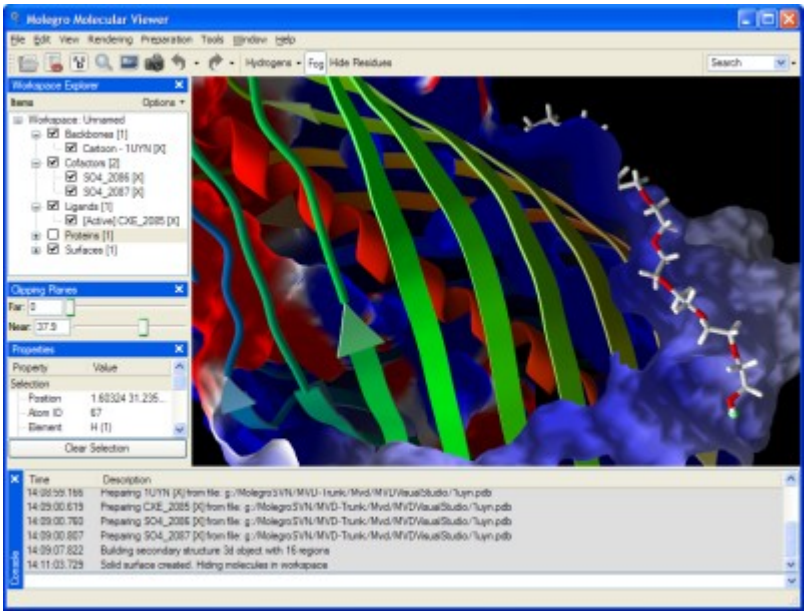
molegro molecular viewer

user manual

MMV 1.2 for Windows, Linux, and Mac OS X



Molegro ApS

Table of Contents

1 Introduction

Molegro Molecular Viewer (MMV) is an application for studying and analyzing how ligands interact with macromolecules.

MMV can be used to:

- Inspect docking results consisting of high-scoring poses found by Molegro Virtual Docker (MVD) – the molecular docking software product offered by Molegro.
- Inspect and visualize molecular structures obtained from other sources, such as the Protein Data Bank.

This manual describes various aspects of MMV from how to use the GUI, importing, preparing, and visualizing molecules to inspecting and analyzing docking results from Molegro Virtual Docker.

Notice: The main focus of MVD and MMV is on studying protein-ligand interactions. MMV is currently not supporting DNA and RNA molecules. It is possible to import DNA and RNA molecules in MMV but they will appear as ligand molecules.

1.1 Contact Information

Molegro Molecular Viewer is developed by:

Molegro ApS
Høegh-Guldbergs Gade 10, Building 1090
DK-8000 Aarhus C
Denmark
www.molegro.com

E-mail:

- General inquiries: info@molegro.com
- Product support: support@molegro.com
- Reporting bugs: bugs@molegro.com

1.2 System Requirements

The system requirements for Molegro Molecular Viewer are:

- Windows Vista, XP, 2003, or 2000
- Linux: Most standard distributions. We provide both 32 and 64 bit builds such as Fedora Core 3 (or later versions) and Red Hat. Please send a mail to support@molegro.com if the program does not work on a particular distribution – and we will try to provide a new build.
- Mac OS X 10.4 (or later versions). We provide both PowerPC and Intel builds.

1.3 Reporting Program Errors

If you discover a program error, please mail the information to:

bugs@molegro.com

Remember to specify how the error can be reproduced, the version number of Molegro Molecular Viewer in question, and the operating system that was used. If possible, inclusion of molecular files used (e.g. Mol2, PDB, MVDML) will make it easier for us to reproduce (and correct) the error.

1.4 Text Formats Used in the Manual

The following formatting styles are used in this manual:

- All GUI text, labels, and keyboard shortcuts are written in bold face with initial capital letters.

  Examples: **Workspace Explorer**, **Fixed Color**, **Ctrl-O**

- Menus and menu items are identified using dividing lines and bold face.

  Example: **View** | **Docking View** indicates that the user should first select the **View** menu and then select the **Docking View** menu item.

- Filenames are written in mono-spaced font.

  Example: `\Molegro\MMV\bin\mmv.exe`

1.5 Keyboard Shortcuts

The keyboard shortcuts used in the manual works for Windows and Linux

versions of MMV. On Mac OS X, the **CTRL** key is replaced by the **command** key and function key shortcuts (e.g. **F1**) should be invoked by pressing the function key and the **fn** key (e.g. **fn+F1**).

1.6 Screenshots Used In the Manual

The screenshots used in the manual are taken from the Windows XP version of MMV. Therefore, dialogs and other GUI related material may slightly differ on Linux and Mac OS X versions.

1.7 Future Updates

Molegro Molecular Viewer contains a built-in version checker making it easy to check for new program updates including new features and bug fixes. To check for new updates, select **Help** | **Check for Updates**. A window showing available updates and details about changes made will appear.

2 User Interface

2.1 Basic Concepts

Molegro Molecular Viewer is based on the notion of workspaces. The *workspace* is the central component and represents all the information available to the user in terms of molecules (proteins, ligands, cofactors, water molecules, and poses), user-defined constraints (visualized as small spheres), cavitities (visualized as a grid mesh), and various graphical objects (molecular surfaces, backbone visualizations, labels, etc.).

By default, an empty workspace is shown when starting MMV. A workspace can be saved, cleared, or replaced by other workspaces. The content of the current workspace is listed in the **Workspace Explorer** window, which also allows for manipulation of the various items available (see Section 2.4 for more details).

Notice: When saving a workspace in the internal MVDML format only molecules, cavities, and constraints are stored – all 3D visualization objects and the color and rendering settings are ignored. For more information about the MVDML format see Appendix I and Section 5.2.

MMV can visualize constraints and cavities when importing MVDML files created by Molegro Virtual Docker (MVD) but MMV does not support creation of constraints and detection of cavities on its own.

2.2 Overview

The user interface in MMV is composed of a central 3D view (referred to as the **Visualization Window** or 3D world, together with a number of dockable windows (introduced below).

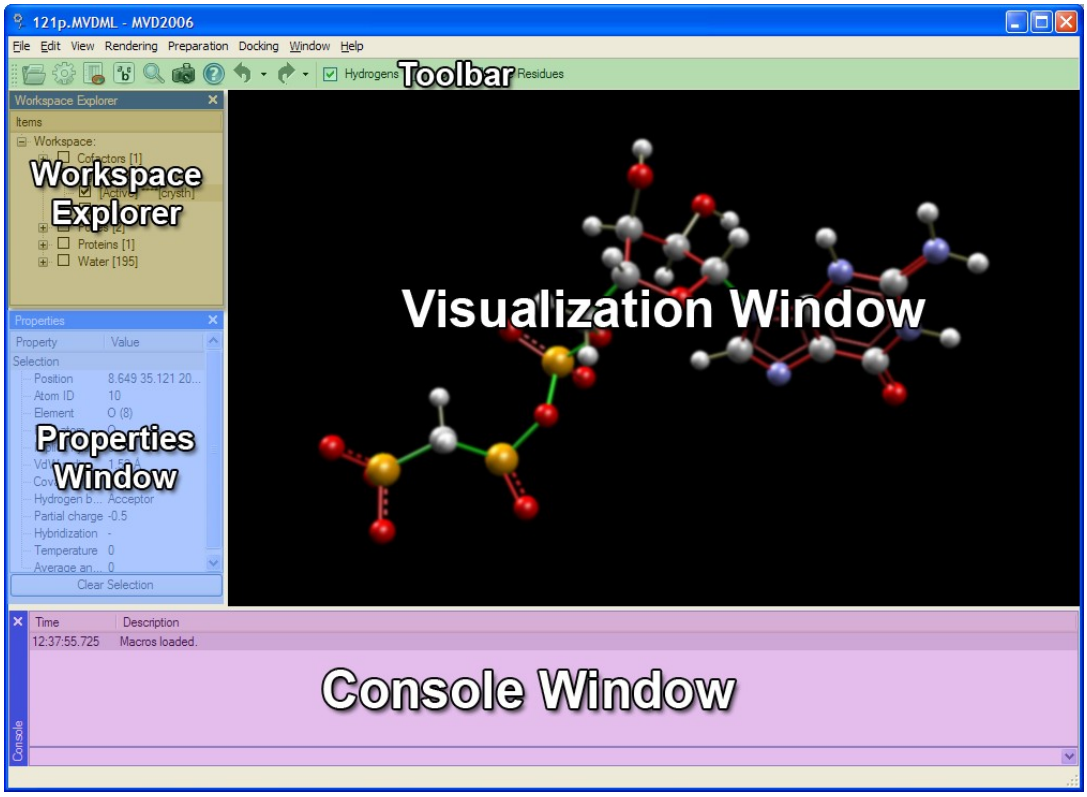


Figure 1: Main application window.

2.3 Toolbar

The **MMV Toolbar** provides easy and fast access to commonly used actions, such as import of molecules and pose inspection using the **Pose Organizer**.

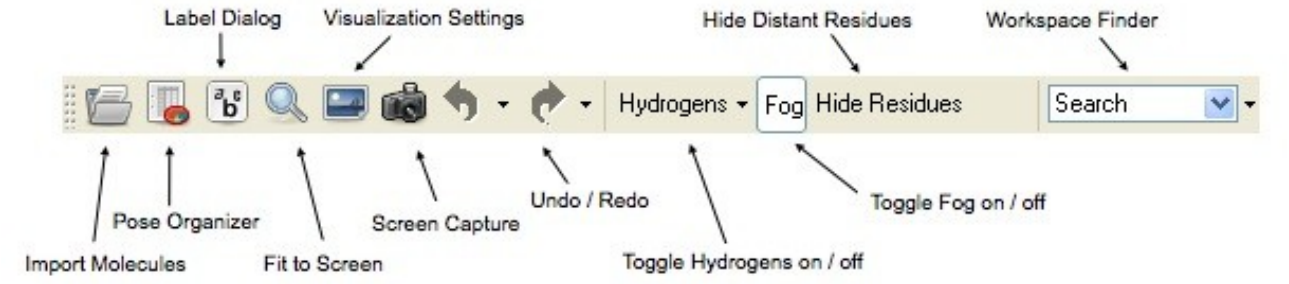


Figure 2: MMV Toolbar.

The **MMV Toolbar** also contains three toggle buttons. The **Hydrogens** button makes it easy to switch between different view modes (**Show all hydrogens**, **Show only polar hydrogens**, and **Hide all hydrogens**). The **Fog** button is used to toggle fog effects on and off. The **Hide Residues** button is used to toggle whether residues should be hidden or not (see Section 2.9 for more details). The **Workspace Finder** located at the far right side of the toolbar can

be used to quickly search for molecule names and residue/atom IDs (see Section 2.10 for more details).

2.4 Workspace Explorer

The **Workspace Explorer** window (see Figure 3) contains information about the 3D objects (both molecules, such as proteins, ligands, and water molecules - but also objects such as labels, surfaces, backbones, and cavities).

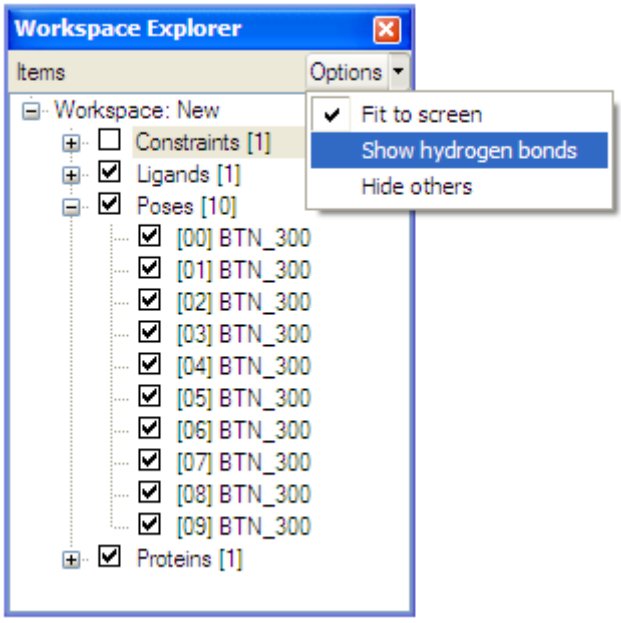


Figure 3: Workspace Explorer window.

The context menu (right mouse button click) allows the user to:

- Export molecules to PDB, Mol2, or SDF format
- Edit workspace properties (workspace title and workspace notes)
- Rename molecules
- Remove items from the current workspace
- Set the currently active ligand or reference ligand (optional)
- Copy ligands to poses (used to inspect ligands with the **Pose Organizer**)
- Clone ligand or protein (makes a copy of the molecule)
- Convert ligand to pose or cofactor
- Convert protein to ligand
- Convert pose to ligand (used when docking poses)
- Inspect poses (using the **Pose Organizer**)

- Prepare molecules
- Create labels, surfaces, and backbones
- Fit the molecule to the visualization window

Inspecting Molecules

The **Workspace Explorer** can also be used to inspect molecules in the **Visualization Window** using the left mouse button to select the molecules or by using keyboard shortcuts (see below).

The **Options** button (see Figure 3) contains settings used to customize the behavior when inspecting molecules. The **Fit to screen** option will automatically zoom selected molecules so that they fit into the **Visualization Window**. The **Show hydrogen bonds** option can be used to display hydrogen bonds (only applicable for ligands and poses). The **Hide others** option toggles whether other checked molecules in the current workspace category are allowed or not.

Keyboard shortcuts are also available for inspecting molecules. Pressing the **Shift** button while clicking the left-mouse button on a molecule in the chosen category (e.g. Ligands or Poses) will fit the selected molecule in the **Visualization Window** and all other molecules located in the same category are hidden.

Alternatively, using **Ctrl+Shift** when clicking on a molecule, hydrogen bonds are shown for the selected molecule.

Instead of using the mouse to select molecules to inspect, **Up** or **Down** keys can be used to browse the molecules present in the currently selected **Workspace Explorer** category. If the **Ctrl** and **Shift** shortcuts are omitted, the settings enabled in the **Options** panel will be used.

