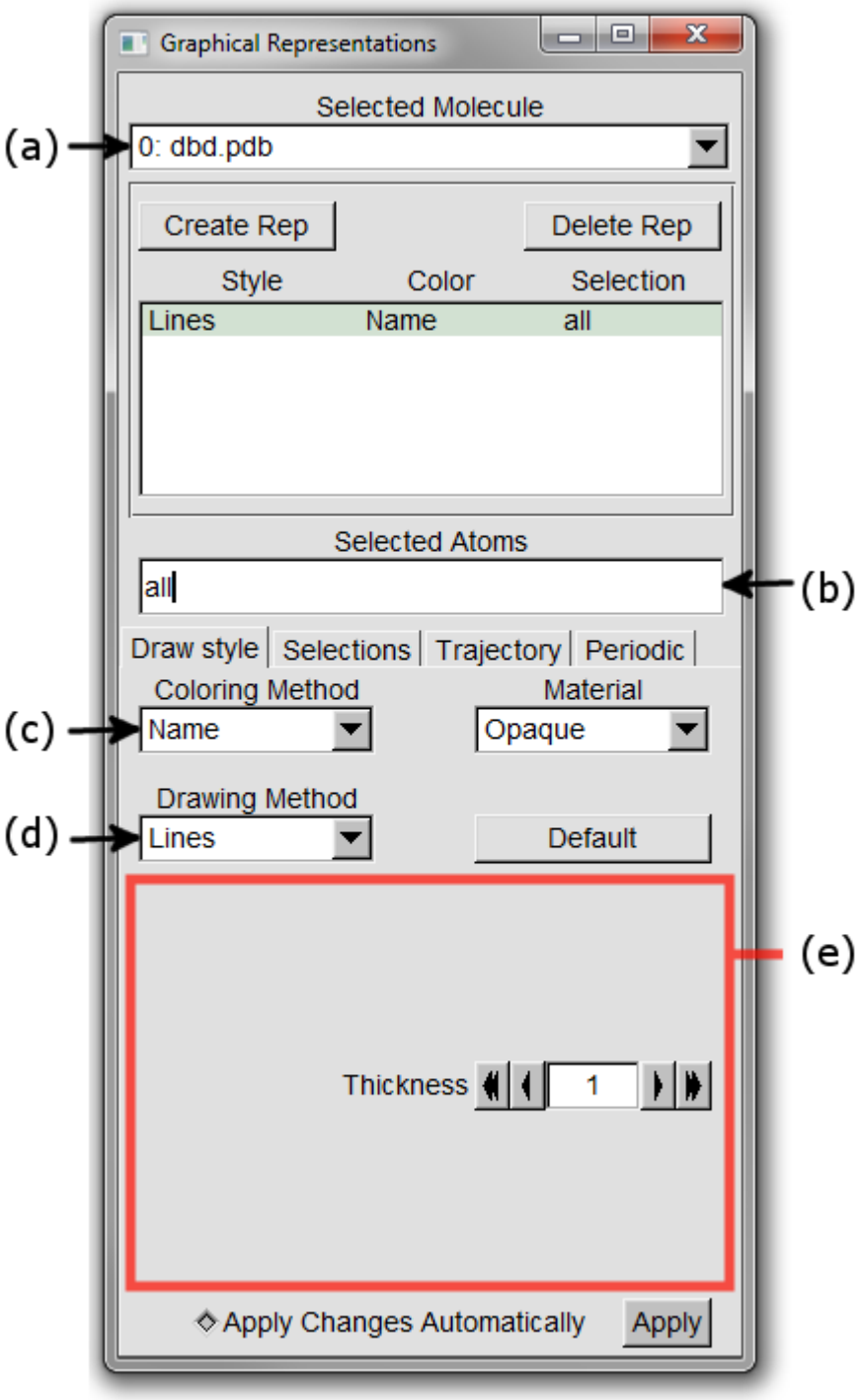


Figure 7: Figure 7. Graphical Representations window.