2.5 Properties Window

The **Properties Window** contains information about the currently selected (or highlighted) 3D object(s) in the **Visualization Window** and provides useful information while preparing and modifying the molecules.

Figure 4 shows an example of different properties for a highlighted atom.

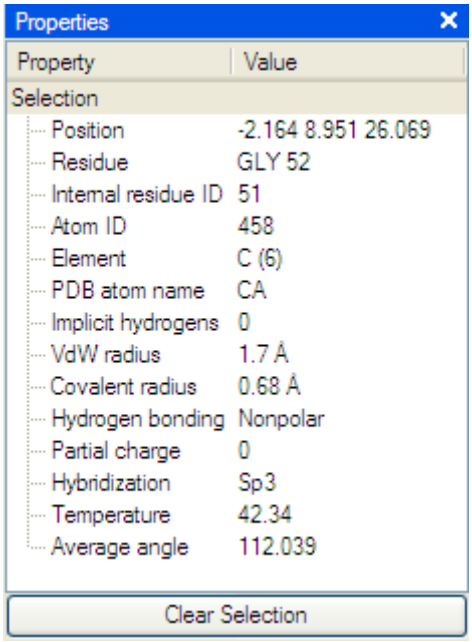


Figure 4: Example of properties for a selected atom.

2.6 Visualization Window

The **Visualization Window** (see Figure 5) visualizes all the selected molecules in the workspace and all custom graphical objects (e.g. labels, annotations, charges, backbones, surfaces, and cavities).

Notice: Cavities can be imported and visualized from MVDML files but not created in MMV. MVD is required to create cavities.

For large molecules it can be computationally slow to display all atoms. Therefore it is recommended to adjust the view to the user's needs. Often it is a good idea to add a molecular surface (perhaps transparent) to give some idea of the 3D structure. Alternatively, switching to wireframe visualization style and hiding non-polar (or all) hydrogens atoms can also improve the visualization speed significantly. Also consider cropping (removing) non-relevant parts of the complex, in order to make the visualization faster. Cropping is described in Section 2.9.

Changing the 3D World Appearance

The visualization engine is highly configurable.

Molecules can be drawn as lines (wireframe), ball-and-sticks, capped-sticks, and space-fill (CPK).

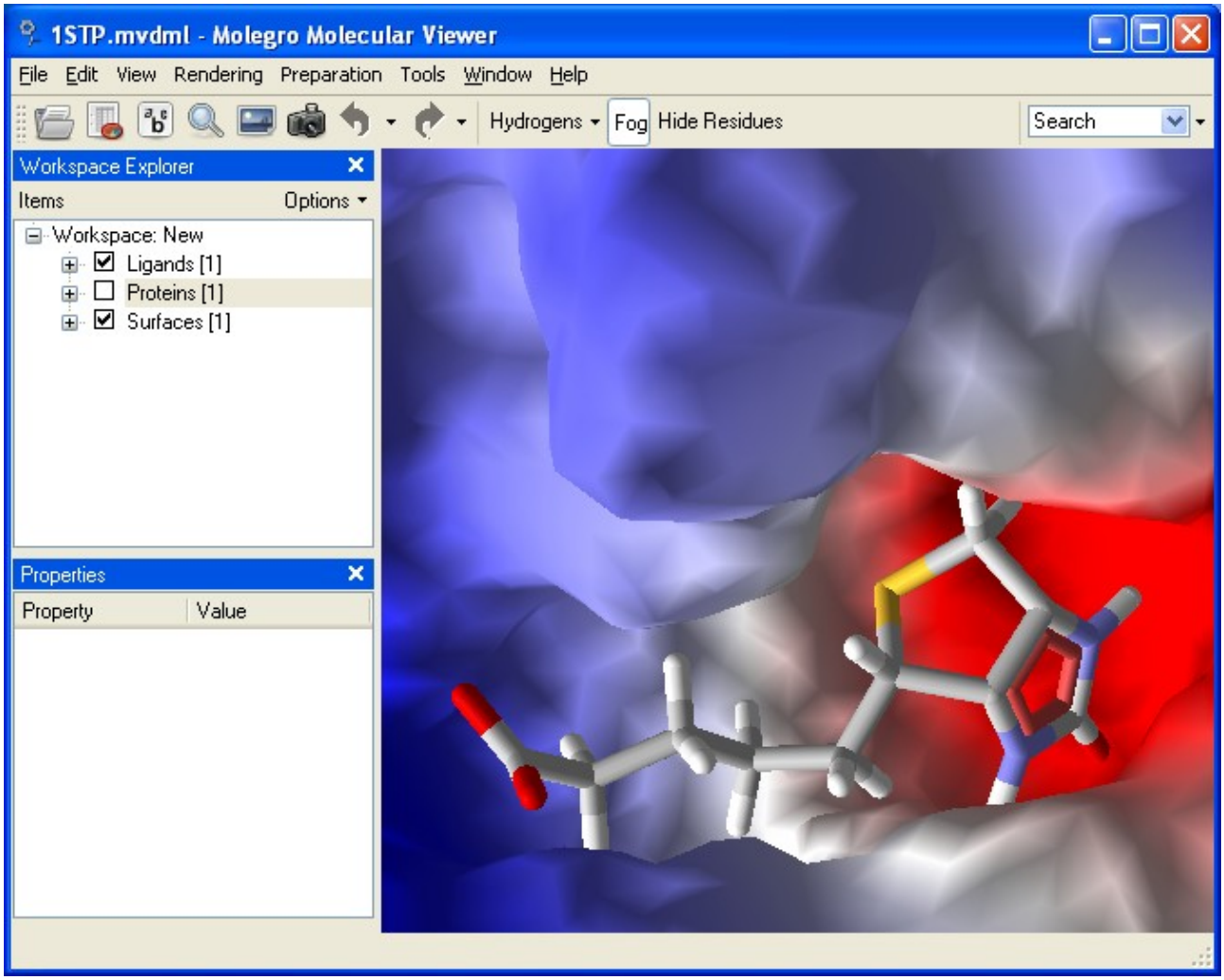


Figure 5: Visualization of Biotin (1STP) in capped-stick style and electrostatic protein surface.

Notice: Ball-and-stick is the preferred style for handling preparation of ligands, since the visualized bond shows bond order, and is color coded to display whether the bond is set rigid (brown or red) or flexible (green).

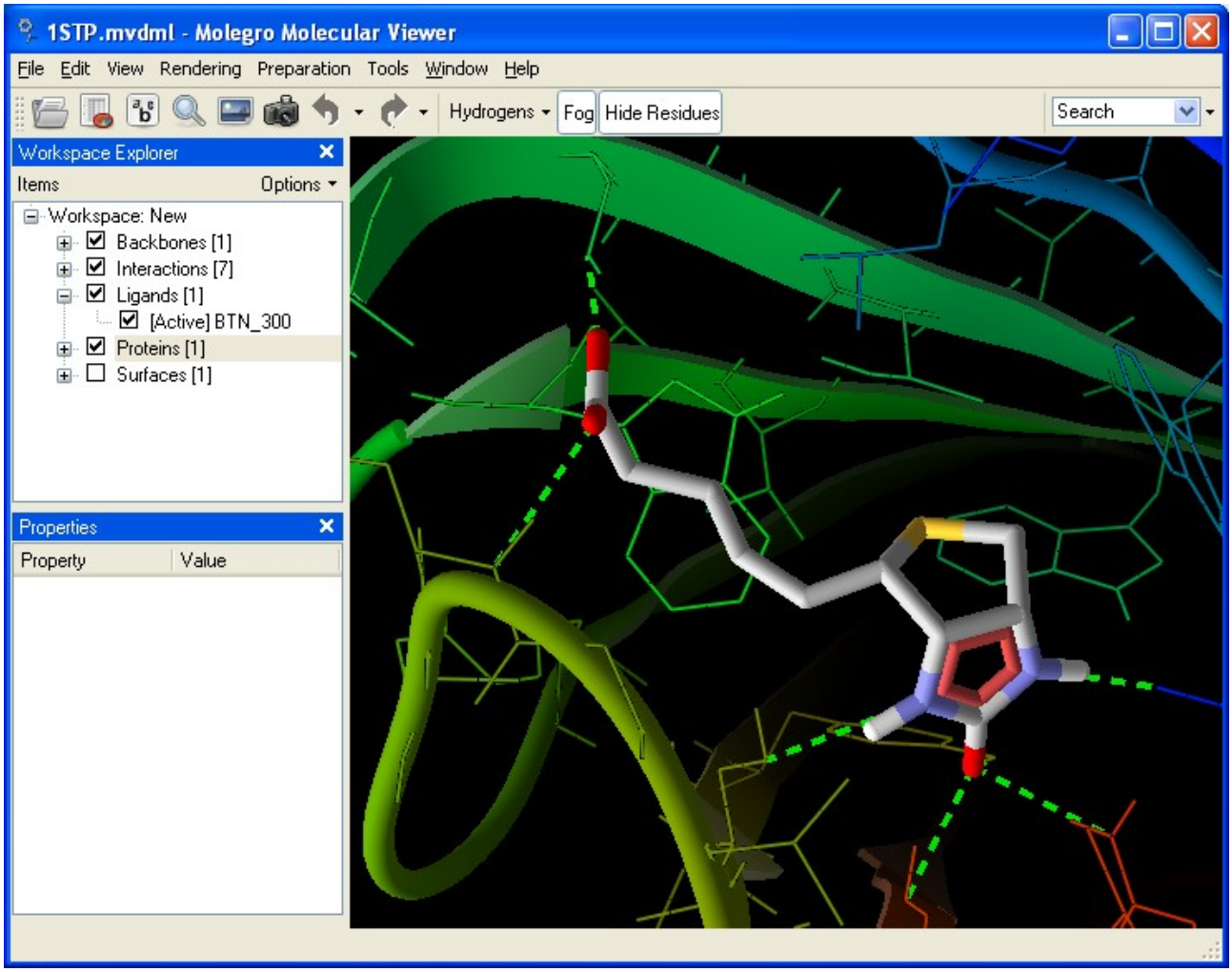


Figure 6: Main window showing different visualization styles.

The easiest way to get acquainted with the different drawing modes is to try the preset modes listed in the **Rendering** menu or to use the **Visualization Settings** dialog to inspect and modify visualization settings (described in Section 2.18).

Navigating the 3D World

Mouse actions available in the 3D world:

Function	Action
Zoom	By pressing both mouse buttons and moving up and down. By using scroll wheel. By using shift and left mouse button.
Free Rotation	Dragging mouse cursor while holding left mouse button down.
Drag Atom Rotation	While holding mouse over an atom: Dragging mouse (left mouse button down) will force the atom to follow the mouse cursor.
Free Translation	Dragging mouse cursor while holding right mouse button down.
Show Context Menu	Click and release right mouse button.

All rotations are centered about the rotational center.

This center can be chosen by invoking the context menu on an atom (right mouse button click) and selecting **Set as Rotational Center**. Another option is to choose **Fit to Screen** from the **Workspace Explorer** context menu. **Fit to Screen** will set the rotational center to the center of the bounding box enclosing the chosen molecule. If **Fit to Screen** is invoked from the **MMV Toolbar** or from the **Visualization Window** context menu, the new rotational center will be the center of the bounding box enclosing all visible molecules in the **Visualization Window**.

Manipulating Visualization Objects

All objects in the 3D world have context menu actions. These can be used for changing their properties, e.g. setting hybridization, partial charge, implicit hydrogens, or hydrogen bond types for atoms and bond order or bond flexibility for bonds. See Section 3.3 for more details.

2.7 Console Window

The **Console Window** (at the bottom of the screen) displays information, warnings and errors. The input field at the bottom of the console window

allows the user to enter console commands. The amount of information in the console can be controlled with the associated context menu (right mouse button click) - e.g. info, warnings, and debug messages can be turned off.

2.8 Clipping Planes

Clipping Planes allows you to change the clipping planes of the visualization window, i.e. how close and how far away objects are drawn. This can for example be useful if you want to visualize the interior of a protein or a ligand deeply buried inside a macromolecule.

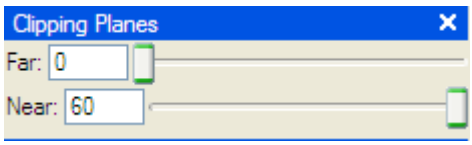


Figure 7: Clipping Planes dockable window.

Clipping Planes can be enabled by choosing **Window** | **Clipping Planes...** from the menu bar. Clipping Planes are enabled when the **Clipping Planes** window is shown and disabled when it is closed. Adjust the near and far slider until the desired region is shown.

2.9 Hiding Distant Residues

The **Hide Residues** dialog (see Figure 8) allows you to hide residues outside of a user-defined sphere: this can for example be used to show only the relevant residues near the binding site of the protein.

It is possible to set the center to of the sphere to the following objects if they are part of the workspace: The center of the protein(s), the center of the **Active Ligand**, the center of the **Reference Ligand**, the center of any cavity in the workspace, the center of the currently defined search space, or the center of a selection of atoms (if any). The residues are dynamically shown/hidden when the **Sphere radius** slider is moved.

The lower pane of the **Hide Residues** dialog allows you to restrict the types of residues shown by toggling the appropriate button. If a given residue type is not contained in the sphere defined in the panel above, the button corresponding to the type will be grayed and can not be toggled.

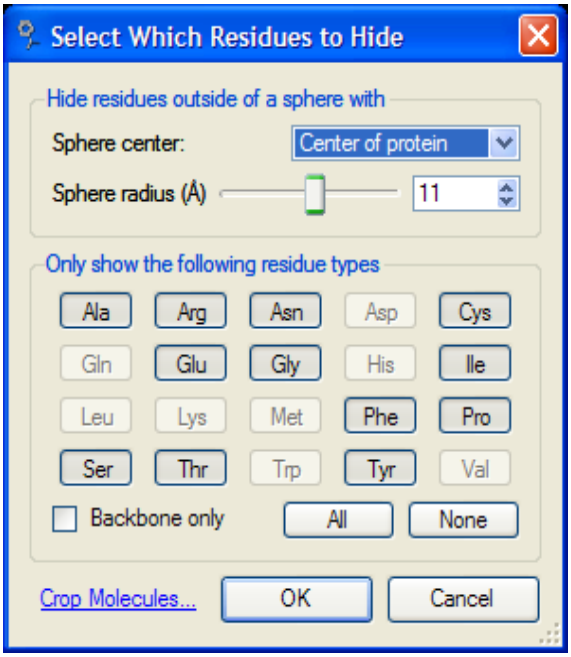


Figure 8: Hide residues dialog.

The **Backbone only** check box can be used to toggle whether side-chains are visible or not.

The **Hide Residues** dialog can be invoked by pressing the **Hide Residues** button in the **MMV Toolbar**. In order to show all protein residues again, select the **Hide Residues** button on the **MMV Toolbar**.

Cropping. It is possible to delete molecules from the workspace in order to remove non-relevant regions. To crop molecules, invoke the **Hide Residues** dialog and adjust the visible sphere to the desired size before clicking the **Crop Molecules...** button. A dialog will show which structures will be kept (the checked molecules) and which will be discarded. Notice that proteins are cropped on a per-residue basis: residues outside the cropping sphere will be discarded. All other molecule types are kept or discarded in their entirety.

2.10 Workspace Finder

The **Workspace Finder** located in the **MMV Toolbar** (see Figure 10) allows you to quickly search for molecule names and residue/atom IDs in the workspace. When a name or ID number (or part of it) is typed in the search box, the **Workspace Finder** will present a list of matches (a maximum of 30 matches is returned). It is also possible to search in atom coordinates by prepending the search with a '!' (e.g. searching for '!1.23' will return atoms where one of the coordinates starts with 1.23).

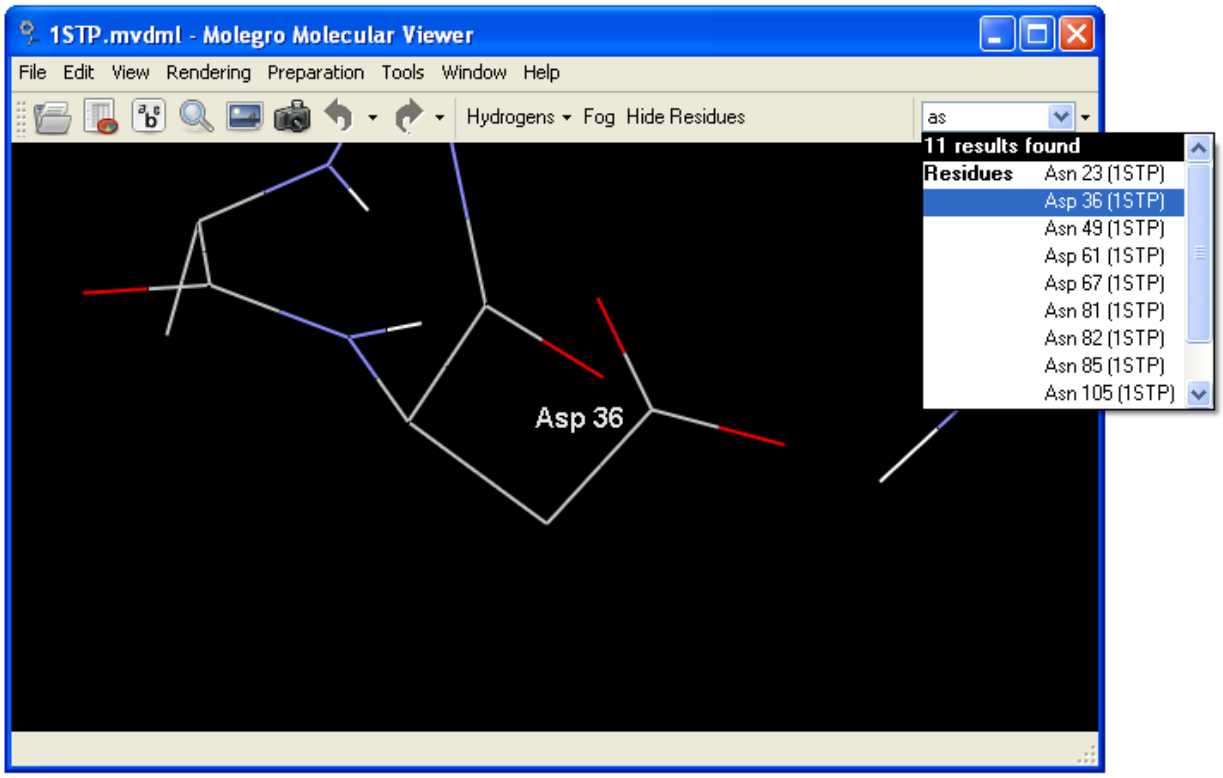


Figure 9: Workspace Finder dialog.

By default, the **Fit to screen** option is enabled so that items (molecules, residues, or atoms) are fitted to the **Visualization Window** while browsing the list of results found. The **Fit to screen** option can be disabled in the options panel invoked by pressing the small button on the right hand side of the **Workspace Finder** search box.

The **Workspace Finder** is invoked by typing characters in the search box (text field) located in the far right side of the **MMV Toolbar**. A result is selected by pressing the **Return** key. Pressing the **Escape** (Esc) key or mouse-clicking outside the **Workspace Finder** window will cancel the current search query.

2.11 Sequence Viewer

The **Sequence Viewer** dialog (see Figure 10) allows you to inspect protein residues in an easy manner. Using the context menu on the **Sequence Viewer** window it is possible to select residue atoms in the **Visualization Window**, hide non-selected residues, change between one and three-letter residue names, and toggle details about secondary structure.

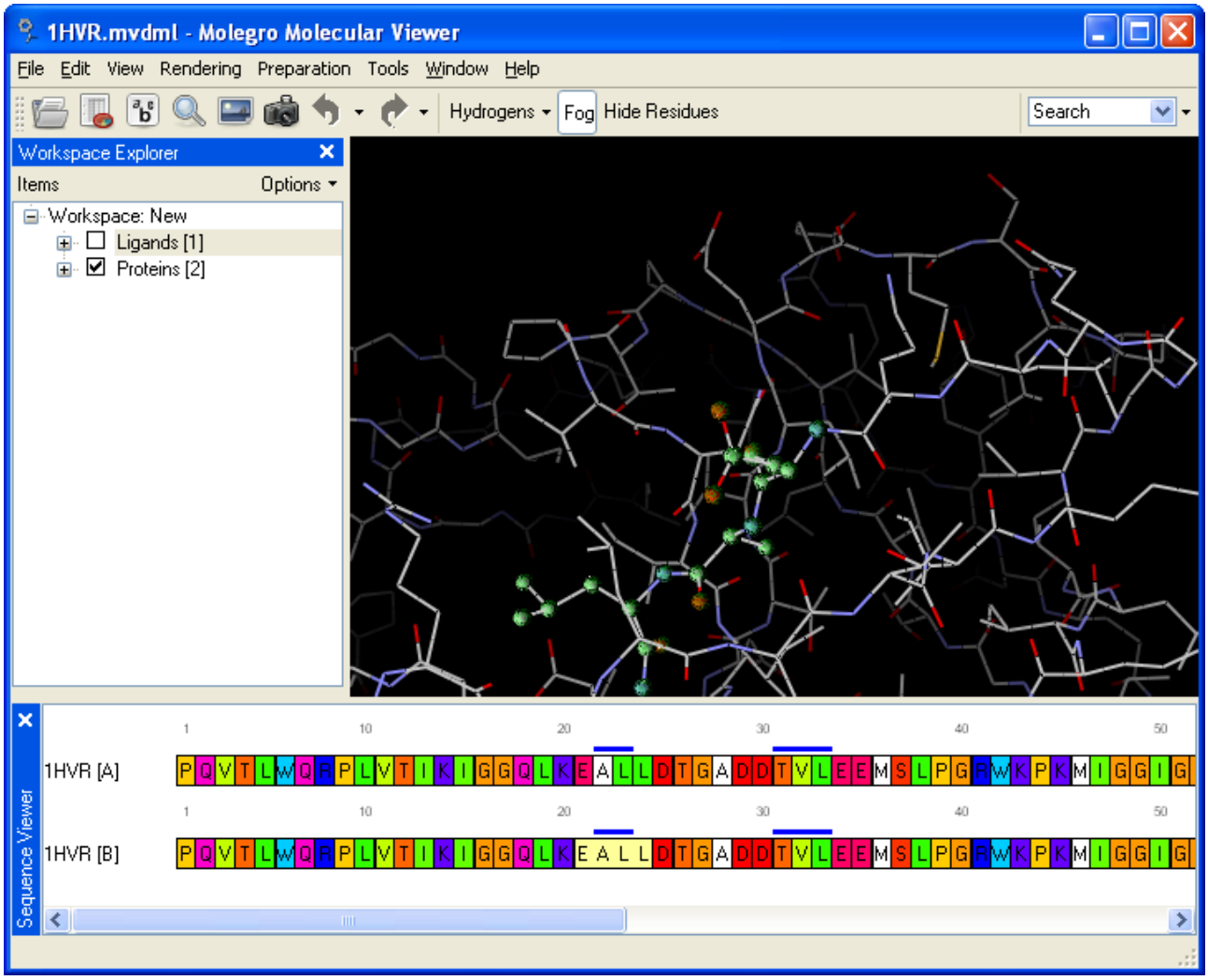


Figure 10: Sequence viewer with selection of four residues highlighted in the Visualization Window.

The **Sequence Viewer** dialog can be invoked by selecting **Window** | **Sequence Viewer**' or using the **Ctrl-Shift-S** keyboard shortcut.

2.12 Workspace Properties

Workspaces can contain user-specified notes. Further, the title of the workspace can be changed using the **Workspace Properties** dialog. The **Workspace Properties** dialog can be found in the **Edit Properties...** context menu on the **Workspace** item in the **Workspace Explorer** or via **Edit** | **Workspace Properties...** (see Figure 11).

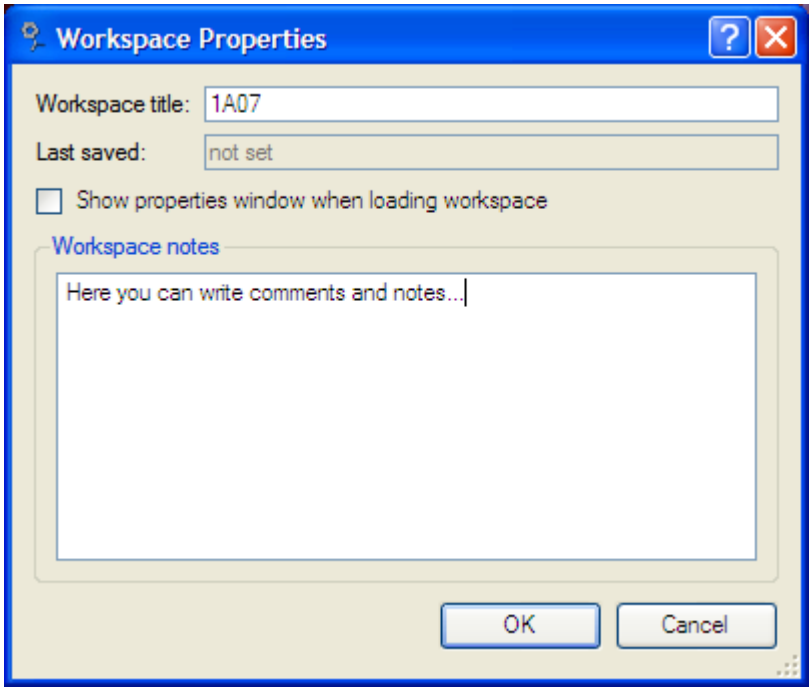


Figure 11: Workspace Properties dialog.

2.13 Measurements and Annotations

Distances and angles can be measured directly in the 3D world (see Figure 12).

If two atoms are selected, the distance between them will be shown in the **Properties Window**.

If three connected atoms are selected, the angle that they span will be shown in the **Properties Window**.

If no atoms are selected, and a bond is highlighted, the field **Torsion Angles** in the **Properties Window** will show the torsion angle(s), defined through this bond.

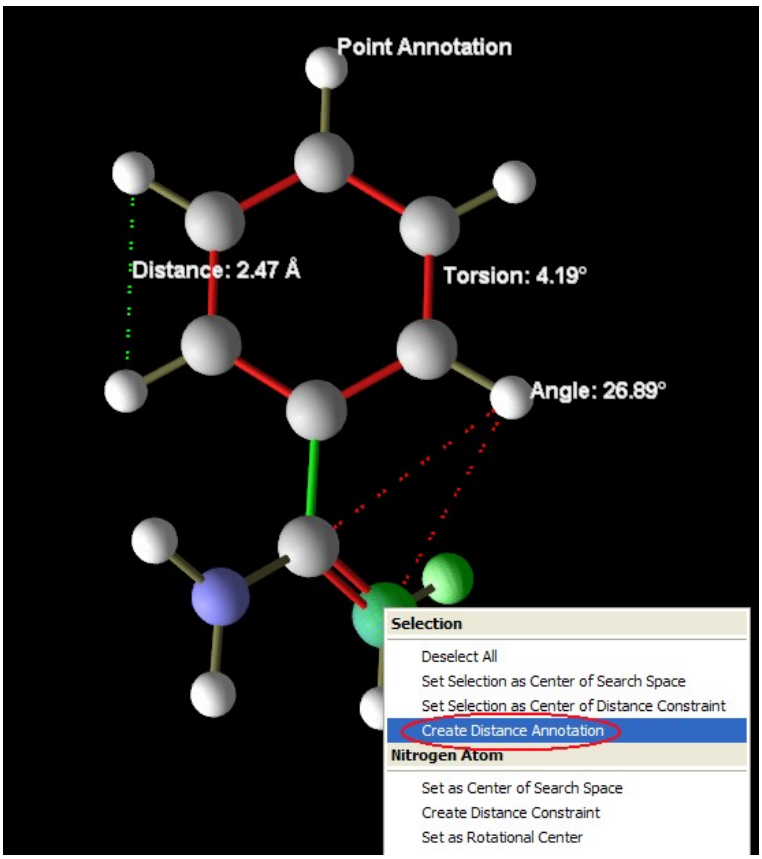


Figure 12: Annotations and measurements.