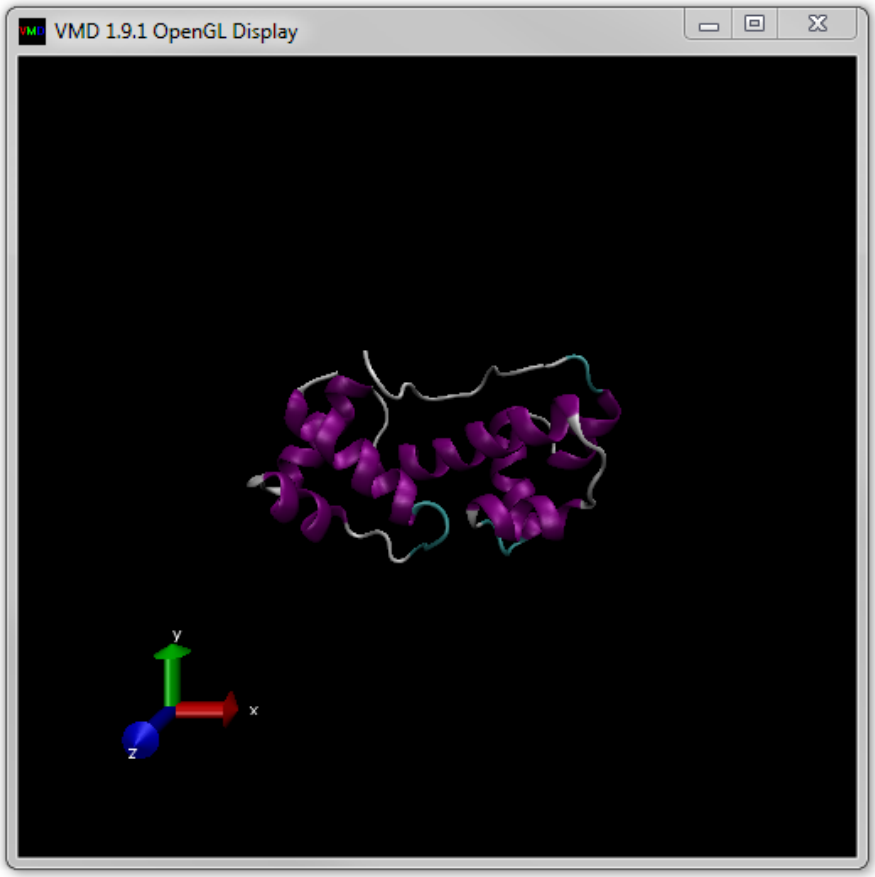


Figure 8: Figure 8. The dbd.pdb structure as displayed in the New Cartoon style, and colored using Secondary Structure.

Loading a Trajectory

In the Main window either select *File -> Load Data into Molecule...* or right click on the line for dbd.pdb and select *Load Data into Molecule...* from the menu (see Figure 9).

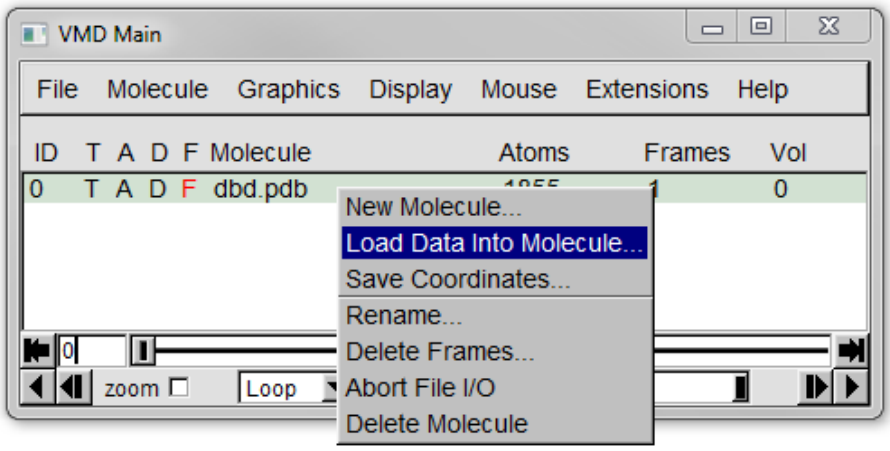


Figure 9: Figure 9. Loading data into an already loaded structure.

Like when we loaded a PDB previously a Molecule File Browser window should appear. *Browse* to the dbd-short-traj.dcd and *Load* it.

As the trajectory loads you will see the number of frames increase in the Main windows and the structure in the OpenGL Display update.

Trajectory Animation

The Main window should now look something like it does in Figure 10.

However, the current frame (circled in red in Figure 10) should show a number one less than the number of frames shown above it in the information panel. This is because frames are counted from zero in VMD. Frame zero in this case is the structure loaded from the PDB. If you load a structure information only file, like a PSF, instead of a PDB then frame zero will be the first frame of the DCD.

You can use the slider highlighted in blue in Figure 10 to move to any frame you desire. It is also possible to type in a specific frame number in the current frame region.

Beneath these controls are the animation controls (highlighted in green in Figure 10). These allow you to play through the frames of the trajectory. Forwards and backwards play buttons can be found at either end of the controls. Between

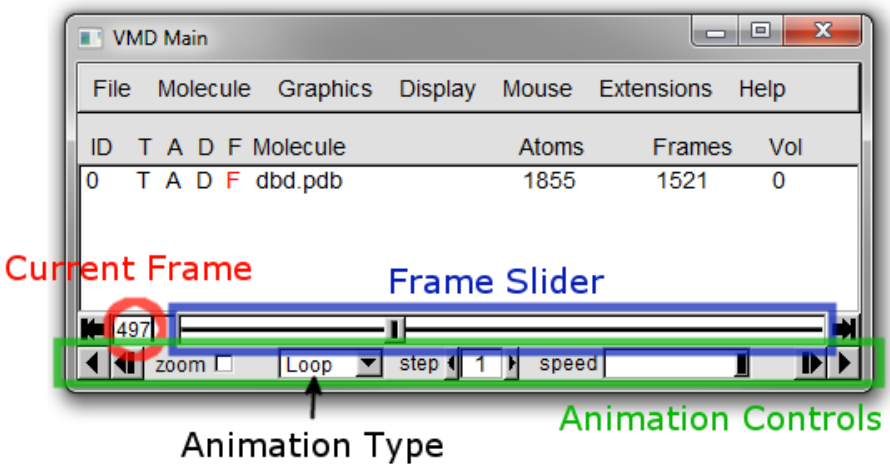


Figure 10: Figure 10. Main window with a trajectory loaded.