Measurements can also be made permanent as annotations. There are different kinds of annotations. To create annotations, select 1-4 atoms and use the context menu (right-click mouse button) and choose **Create … Annotation**. The text can be edited before the annotation label is created. Annotations are added to the **Workspace Explorer** category: **Annotations**

2.14 Creating Labels

To create labels use the **Create Label** dialog, which can be invoked via **Create Labels…** in the **Workspace Explorer** context menu (on molecular categories: **Proteins**, **Ligands**, and **Poses**) or via the **Tools** | **Labels** menus.

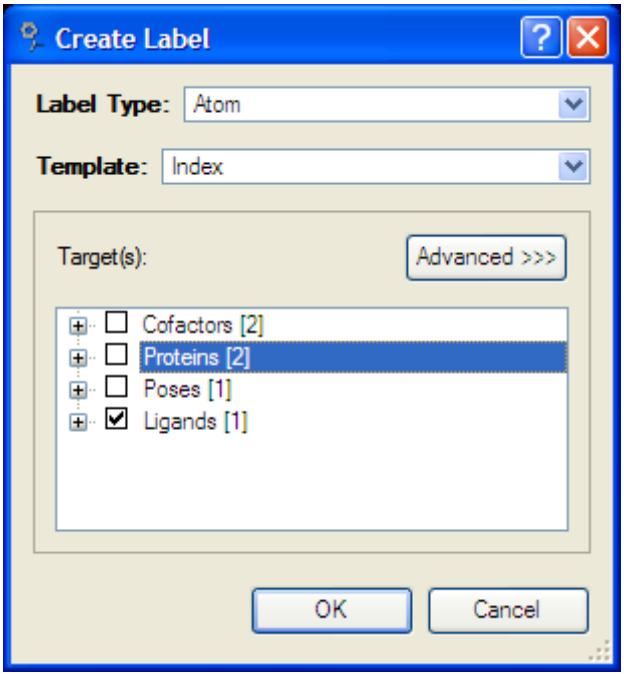


Figure 13: Creating a new label.

The **Create Label** dialog makes it possible to label different *object levels*: atoms, bonds, molecules, residues or torsion trees. The labels can be chosen from a list of standard templates or constructed from a list of available variables (using the **Advanced** tab).

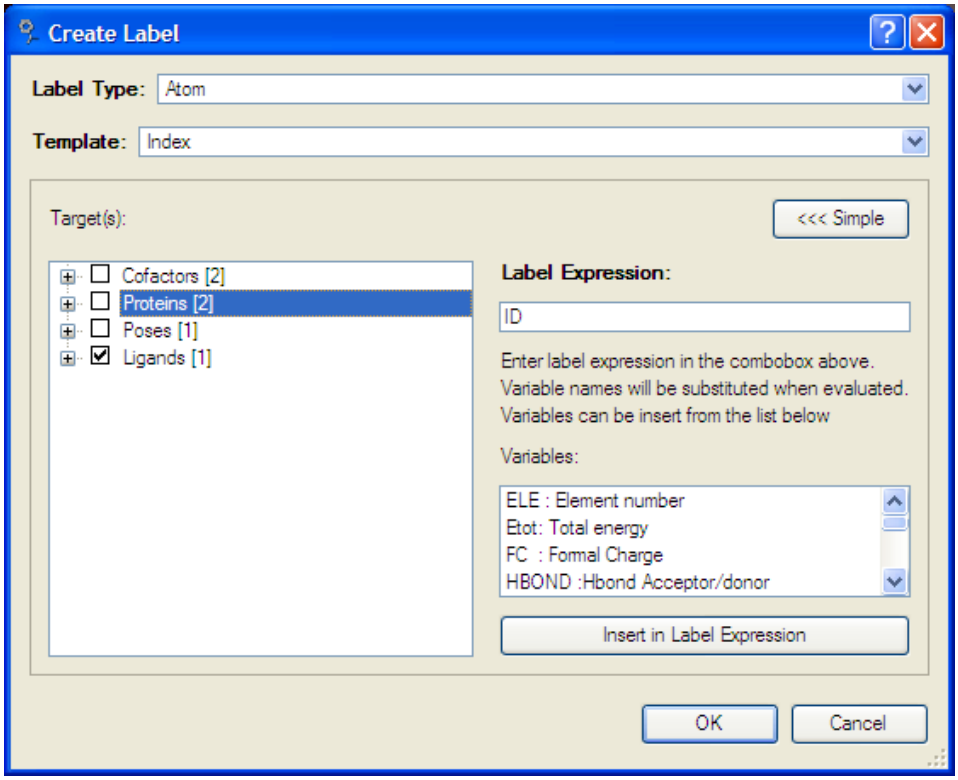


Figure 14: Advanced label expression dialog.

Labels will occur in the **Labels** category in the **Workspace Explorer** and can be removed or hidden using the context menu or by pressing the labels tool bar button.

2.15 Creating Molecular Surfaces

Surfaces can be created for all molecular objects, and subsequently customized.

In MMV surfaces are created by probing points on a uniformly spaced grid. It is possible to adjust the grid resolution (**Resolution**) and probe size (**Probe Radius**) under **Advanced** settings.

Two types of surfaces are available:

Expanded Van der Waals – this is an approximation to the surface created by expanding the Van der Waals radius of each atom with the **Probe Radius**.

Molecular surface – this is an approximation to the surface defined by the contact area of the probe and Van der Waals sized spheres.

Surfaces can be colored by **Hydrophobicity**, **Electrostatic Potential**, or **Solid Color**. Surfaces can be drawn transparently, as dots, lines, or solid polygons.

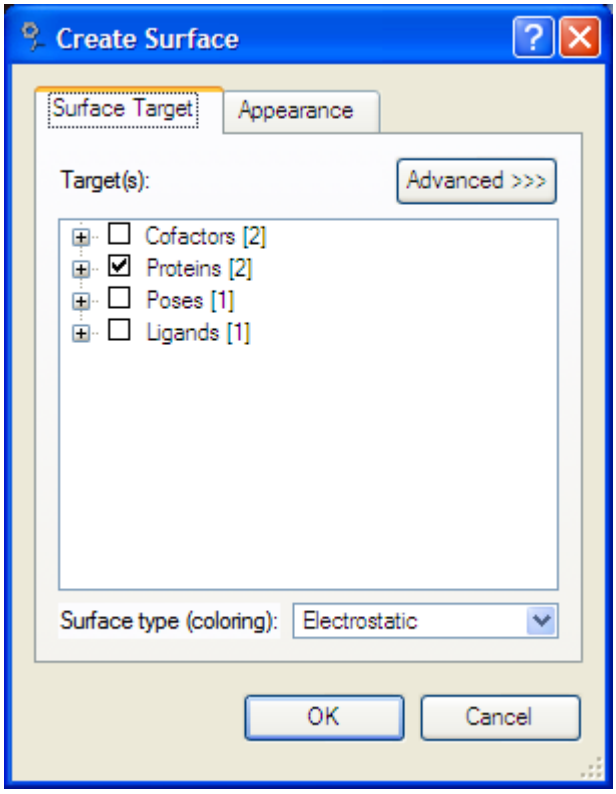


Figure 15: Creating a new surface.

Surfaces can be created via **Create Surface...** from the context menu in the **Workspace Explorer** or via **Tools** | **Surfaces**.

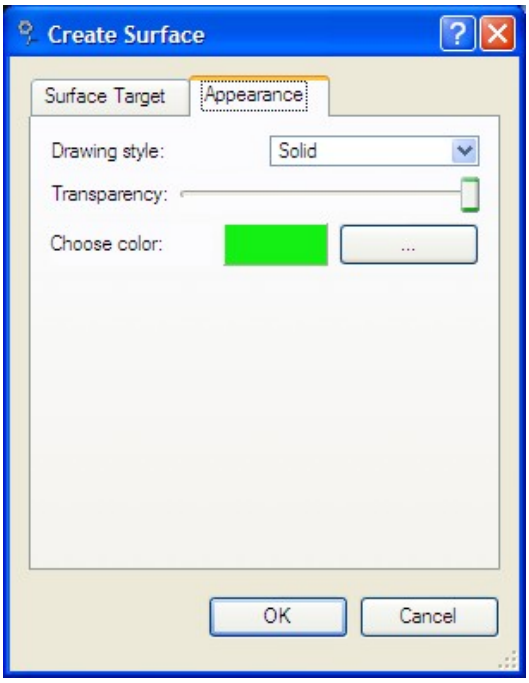


Figure 16: Changing surface appearance.

2.16 Creating Protein Backbone Visualizations

The backbone of the protein can be visualized by using the **Create Backbone Visualization** dialog. The dialog can be invoked by using the context menu on the **Proteins** category (or a single protein item) in the **Workspace Explorer**.

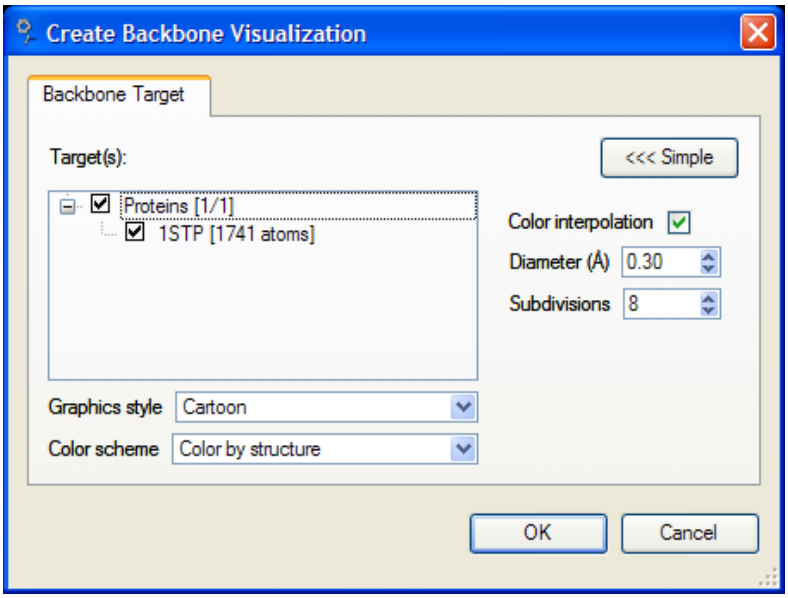


Figure 17: Creating a new backbone.

The **Create Backbone Visualization** dialog allows you to select which proteins (or protein chains) the backbone should be visualized for.

Two main graphics styles can be used. The **Cartoon** style visualizes the secondary structure of the protein(s) using arrows to represent beta sheets and helical lines for alpha helices (see Figure 18).

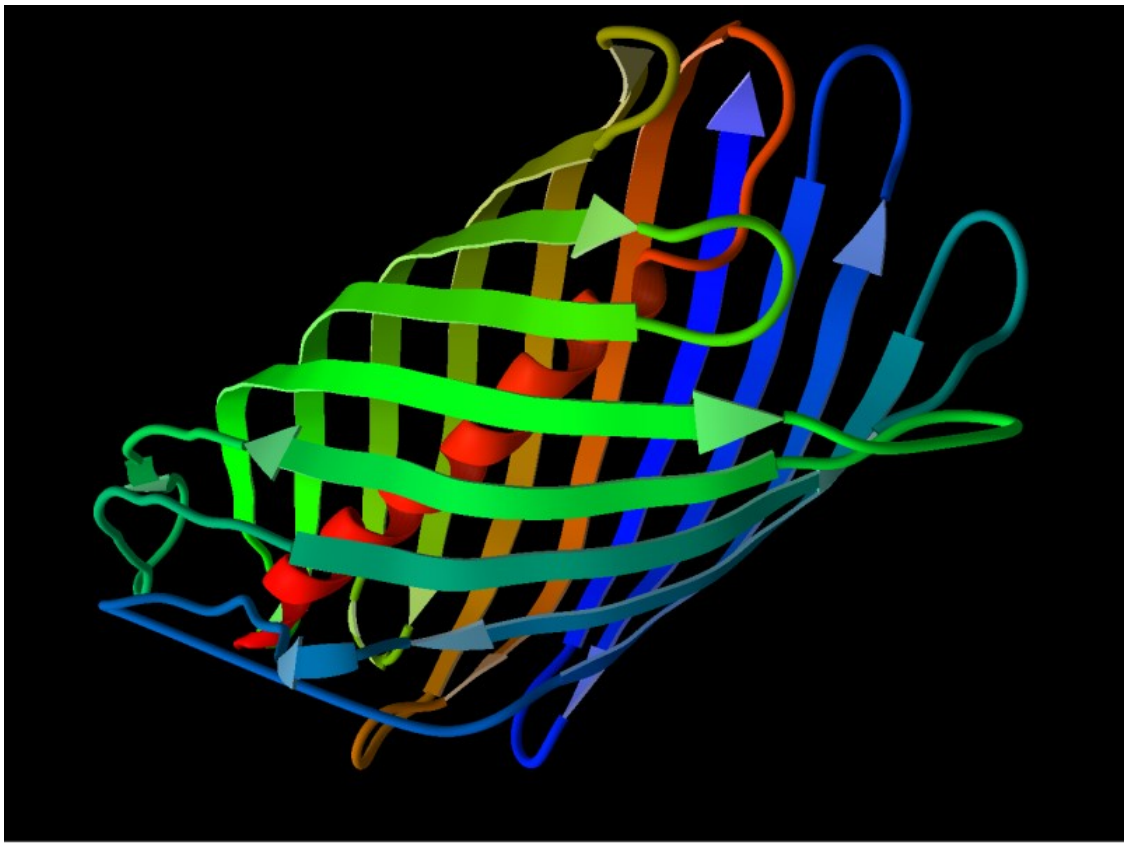


Figure 18: Cartoon graphics style.