these are controls which alter the way the animation plays. The 'step' controls the gap between the frames which are displayed and 'speed' is self explanitory. The animation type drop down menu controls what happens when the animation has played the last frame in the current direction:

Once: The animation terminates.

Loop: The animation plays again from the start (opposite end).

Rock: The animation reverses direction.

Play with these controls until you are happy you understand how they work. Once you are at that stage our work here is done.

Want to Know More About VMD?

VMD is not only a molecular viewer but also allows you to analyse structures and trajectories, convert file format, render graphics for publication and many more things well beyond the scope of this tutorial. When you want to try out these features best place to start is the User Guide.

When you, inevtiably, run into difficulties a good first port of call is the mailing list.

Return to Main Documents Page Go to top

Supported via CCP-SAS a joint EPSRC (EP/K039121/1) and NSF (CHE-1265821) grant

Return to Main Documents Page

PDB Scan

Generates a report that characterizes the user supplied PDB file (alpha).

Accessibility

The Monomer Monte Carlo module is accessible from the Build section of the main menu.

Basic Usage

The purpose of this module is to assess whether an input PDB is ready for simulation and where possible to provide files enabling CHARMM forecfield parameterization. When a header is available, two scans of an input PDB are made, one accounting for information extracted from header records and one based solely on information contained within the coordinate records. Both scans produce independent reports and in cases where a header is available this is deemed the most reliable guide to the files contents.

Information on missing atoms and residues and those not covered as standard by the CHARMM 27 forcefield are reported. Furthermore, when header information is available the BIOMT records are extracted to indicate if the protein unit in the coordinates is represenative of the biological unit or whether a coordinate transform is needed to complete it (e.g. in cases where symmetry can be used to create a dimer or higher order oligomer). Both header and no header reports provide information on (predicted) disulphide bonds found within the structure.

Notes

- The text area report provides summary information on missing atoms and residues and those not covered as standard by the CHARMM 27 forcefield
- All Header scan information is extracted from the PDB header records
- Header scan checks for consistency between the SEQRES and coordinate sequences

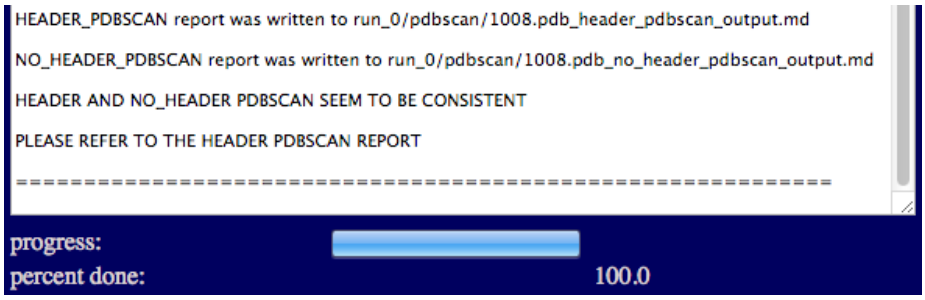
- No Header scan reports missing atoms by comparing coordinates to the CHARMM 27 residue descriptions
- No Header scan detects possible missing residues by calculating the backbone C-N distance between two adjacent residue listed in the PDB coordinate entry.
- No Header scan detects possible disulfide bonds by calculating the pair-wise distance between all Cys-S atoms.

Screen Shots and Description of Input Fields

These examples show the analysis and file output of a run scanning a PDB with and without header records. The former system requires editing before it can be prepared for simulation, the latter is ready to simulate.

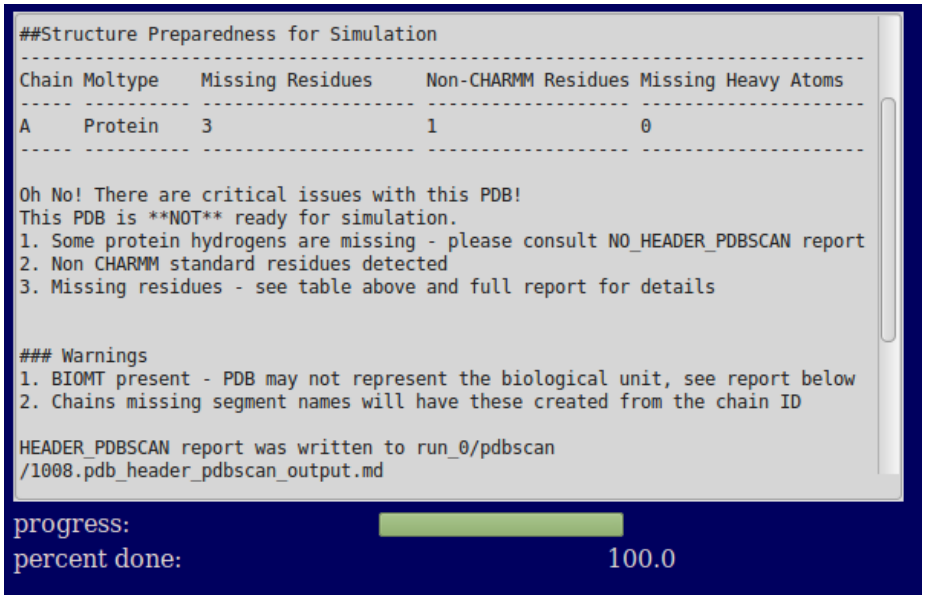
- **run name** user defined name of folder that will contain the results.
- **pdb input file** PDB file to be scanned for simulation suitability.

Indication of report generation A brief summary report is shown in the textbox indicating whether the PDB is ready for simulation and that a full report has been successfully generated.



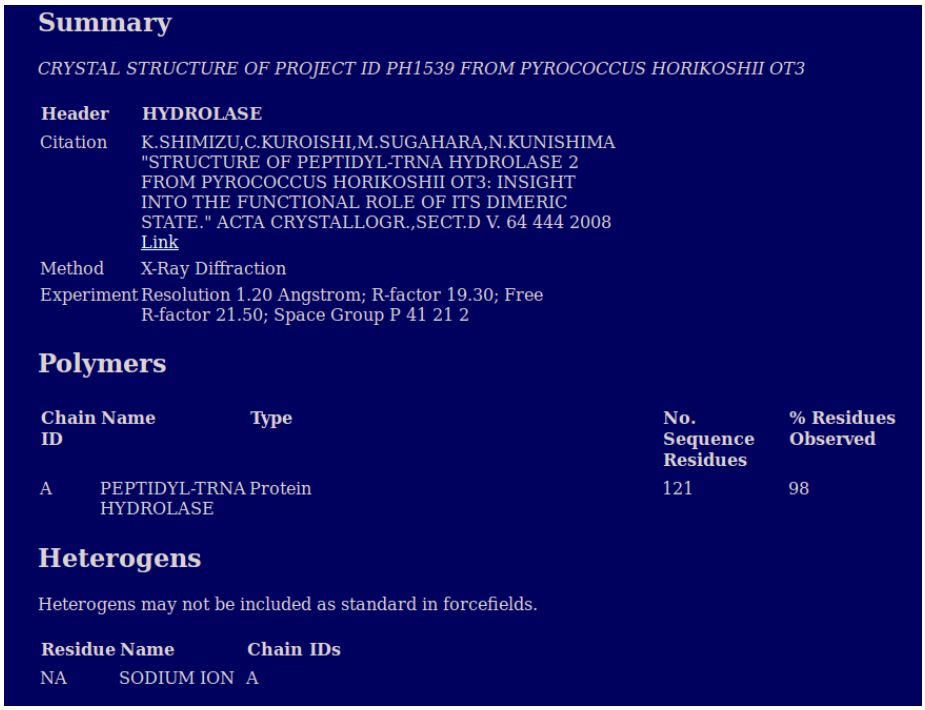
Case 1: PDB with Header

This example shows the report for a PDB containing a single chain in the coordinates but which the header indicates has a dimeric biological unit.



The text output region provides a brief summary of the contents of the file. It focusses on issues that might prevent the creation of a correct CHARMM parameterized system (missing residues and atoms and unusual residues not standard in CHARMM). If the PDB is not ready for simulation (as is the case in this example) many of the lines point to the full report for more details.

HEADER_PDBSCAN Report A full report is provided which details all of the information extracted from the PDB header and coordinates. The first part of the report reproduces the text area report for future reference.



Summary

CRYSTAL STRUCTURE OF PROJECT ID PH1539 FROM PYROCOCCUS HORIKOSHII OT3

Header	**HYDROLASE**
Citation	K.SHIMIZU,C.KUROISHI,M.SUGAHARA,N.KUNISHIMA "STRUCTURE OF PEPTIDYL-TRNA HYDROLASE 2 FROM PYROCOCCUS HORIKOSHII OT3: INSIGHT INTO THE FUNCTIONAL ROLE OF ITS DIMERIC STATE." ACTA CRYSTALLOGR.,SECT.D V. 64 444 2008 Link
Method	X-Ray Diffraction
Experiment	Resolution 1.20 Angstrom; R-factor 19.30; Free R-factor 21.50; Space Group P 41 21 2

Polymers

Chain ID	Name	Type	No. Sequence Residues	% Residues Observed
A	PEPTIDYL-TRNA HYDROLASE	Protein	121	98

Heterogens

Heterogens may not be included as standard in forcefields.