If the **Tube** graphics style is used, the backbone is visualized as a spline (a piecewise parametric polynomial curve) interpolating the positions of the alpha carbons in the backbone (see Figure 19).

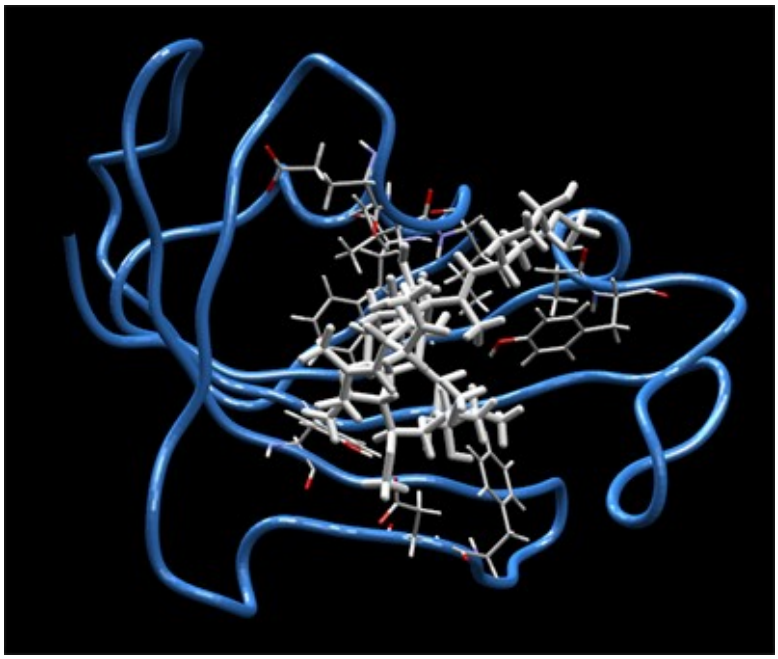


Figure 19: An example of a protein backbone using the Tube graphics style.

It is also possible to set the color scheme for the backbone. **Color by structure** colors the backbone based on the secondary structure information (alpha helices are colored yellow, beta sheets are colored blue, and coil is colored gray). **Color by residue position** colors the backbone based on the residues order of occurrence creating a rainbow color effect. **Color by chain** colors each individual protein chain in a different color. **Color by atom** colors the backbone by using the currently shown color of the protein backbone atoms (the color used is taken from the C-alpha atom).

On the advanced panel, the **Color interpolation** check box allows you to determine whether the backbone color should be interpolated between the atoms it passes through or should be held constant between atoms. **Diameter (Å)** sets the width of the backbone in angstrom, **Subdivision** sets the resolution of the backbone (the number of subdivisions between each residue in the protein).

Backbones appear in the **Backbones** category in the **Workspace Explorer** and can be removed via the context menu or hidden using the check box.

2.17 Making Screenshots

Screenshots can be made by choosing **Window** | **Capture Screen**. It is possible to specify whether to capture the Visualization Window only (the 3D view) or the entire Desktop (see Figure 20). The captured region can be saved in JPG, BMP, or PNG file formats.

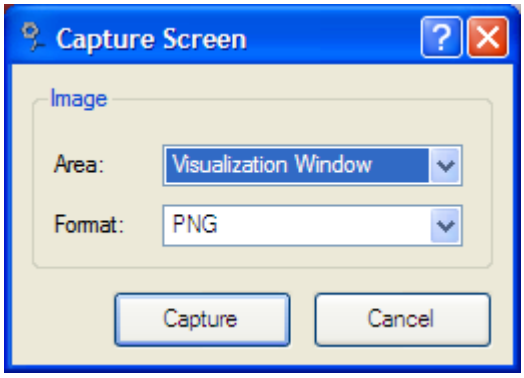


Figure 20: Screen Capture dialog.

2.18 Visualization Settings Dialog

The graphical settings for the 3D visualization can be adjusted by selecting **Rendering** | **Visualization Settings Dialog**.

Graphical Styles and Coloring Schemes

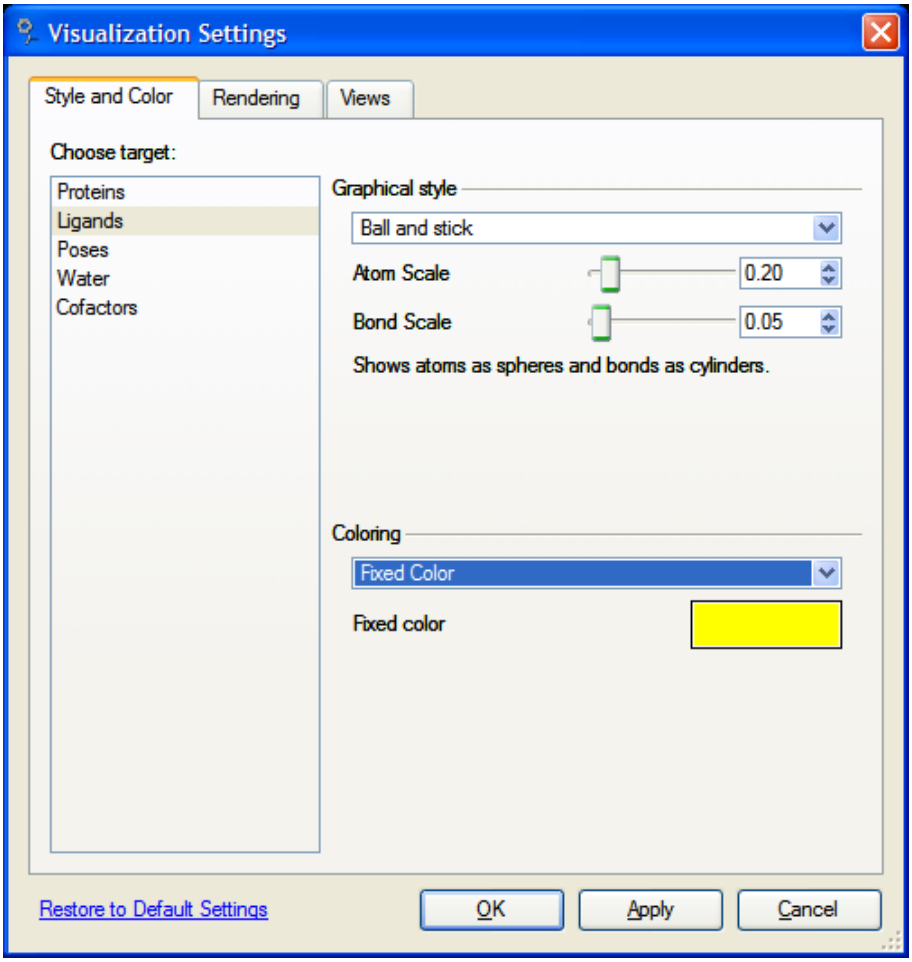


Figure 21: The Visualization Settings dialog.

From the **Style and Color** tab, select a category from the list on the left side of the tab (one of 'Proteins', 'Ligands','Poses','Water', and 'Cofactors') and adjust either its graphical style or color scheme.

The following graphical styles can be chosen:

- **Ball and Stick**. Atoms are drawn as spheres (balls), and bonds are drawn as cylinders (sticks). The **Atom Scale** parameter sets the fraction of the Van der Waals radius that is used as radius for the sphere. **Bond Scale** is the diameter of the bonds in Ångstrom. This is the preferred graphical style for modifying and inspecting bond and atom properties (since the bond order is visualized and the atoms are easy to select).
- **Stick**. Bonds are drawn as cylinders. **Bond Scale** is the diameter of the bonds in Ångstrom.
- **Spacefill (CPK)**. Atoms are drawn as spheres (balls). Bonds are not drawn. The **Atom Scale** parameter sets the fraction of the Van der Waals radius that is used as radius for the sphere.
- **Wireframe**. This is by far the fastest way to draw molecules. Bonds are drawn as lines between atoms. No atoms are drawn (but notice that it is still possible to do atom selections in the GUI). Notice all bonds are drawn as single lines (double bonds and delocalized bonds are also drawn as single lines). It is possible to adjust the line width in pixels (Notice that not all OpenGL implementations support non-integer line widths).

The following coloring styles can be applied to all molecules:

- **Fixed Color** - A user-defined color.
- **Color By Element (CPK)** - Atoms are colored according to element type.
- **Color By Id (or Chain)** - Molecules are colored according to their internal molecule ID (i.e. a single ligand will be uniformly colored, but all ligands will have different colors).
- **Color By Id (carbons only)** - Same as above, except only carbons are colored using this scheme. Other atoms are colored according to element type.
- **Color By Hydrogen Bond Type** - Colors atoms according to hydrogen bonding properties (donors are red, acceptors green and atoms capable of both donating and accepting hydrogens are yellow).
- **Color By Partial Charge** - Colors according to electrostatic partial charge (blue corresponds to positive charge, red to negative charge).

The following can only be applied to proteins:

- **Color By Temperature (B-Factor)** - The temperature factor is a measure of how much a given atom vibrates around its position in the crystal structure. Notice that this information is not always present in PDB-files, and that it is sometimes used for other purposes. The colors will be interpolated between blue for the minimum temperature and red for the maximum temperature.
- **Color By Amino Acid Type** - Colors proteins according to their residue type.
- **Color By Shapely Residue Scheme** - Same as above with alternative colors.
- **Color By Residue ID** - Colors according to residue ID (rainbow effect).
- **Color By Secondary Structure** - Colors according to secondary structure (red for helices, blue for strands and yellow for turns).
- **Color By Hydrophobicity** - Green for non-polar atoms, red for polar atoms.

Rendering Settings

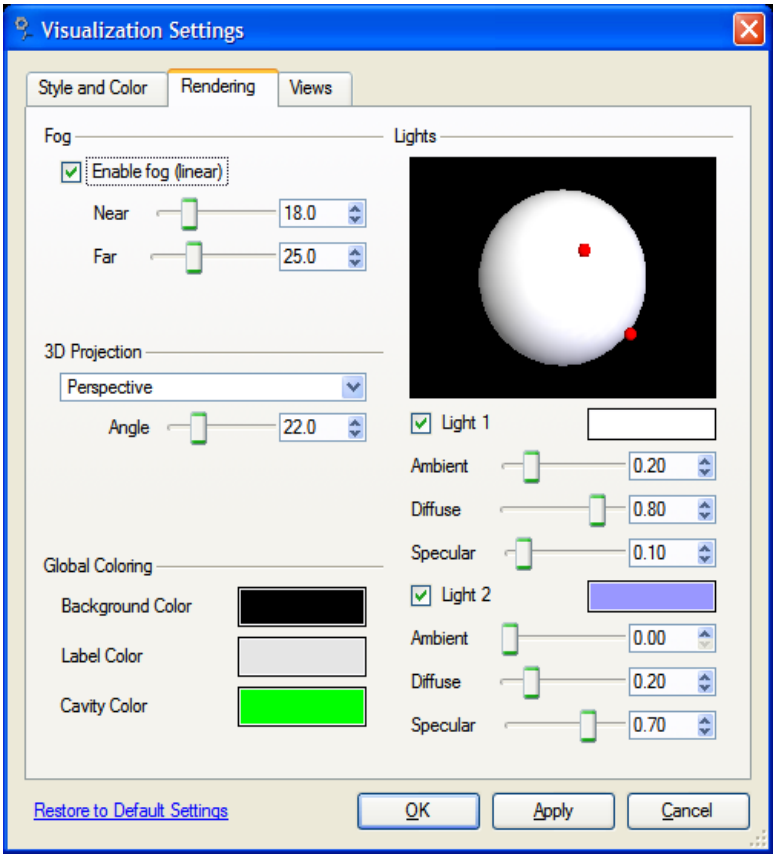


Figure 22: The Visualization Settings Rendering options.

The **Rendering** tab (Figure 22) on the **Visualization Settings** dialog allows you to customize the rendering behavior.

The **Fog** settings enables or disables fog. It is possible to adjust when the fog should begin (the **Near** value) and when the fog should reach its maximum density (the **Far** value).

The **3D Projection** settings manage the perspective projection. In **Perspective** projection objects farther away from the viewer appear smaller (the magnitude of this effect can be controlled by adjusting the field-of-view **Angle** parameter). In **Orthographic** projection object sizes are independent of their distance from the viewer.

The **Global Coloring** settings allow you to adjust the background color, the color labels are drawn with, and the color cavities (predicted binding pockets) are drawn with.

The **Lights** section controls the global lightning of the 3D world. It is possible to enable one or two light sources. Their positions can be adjusted directly in the 3D sphere view. The light source color can be changed by clicking the color selector next to the light checkbox.

OpenGL Lights contain three different parts: **Ambient** light always reaches an object, independent of its position relative to the light source. **Diffuse** lightning is dependent on whether the object faces the light source or faces away from it. The reflected light is emitted equally in all directions. **Specular** lightning is also dependent on the objects' orientation towards the light source, but the reflected light is emitted mainly in the direction of the reflected light ray (creating 'highlights').

Preset Views

The **Views** tab (Figure 23) in the **Visualization Settings** dialog controls the preset views (the macros residing under the **View** menu item on the main window menu bar).

The upper panel on the tab allows you to activate a preset view (by pressing the **Select** button') or delete a view (the **Delete** button). Notice that when deleting a view, you are not able to recover it.