Residue	Name	Chain IDs
NA	SODIUM ION	A

The next section of the gives information on the totality of the contents of the PDB.

- The summary section of the main report gives the high level information on the protein provided in the PDB header relating to what it represents and the origin of the data.
- The polymers section gives information on the residues in each polymer chain in the file.
- The heterogens section describes all of the unique non-standard residues reported in the header

System statistics

Metrics derived from PDB coordinates:

Measurement	Value
Number of models	1
Mass	15324.48
Centre of mass	17.85 16.43 33.52
Radius of gyration	14.96
Coordinate minimums	0.88 -6.20 14.78
Coordinate maximums	38.28 40.67 49.46

The system statistics section provides basic information on the geometry and mass of the protein calculated from the atomic coordinates. A similar table is also provided per chain later in the report.

Biological Unit

Biomolecule: 1 - Author asserted biological unit for chain A: Dimeric

#	Rotation	Translation
1	1.000000	0.00000
	0.000000	
	0.000000	
1	0.000000	0.00000
	1.000000	
	0.000000	
1	0.000000	0.00000
	0.000000	
	1.000000	
2	0.000000	0.00000
	1.000000	
	0.000000	
2	1.000000	0.00000
	0.000000	
	0.000000	
2	0.000000	85.79100
	0.000000	
	-1.000000	

The biological unit section describes the oligomeric state suggested by the authors of the PDB (and/or the software they used) in the BIOMT records.. It also shows the geometric transformation matrix provided in the PDB header to transform the provided coordinates where multiple copies are needed to produce the biological unit.

Chain A

Sequence:

```
      123456789|123456789|123456789|123456789|123456789|
1     mikMFKYKQVIVARADLKLSKGKLAAQVAHGAVTAAFEAYKKKREWFEAW
51    FREGQKKVVVKVESEEELFKLKAEAEKLGLPNALIRDAGLTEIPPGTVTV
101   LAVGPAPEEIVDKVTGNLKLL----------------------------
151   ----------------------------------------------
201   ----------------------------------------------
251   ----------------------------------------------
301   ----------------------------------------------
351   ----------------------------------------------
401   ----------------------------------------------
451   ----------------------------------------------
501   ----------------------------------------------
551   ----------------------------------------------
601   ----------------------------------------------
651   ----------------------------------------------
701   ----------------------------------------------
751   ----------------------------------------------
801   ----------------------------------------------
851   ----------------------------------------------
901   ----------------------------------------------
951   ----------------------------------------------
1001  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
1051  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
1101  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
1151  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

Gaps:

1-3

Heterogens:

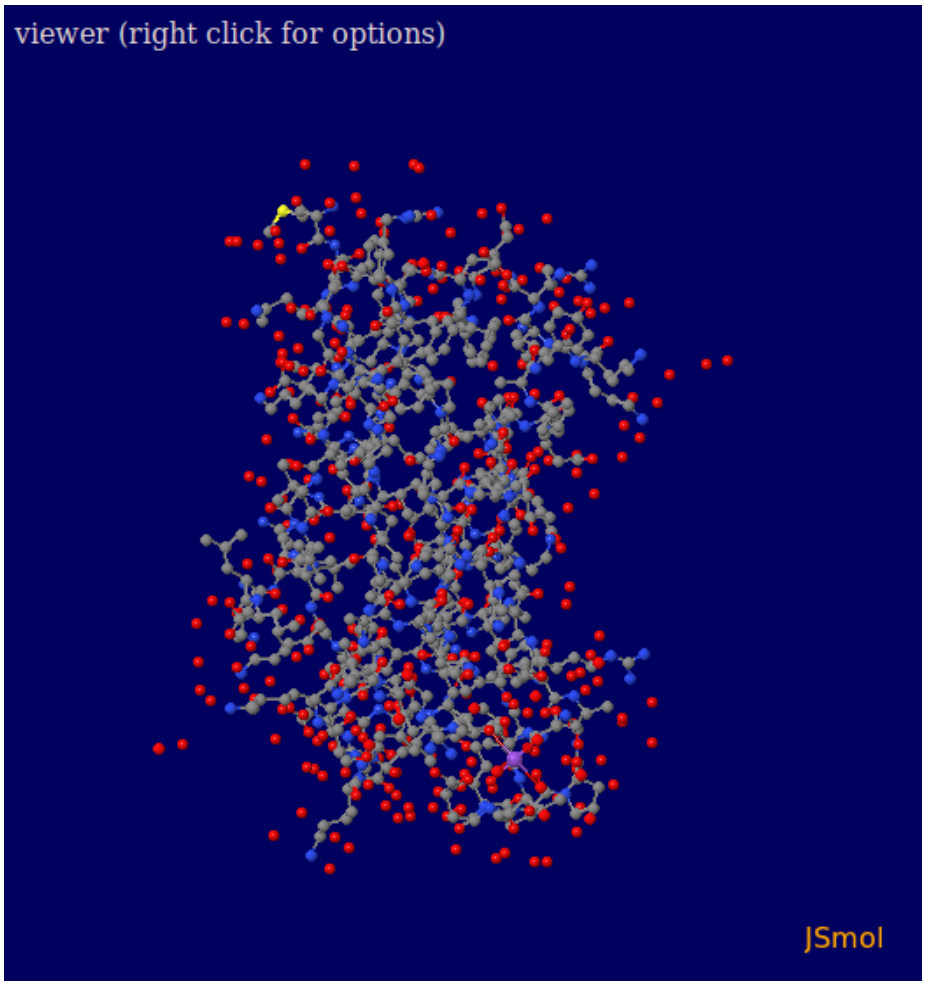
Residue No.	Resdiue ID	Description
1001	NA	SODIUM ION
1002-1188	HOH	Water

A report is provided for each chain which reports on the sequence in FASTA format and any heterogens present. Lower case letters in the sequence indicate missing residues

---

PSFGEN preparation files In addition to producing a report detailing the contents of the files two PDB files are generated for each chain. The first contains full entries including the coordinates of all the atoms present in the input PDB. the second a single CA atoms for each residue in the full sequence described in the header (whether or not coordinates were found in the input PDB). These files are used in PSFGEN to create completed PDB/PSF pairs for simulation.

Visualization As part of the outpt a JSmol vizualization of the protein is produced. Holding down the left mouse button and moving the cursor over the picture allows you to rotate the view, the scroll wheel facilitates zooming in and out. Right clicking on the image allows you to access all of the JSmol options.



Files Used and Created in Example

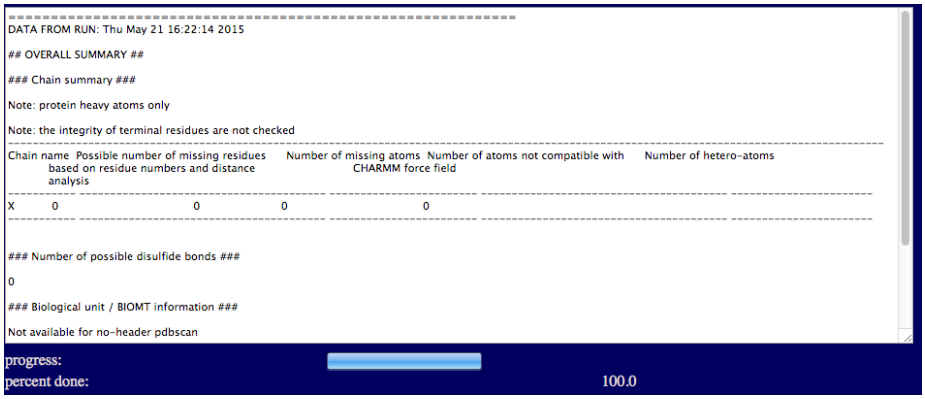
- input files