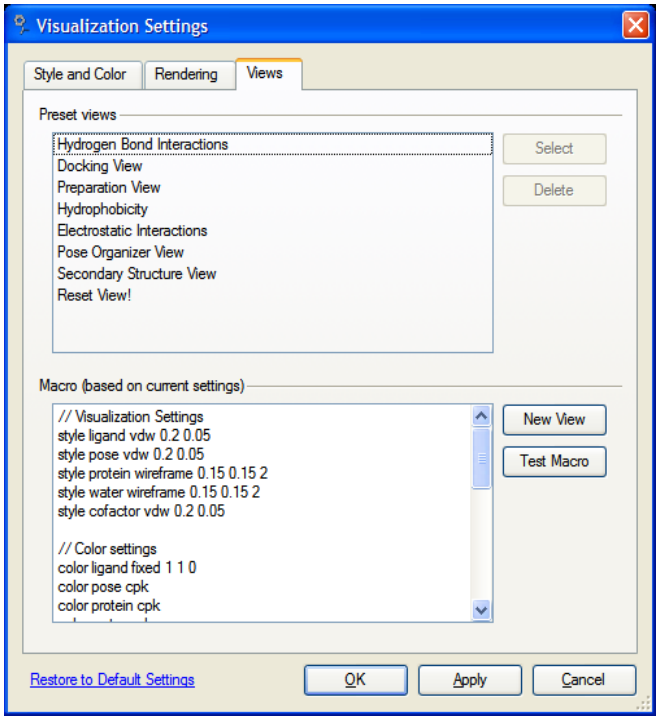


Figure 23: The Visualization Settings Views tab.

The lower panel allows you to create new views based on the current visualization settings. By pressing **New View** a dialog allows you to specify the name for the new view, after which it is added to the list of views on the main window menu bar. Views are stored as parts of the `viewermacros.xml` file and appear under the **View** menu item.

It is also possible to modify the macro in the text-area before committing it as a macro. Modified macros can be tested by pressing **Test Macro** before they are stored permanently.

2.19 Biomolecule Generator

Some PDB files contain transformation information for generating biomolecules. To apply these transformations, invoke the **Biomolecule Generator** by choosing **Tools** | **Biomolecule Generator**.

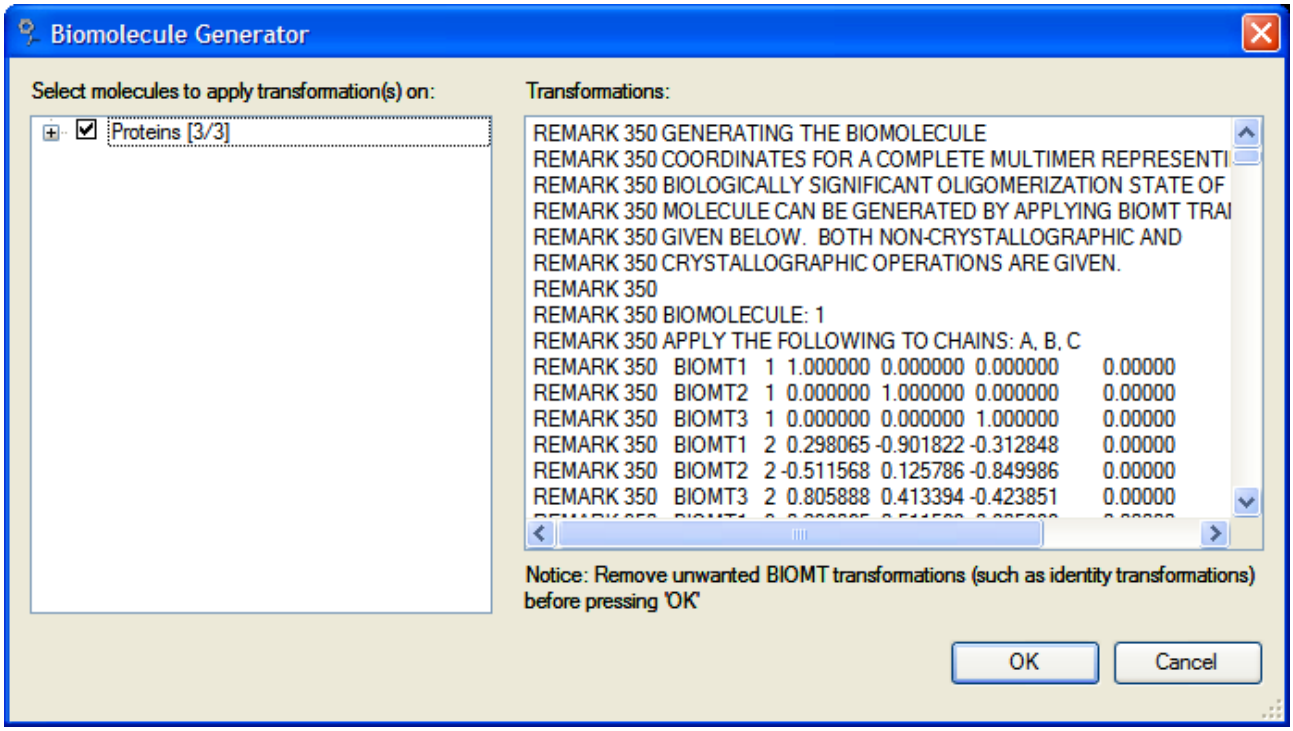


Figure 24: The Biomolecule Generator.

The left panel on the dialog controls which molecules the transformation should be applied to. This is normally the proteins (or protein chains), but ligands, water and cofactors can also be transformed.

The right panel contains a text box where a transformation description can be pasted. Notice that if a transformation remark was present in the last loaded PDB file it will automatically appear here.

It can be necessary to manually edit the transformation remarks. For instance the remarks may contain redundant identity transformations which should be removed:

```
// Example of identity transformation.

REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

PDB transformation remarks are triplets of remark lines, named BIOMT1-3. The first three columns constitute a rotation matrix, and the last column is a translation vector.

For some complex structures the transformation description may contain several steps where different transformations are applied to different subsets of the molecules. In this case it is necessary to run the **Biomolecule Generator** multiple times.

Also notice that biomolecules can be very large. Always render the protein in wireframe before attempting to generate large biomolecules.

2.20 Structural Alignment of Proteins

It is possible to structurally align proteins in MMV.

A structural alignment is done by matching a number of residues in two proteins and calculating the translation and rotation that minimizes the RMSD between the alpha-carbons in the matched residues.

The **Structural Protein Alignment** dialog can be invoked by selecting **Tools | Structural Protein Alignment** from the main menu.

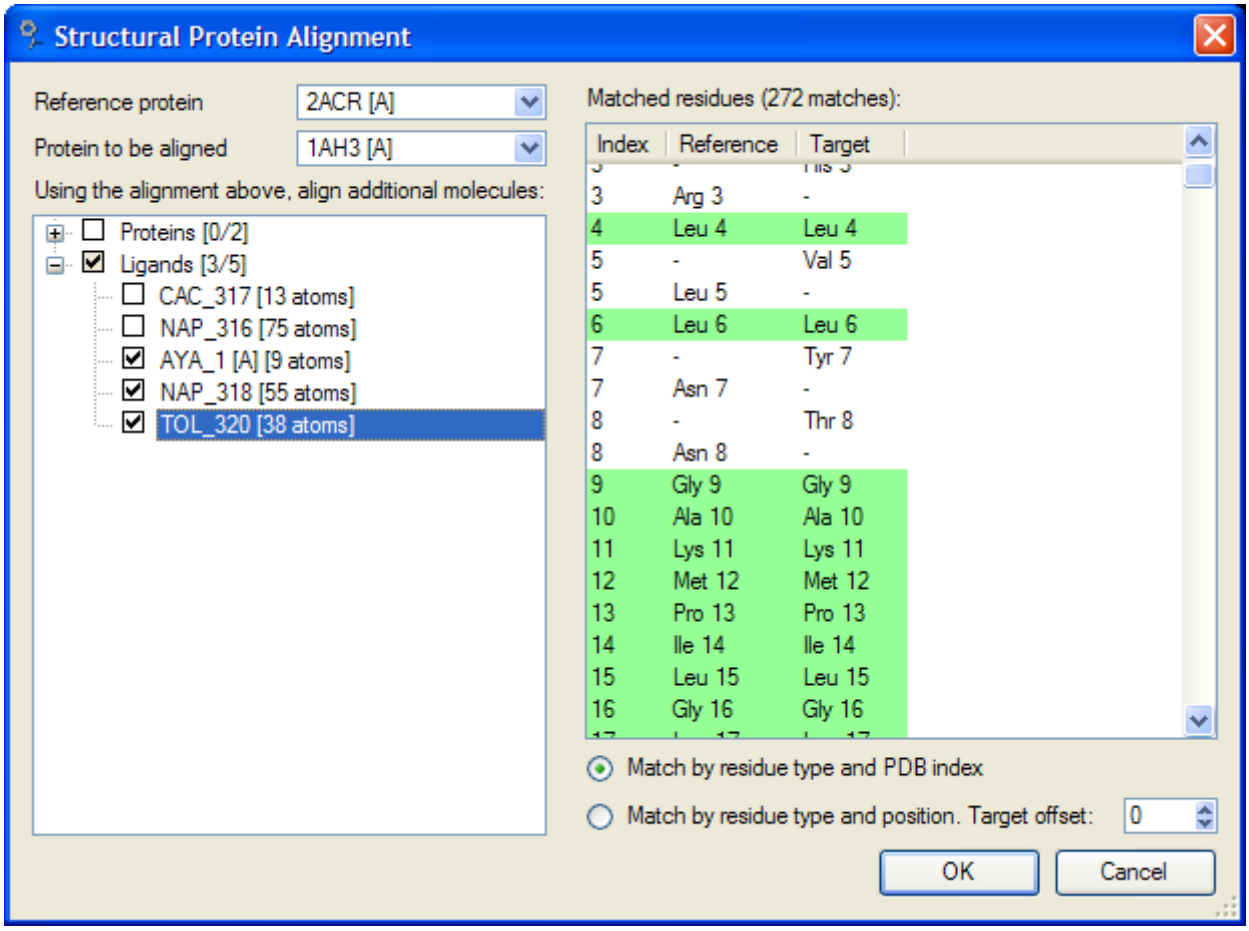


Figure 25: The Structural Protein Alignment dialog box.

The first step is to choose a reference protein and a protein to be aligned (the target protein). The target protein is the protein which will be translated and re-oriented.

When two proteins have been chosen, the list on the right side of the dialog will suggest a matching between residues in the proteins. Green entries indicate which residues that will be aligned. By default the matching will be done using **Match by residue type and PDB index** – where two residues will

be matched if they are of the same kind and have identical PDB residue identifiers.

Two PDB crystal-structures may have similar sequences, but different PDB residue identifiers. In this case it is possible to **Match by residue type and position.** This will match two residues if their positions in the sequences are identical. It is also possible to add a index offset to the target protein index.

Sometimes a number of other molecules are associated with a protein (a bound ligand or cofactor, or another protein chain). It is possible to select a number of additional molecules and apply the same transformation that aligns the target protein to the reference protein to the additional molecules. This is done by checking the desired molecules in the workspace view on the left side of the dialog. Notice that if the reference or target protein is selected as part of an additional alignment they will be ignored (since they are already considered).

2.21 Structural Alignment of Small Molecules

It is possible to perform a simple alignment of small molecules in MMV. By selecting three atoms in one ligand, and selecting three atoms in another ligand, a new context menu appears when clicking on an atom in one of the molecules - **Align...**. This will align the molecules. The atoms are aligned in the same order as they are selected, that is, the first selected atom in ligand 1 is aligned to the first selected atom in ligand 2 etc. Therefore, it is important to ensure that the selection order is correct and that no other atoms are selected. Alignments can be undone (click the undo button in the tool bar).

Notice: Only alignments with three selected atoms in each molecule are possible.

3 Preparation

3.1 Import of Molecules

Molecules can be imported into MMV using the **Import Molecule...** menu option located in the **File** menu. A shortcut is provided from the tool bar by clicking on the **File** folder icon or using the **Ctrl-O** keyboard shortcut. Molecules can also be imported by dragging-and-dropping the molecular file into the main application window.

Currently, MMV supports the following file formats:

- Protein Data Bank (pdb/ent)
- Sybyl Mol2 (mol2)
- MDL (sdf/sd/mol/mdl)

Notice that only PDB and Mol2 files can contain proteins and water molecules. In general, it is recommended to use Mol2 or SDF files for ligands since they can contain bonding information.

From the **Import Molecules** dialog shown in Figure 26, it is possible to select which molecules to import, prepare molecules, and inspect warnings found during parsing of the imported file.

Notice: If more than 10 ligands are present in the file (typically SDF or Mol2 files), a subset of the ligands can be selected for import using the **Specify ligand range** option (see Figure 26). Since it is computationally slow to display a large number of molecules (e.g. thousands of compounds), ligands and poses are not automatically shown in the **Visualization Window** if the number of molecules imported exceeds 50 (for each category).

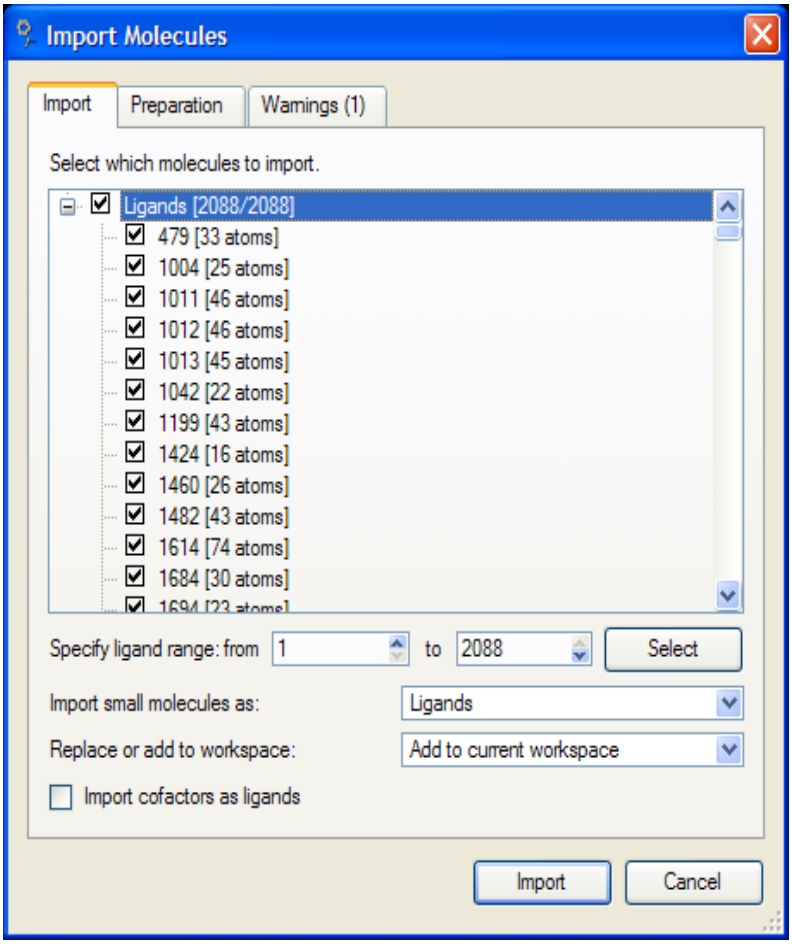


Figure 26: Import Molecules dialog.

When all relevant molecules have been imported, the molecules can be automatically prepared (see next section).

MMV automatically tries to identify cofactors: a molecule is considered a cofactor if it has less than 5 heavy atoms or its name is included in a list of common cofactor names (like 'HEM', 'SO4', 'PO4', ...). If this is not desired, it is possible to override cofactor recognition by checking the **Import cofactors as ligands** option.