  1008.pdb
- output files
  1008.pdb_header_pdbscan_output.html
  1008.pdb_header_pdbscan_output.md
  1008.pdb_no_header_pdbscan_output.html
  1008.pdb_no_header_pdbscan_output.md

sequence_chain_A_1008.pdb

coor_chain_A_1008.pdb

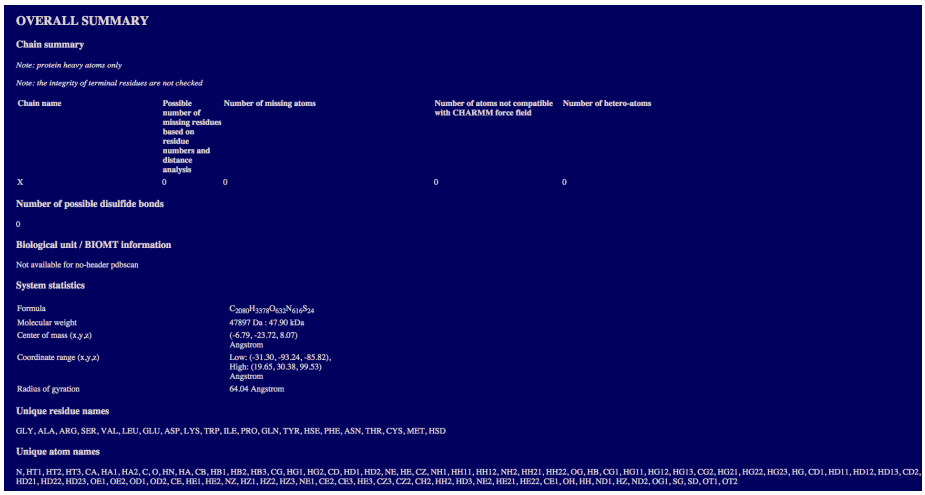
Case 2: PDB with no header

This example shows the report for a PDB that contains no header.



The text output region provides a brief summary of no header PDBScan report, that include the number of missing residues, missing atoms, and hetero-atoms for each individual chain, and the number of disulfide bonds. The number of missing atoms and the number of disulfide bonds are given based on distance analysis as described in the previous section.

NO_HEADER_PDBSCAN Report A full report is provided which details all of the information extracted from the PDB coordinate entry.



OVERALL SUMMARY

Chain summary

Note: protein heavy atoms only

Note: the integrity of terminal residues are not checked

Chain name	Possible number of missing residues based on residue numbers and distance analysis	Number of missing atoms	Number of atoms not compatible with CHARMM force field	Number of hetero-atoms
X	0	0	0	0

Number of possible disulfide bonds

0

Biological unit / BIOMT information

Not available for no-header pdbscan

System statistics

Formula	$C_{2080}H_{3379}O_{632}N_{616}S_{24}$
Molecular weight	47897 Da : 47.90 kDa
Center of mass (*x,y,z*)	(-6.79, -23.72, 8.07) Angstrom
Coordinate range (*x,y,z*)	Low: (-31.30, -93.24, -85.82), High: (19.65, 30.38, 99.53) Angstrom
Radius of gyration	64.04 Angstrom

Unique residue names

GLY, ALA, ARG, SER, VAL, LEU, GLU, ASP, LYS, TRP, ILE, PRO, GLN, TYR, HSE, PHE, ASN, THR, CYS, MET, HSD

Unique atom names

N, HT1, HT2, HT3, CA, HA1, HA2, C, O, HN, HA, CB, HB1, HB2, HB3, CG, HG1, HG2, CD, HD1, HD2, NE, HE, CZ, NH1, HH11, HH12, NH2, HH21, HH22, OG, HB, CG1, HG11, HG12, HG13, CG2, HG21, HG22, HG23, HG, CD1, HD11, HD12, HD13, CD2, HD21, HD22, HD23, OE1, OE2, OD1, OD2, CE, HE1, HE2, NZ, HZ1, HZ2, HZ3, NE1, CE2, CE3, HE3, CZ3, CZ2, CH2, HH2, HD3, NE2, HE21, HE22, CE1, OH, HH, ND1, HZ, ND2, OG1, SG, SD, OT1, OT2

The overall summary section provide:

- the number of missing residues, missing atoms, and hetero-atoms for each individual chain;
- the number of disulfide bonds;
- the system statistics including the formula, molecular weight, and geometric statistics;
- and the unique residue names and atom names present in the coordinate entry.

CHAIN "X" REPORT

Sequence

Note: missing internal residues represented by ''*

Note: protein heavy atoms only

```
  1 GARASVLSGGELDKWEKIRLRPGGKKQYKLKHIVWASRELERFAVNPGLL
 51 ETSEGCRQILGQLQPSLQTGSEELRSLYNTIAVLYCVHQRIDVKDTKEAL
101 DKIEEEQNKSKKKAQQAAADTGNNSQVSQNYPIVQNLQGQMVHQAISPRT
151 LNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQML
201 KETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMT
251 NNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFY
301 KTLRAEQASQEVKNAATETLLVQNANPDCKTILKALGPAATLEEMMTACQ
351 GVGGPGHKARVIAEAMSQVTNSATIMMQKGNFRNQRKTVKCFNCGKEGHI
401 AKNCRAPRKKGCWKCGKEGHQMKDCTERQAN
```

Possible missing internal residues based on distance analysis

Note: the two residues inside each bracket represent the residues before and after the missing residues

Note: protein heavy atoms only

None

Atoms missing from CHARMM force field

Note: each atom represented as ResidueName-ResidueNumber:AtomName

Note: protein heavy atoms only

None

Atoms not compatible with CHARMM force field

Note: each atom is represented as ResidueName-ResidueNumber:AtomName

Note: protein heavy atoms only

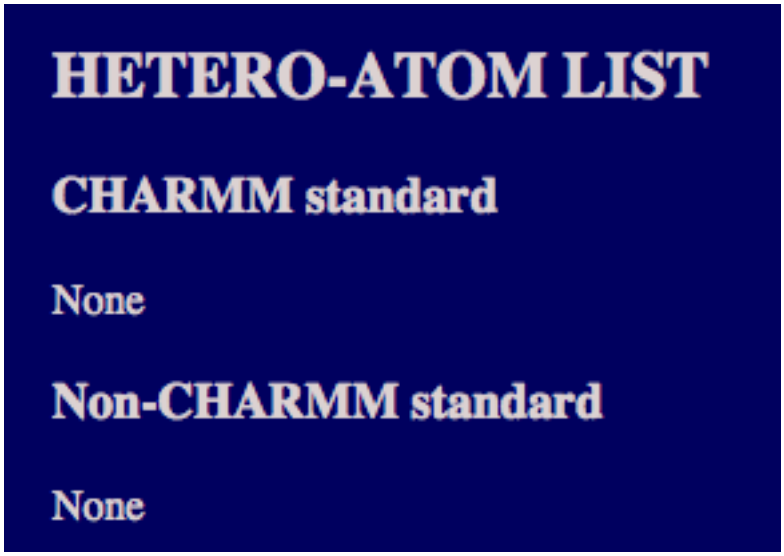
None

Chain statistics

Formula	$C_{2080}H_{3378}O_{632}N_{616}S_{24}$
Molecular weight	47897 Da : 47.90 kDa
Center of mass (x,y,z)	(-6.79, -23.72, 8.07) Angstrom
Coordinate range (x,y,z)	Low: (-31.30, -93.24, -85.82), High: (19.65, 30.38, 99.53) Angstrom
Radius of gyration	64.04 Angstrom
Distance between N- & C-terminals	149.74 Angstrom

The chain report section may contain multiple sections, each of which provides for the individul chain:

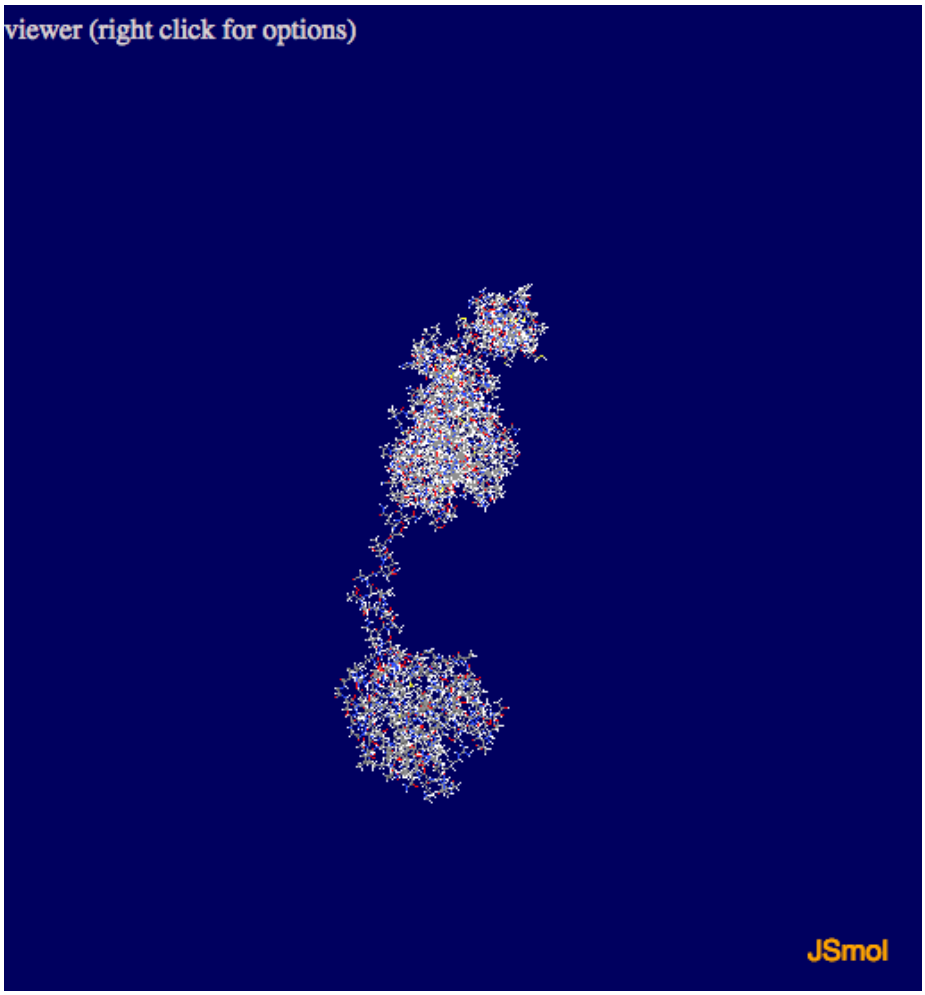
- the sequence;
- the number of missing regions based on distance analysis;
- the number of missing atoms from the CHARMM force field;
- the number of atoms that are not compatible with CHARMM force field;
- and the chain statistics including formula, molecular weight, and geometric statistics.



The hetero-atom section lists the hetero-atoms that are compatible and not compatible with CHARMM force field.

The disulfide bond section lists the disulfide bonds that are found by distance analysis.

Visualization As part of the outpt a JSmol vizualization of the protein is produced. Holding down the left mouse button and moving the cursor over the picture allows you to rotate the view, the scroll wheel facilitates zooming in and out. Right clicking on the image allows you to access all of the JSmol options.



Files Used and Created in Example

- input files
  hiv1_gag.pdb
- output files
  hiv1_gag.pdb_no_header_pdbscan_output.html
  hiv1_gag.pdb_no_header_pdbscan_output.md

Reference(s) and Citations

Return to Build

Return to Main Documents Page Go to top

Supported via CCP-SAS a joint EPSRC (EP/K039121/1) and NSF (CHE-1265821) grant