3.2 Automatic Preparation

Some molecular file formats support information about bond type and charge (e.g. Mol2) while others do not (e.g. PDB). In order to maker proper predictions, it is important that the structures have been properly prepared. That is, that the atom connectivity is known and that the correct bond order and charges have been assigned.

The **Prepare Molecules** dialog allows the user to perform the necessary preparation. It is invoked automatically when importing Mol2, SDF, or PDB files, and can be invoked manually by selecting **Preparation** | **Prepare**

Molecules or by using the context menu (e.g. **Prepare Ligand...**) on molecules in the **Workspace Explorer**.

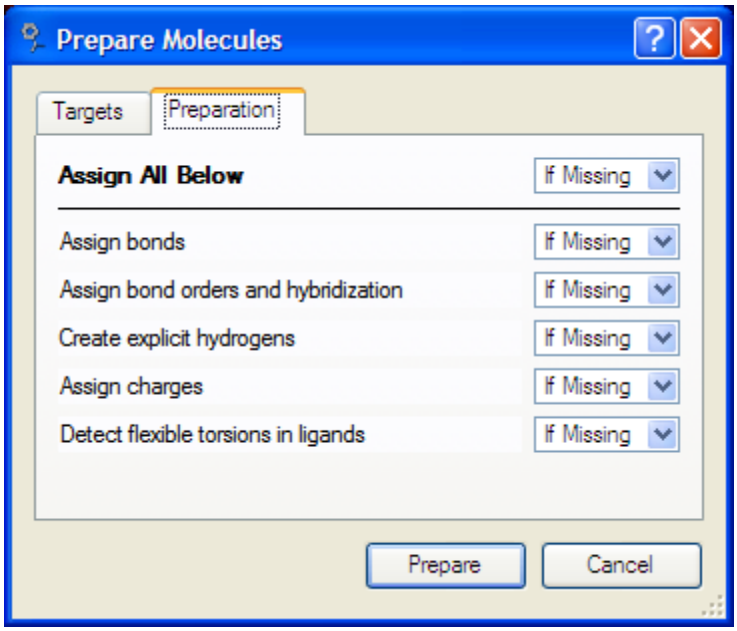


Figure 27: Preparing molecules.

Within all preparation types the following four different possibilities are available (see Figure 27):

- **Always**. Unconditionally performs the preparation by MMV.
- **Never**. Skips the preparation.
- **If Missing**. The preparation will only be performed if no knowledge is already present (e.g. if bond orders exist in the Mol2 file, bond orders are not assigned by MMV. However, if bond order information is not included, MMV will assign it).
- **Remove**. Tries to remove preparation (e.g. if 'Assign bond orders...' is set to 'remove', all bond orders will be set to single bonds. If 'Create explicit hydrogens' is set to 'remove' all hydrogen atom are removed).

Notice: The preparation options (Always, Never, If Missing, Remove) applies to each individual molecule (not each individual bond or atom). For instance, setting 'Assign bonds' to 'If Missing' results in covalent bonds being created for molecules not containing any bonds at all while molecules with bond information will preserve their bond assignments. Likewise, setting 'Create explicit hydrogens' to 'If Missing' will not add additional hydrogens to molecules containing e.g. polar hydrogens only. In this case, 'Always' should be used if all hydrogens should be created.

Assign Bonds

This option allows to determine which atoms are connected (covalently bound). Two atoms are connected if their distance is more than 0.4Å and less than the sum of their covalent radii plus a threshold of 0.45Å (the threshold is set to 0.485Å if one of the atoms is Phosphorus).

Assign Bond Order and Hybridization

This options allows recognition of bond orders (whether bonds are single, double or triple, ...), the number of hydrogens attached to the atoms, and their hybridization (SP, SP2, SP3). Also aromatic rings will be detected. It should be noted that this assignment is not always perfect - different protonation states can be difficult to assign properly. A detailed description can be found in Appendix II: Automatic Preparation.

Notice: The algorithm only assigns the number of implicit hydrogens to each atom. No actual atoms will be added. The next option **Create explicit hydrogens** allows you to add explicit hydrogens based on the implicit ones.

Create Explicit Hydrogens

Creates hydrogens matching the predicted number of hydrogens in the step above. The hydrogens are placed according to geometric criteria (i.e. SP3 hybridized atoms are kept at a 109 degrees geometry). The hydrogens are placed at standard distances according to the atom they are connected to. No energy minimization is performed.

Assign Charges

This option allows to assign partial charges to each atom based on the scheme described in Appendix III (Table 2).

Detect Flexible Torsions In Ligands

This option determines which bonds that should be considered flexible during docking. It is advisable always to set this option to either **If Missing** or **Always**. If this option is set to **Remove**, the ligand will be considered rigid during docking.

Hydrogen Bonding Type

Atom hydrogen bonding types (acceptor, donor, both or non-polar) are always set during preparation.

3.3 Manual Preparation

Molecules can be manually prepared using the context menus of highlighted atoms or bonds (see below).

Set Hybridization

Hybridization (SP, SP2, SP3) can be manually assigned to atoms by right-clicking on the atom in question and selecting the **Set Hybridization** menu option.

Set Hydrogen Bond Type

Hydrogen bond type (donor, acceptor, both, non-polar) can be manually assigned to atoms by right-clicking on the atom in question and selecting the **Set Hydrogen Bond Type** menu option.

Set Hydrogen Count

The **Set Hydrogen Count** menu option can be used to set the number of explicit hydrogens attached to the highlighted atom.

Assign Charges

Currently, the MVD scoring function (MolDock Score, see Appendix III for more details) uses partial charges assigned when running the **Preparation** dialog. However, the assignment of charges is based on standard templates and charge assignments can be missing in some cases. It is possible to manually assign partial charges to atoms by right-clicking on the atom in question and selecting the **Set Partial Charge** menu option.

Set Bond Order

Bond orders can be manually assigned by right-clicking on the bond in question and selecting the **Set Bond Order** menu option.

Notice that bonds are not visible in some visualization styles. The most suitable view is the ball-and-stick style, which can be set from the **Rendering** menu in the menu bar.

Set Ligand Flexibility

Flexible torsions in the ligand can manually be set rigid or flexible by right-clicking on a bond and selecting the **Set Flexibility** menu option.

4 Data Sources

There are several ways to import ligands and prepare them in Molegro Molecular Viewer.

- Ligands can be imported in the GUI (using **Import Molecules…** from the **File** menu) and included in the workspace. This is the easiest way to import data, but it can be slow if working with thousands of ligands.
- Ligands can be read from a *Data Source*. Ligands are 'streamed' from a source (such as a large file) and the selected molecules are imported. This can be particularly useful when importing a subset from e.g. an SDF file containing a large number of compounds since only the selected molecules are loaded into main memory.

Currently two types of data sources are available:

- *File data sources.* These are single files containing multiple structures (such as SDF or multi-molecule Mol2). It is possible to read a subset of the molecules contained in the file.
- *Multifile data sources*. These can be used when the input structures are split over several different files. A multifile data source may contain files with a mixture of different data formats.

4.1 Data Sources Syntax

File Data Sources

File data sources are identified by a 'File=' identifier. Examples:

```
File=\\fileserver\molecules\mol23.mol2
File="C:/Test Molecules/steroids.sdf";Index=2,4-8,12,34-
```

It is possible to import a subset of the structures in a file using the 'Index' specifier.

Molecules must be separated either by '$$$$' for SDF files or '@<TRIPOS>MOLECULE' for multi-molecule Mol2 files. Only one molecule will be extracted from each section separated by these separators. For PDB files only the first HETATM molecule will be imported.

Notices that all input structures are expected to be ligands. Molecules recognized as proteins or water molecules will be ignored.

The optional 'Index' specifier must be a comma-separated list of either single values or intervals. Notice that open intervals are allowed (e.g. '5-' or '-19'). Indices should be ordered strictly increasing. Invalid or non-existent indices will be ignored. The 'Index' specifier is 1-based (the number of of the first molecule is 1 and not 0).

Filenames containing spaces must be enclosed in quotation marks. It is possible to specify files on shared network drives and folders.

Multifile Data Sources

Multifile data sources are identified by a 'Dir=' identifier. Examples:

```
Dir="C:/Test Molecules";Pattern="*.sdf;*.mol2";Index=10-100
Dir=C:/Test;Pattern=Stereo*.sdf;Index=10-100
```

The Multifile data source takes a directory and scans it for the given pattern. Patterns are specified using '*' as a wildcard. Notice that on Linux and Mac operating systems, file patterns are case sensitive.

It is possible to specify more than one pattern by separating sub-patterns with semi-colons. Patterns with semi-colons must be surrounded by quotes.

As with file data sources it is possible to specify a subset using the molecule index specifier ('Index'). Notice, that the 'Index' specifier refers to the molecule index – not the file index.

4.2 Loading Data Sources Directly into the Workspace

By using the **File | Import From Datasource...** menu item it is possible to directly load a number of molecules into the workspace. This can be useful for importing a small subset of the molecules in a data source to check that the parsing and preparation is okay. Notice that all molecules are loaded into memory which can make the system slow to work with.

The **Data Source** dialog for defining the data source is shown in Figure 28.

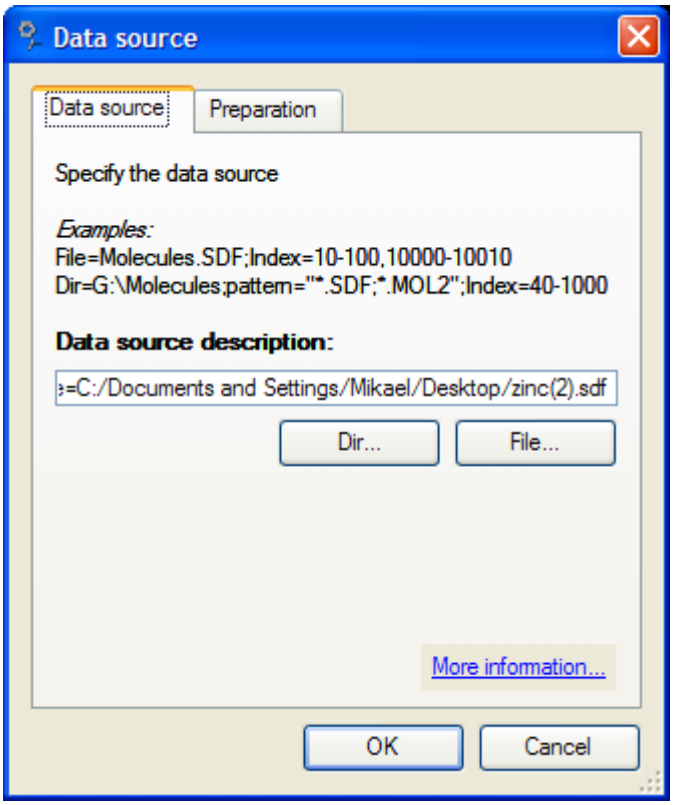


Figure 28: Specifying a data source.

Specify the data source on the **Data source description** line input or use either the **Dir...** or **File...** button to choose a directory or file from a dialog.

The **Preparation** tab determines how the data source should be prepared. These settings are described in Section 3.2.

5 Analyzing Docking Results

5.1 Pose Organizer

The **Pose Organizer** is used to inspect poses found by Molegro Virtual Docker (see Figure 29). It allows you to browse the list of current poses, to see detailed information about specific energy contributions, to visualize hydrogen bonds, electrostatic interactions, and to calculate ranking scores.

The **Pose Organizer** can be invoked in several ways. It is automatically displayed after a docking result file (with `mvdresults` file extension) has been imported to MMV by dragging-and dropping the file into MMV or using **File** | **Import Docking Results (*.mvdresults)...**.

Otherwise it can be invoked by using the context menu on the **Poses** category in the **Workspace Explorer** or by pressing the table icon in the **MMV Toolbar**.

When the **Pose Organizer** is invoked it displays a list of poses parsed from the `mvdresults` file (or poses currently in the workspace). The table in the middle of the dialog window shows various columns with information about different energy contributions and other data for each pose. The columns can be changed under the **Settings** tab pane. A panel in the bottom of the dialog (**Sorting Criteria**) allows the user to sort the table by up to three different criteria.

By default the table in the middle supports multiple selection, i.e. more than one pose can be highlighted. Only highlighted poses will be visible in the 3D window. This setting is useful for quick comparison of different poses.

This default behavior can be changed by selecting **Dynamic update (notice: disables multiple poses selection)**. In this mode only one pose is shown at a time. In return it offers the possibility to visualize different interactions for

the current selected pose (e.g. hydrogen bonds).

Even though **Dynamic Update** is a single-selection mode, it is possible to *lock* poses which keeps them visible even when not selected. A pose can be locked by using the context menu on its entry in the table and selecting **Lock** or **Unlock**. Locking is purely a visualization aid, and has no other consequences for the pose.

When inspecting poses obtained from different ligands, the **Only show top …** option can be used to focus on the most promising poses for each ligand. The selection of the top poses are based on the currently chosen **Sorting criteria**.

Pressing the **Open checked poses in Data Analyzer…** button makes it possible to further inspect poses using the Data Analyzer (introduced in Chapter).

Notice: A detailed energy analysis is available by right-clicking poses in the table and selecting **Energy Inspector…**. Additional options are available in the context menu allowing the user to select, remove, and export poses. These options are also available from the **File** and **Edit** menus located in the **Pose Organizer** dialog.

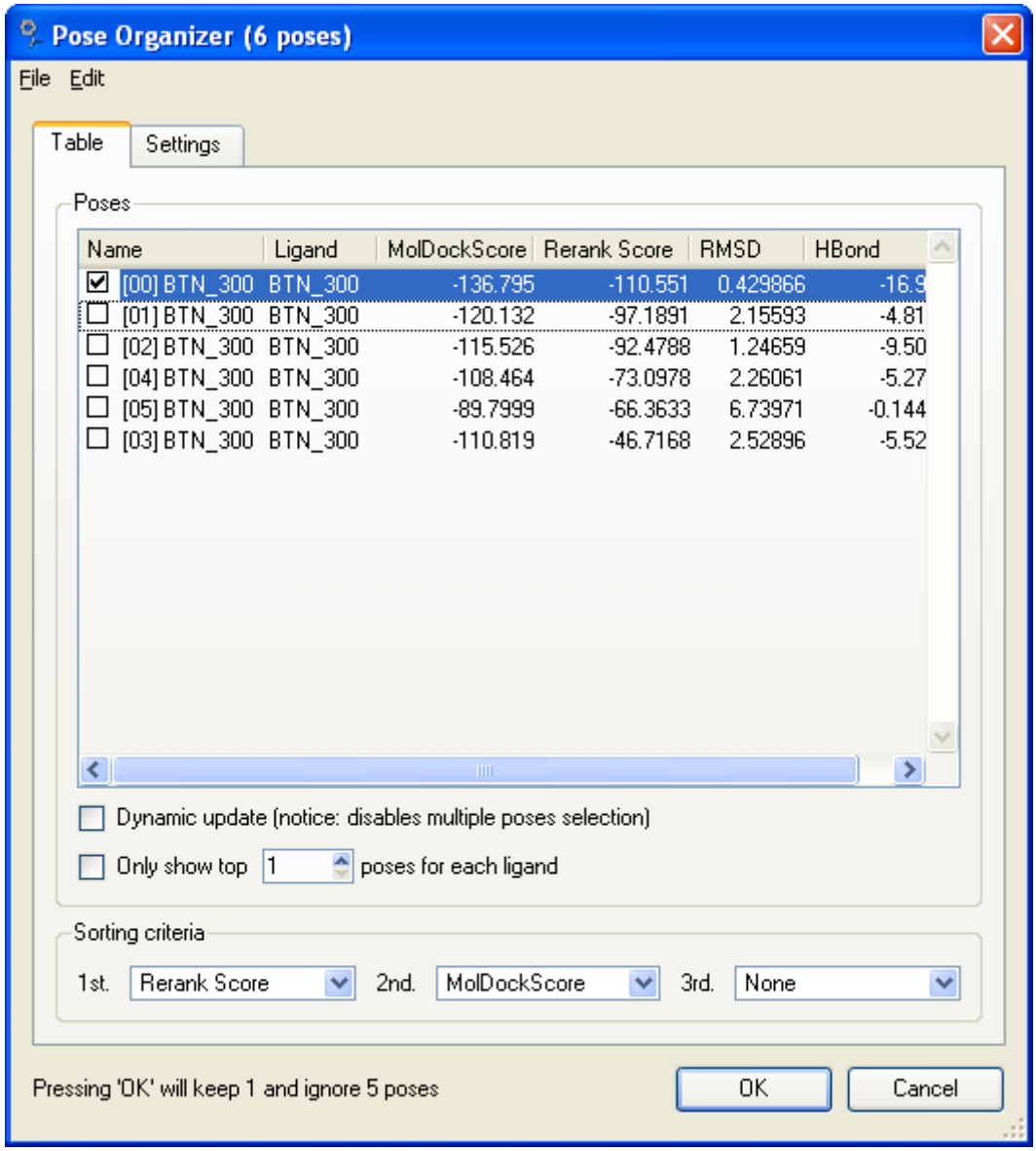


Figure 29: Pose Organizer dialog.

The **Settings** Tab Pane of the **Pose Organizer** can be used to customize the **Pose Organizer** (see Figure 30).

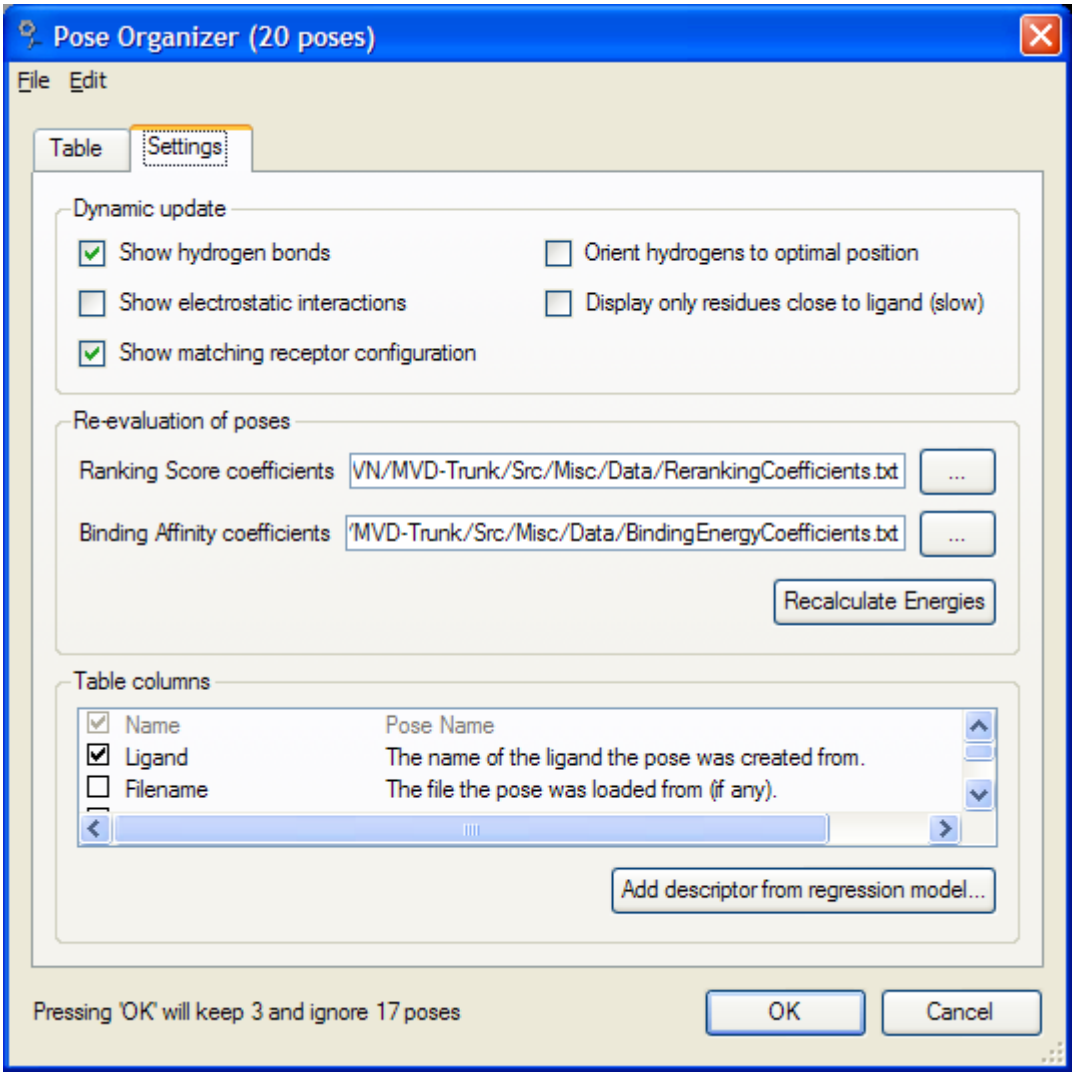


Figure 30: Pose Organizer settings.

The Dynamic Update Panel

The top panel (**Dynamic update**) chooses how the **Pose Organizer** behaves when single pose selection (**Dynamic update**) is enabled. It allows you to visualize hydrogen bonds, electrostatic interactions, orient hydrogens in the protein and ligand to their optimal position, and dynamically show residues close to the chosen pose. The **Orient hydrogens to optimal position** option is useful when inspecting poses as this makes it easier to see if the hydrogen bond is optimal.

Working with Receptor Conformations

When docking with sidechain flexibility in MVD a receptor conformation is saved together with each pose. When a new docking results file is imported,

MMV automatically checks whether any '.receptorConfiguration' files exist together with the poses.

If this is the case, the option **show matching receptor configuration** under **dynamic update** is enabled. When in dynamic update mode the pose organizer will now automatically change to the receptor conformation corresponding to the selected pose. If poses are imported into the workspace, their corresponding receptor conformations will automatically be added to the workspace.

The Re-Evaluation of Poses Panel

The middle panel allows for recalculation of the MolDock Score and re-ranking score terms. These scoring function values are already calculated if the poses are imported from a `mvdresults` file. Pressing the **Recalculate Energies** button will recalculate the energy terms (using the coefficients specified in the file for the re-ranking scores). Notice that the default evaluator settings will be used (e.g. internal ligand hydrogen bonds are not enabled).

The reranking score function is computationally more expensive than the scoring function used during the docking simulation but it is generally better than the docking score function at determining the best pose among several poses originating from the same ligand. The default reranking coefficients are listed in the file: `\Misc\Data\RerankingCoefficients.txt`

The Table Columns Panel

The bottom panel (**Table columns**) determines which columns (descriptors) that are shown in the table on the first tab. Table 1 describes the descriptors that are available.

New descriptors can be added from regression models created using the Molegro Data Modeller (MDM) software product (see www.molegro.com for more details about MDM). To add a new descriptor, simply press the **Add descriptor from regression model...** button and chose the regression model from a saved *Molegro Data Modeling* (MDM) file. Notice that the regression model should only be using the same descriptors as the ones that are available in the *DockingResults* files (only valid regression models will be available in the dialog).

The Pose Organiser shows a subset of the terms in the mvdresults file as columns in the Poses table. Some of the terms use the same terminology as in the mvdresults file (specifically Name, Ligand, Filename, Workspace, RerankScore, Torsions, RMSD, MW, LE1, LE3, Hbond, Similarity Score, Electro, Hbond and Heavy Atoms), but a few terms are renamed (in order to better fit the column layout and for clarity).

Column Name	Description
Name	The internal name of the pose (a concatenation of the pose id and ligand name)
Ligand	The name of the ligand the pose was created from
Workspace	The workspace (.mvdml file) containing the protein.
Filename	The file the pose is stored as (only available when inspecting docking results from a mvdresults file)
MolDockScore	The energy score used during docking (arbitrary units) [This is the 'Energy' term in a mvdresults file]
Rerank Score	The reranking score (arbitrary units)
RMSD	The RMS deviation from a reference ligand (if available)
Interaction	The total interaction energy between the pose and the target molecule(s) [This is the 'E-Inter total' term in a mvdresults file]
Cofactor	The interaction energy between the pose and the cofactors [This is the 'E-Inter (cofactor – ligand)' term in a mvdresults file]
Protein	The interaction energy between the pose and the protein [This is the 'E-Inter (protein - ligand)' term in a mvdresults file]
Water	The interaction energy between the pose and the water molecules [This is the 'E-Inter (water – ligand)' term in a mvdresults file]
Internal	The internal energy of the pose [This is the 'E-Intra (tors, ligand atoms)' term in a mvdresults file]
Torsions	The number of (chosen) rotatable bonds in the pose
Soft Constraints	The energy contributions from soft constraints [This is the 'E-Soft Constraint Penalty' term in a mvdresults file]
Electro	Short-range electrostatic protein-ligand interations (r<4.5Å)
ElectroLong	Long-range electrostatic protein-ligand interations (r>4.5Å)
HBond	Hydrogen bonding energy
Heavy Atoms	Number of heavy atoms in ligand
MW	Molecular weight (in dalton)
LE1	Ligand Efficiency 1: MolDock Score divided by Heavy Atoms count
LE3	Ligand Efficiency 3: Rerank Score divided by Heavy Atoms count
Docking Score	The score actually assigned to the pose during the docking. Notice that since score is calcalculated by the scoring function choosen in the Docking Wizard there may be small differences to the MolDock score reported in the 'MolDockScore' entry (for instance when using the grid based version of the MolDock score the grid interpolation may result in slighty different energies

Column Name	Description
	than the non-grid MolDock score version) [This is the 'PoseEnergy' term in a mvdresults file]
Similarity Score	The similarity score if docking with templates

Table 1: Column names available in the Pose Organizer dialog.

5.2 Saving Molecules and Solutions Found

Saving Workspace

After importing and preparing molecules, all information can be saved in a MVD Workspace (MVDML) file, which contains all relevant information (position of atoms, charges, hybridization, bond orders, ligand flexibility, ...). To save a workspace, select **File** | **Save Workspace As...**. Alternatively, use the keyboard shortcut **Ctrl-S**.

Notice: Visualization objects (surfaces, labels, interactions, ...) are not saved in MVDML files.

Exporting Molecules

The **Export Molecules** dialog can be used to export all (or a selection of) the molecules available in the workspace (see Figure 31).

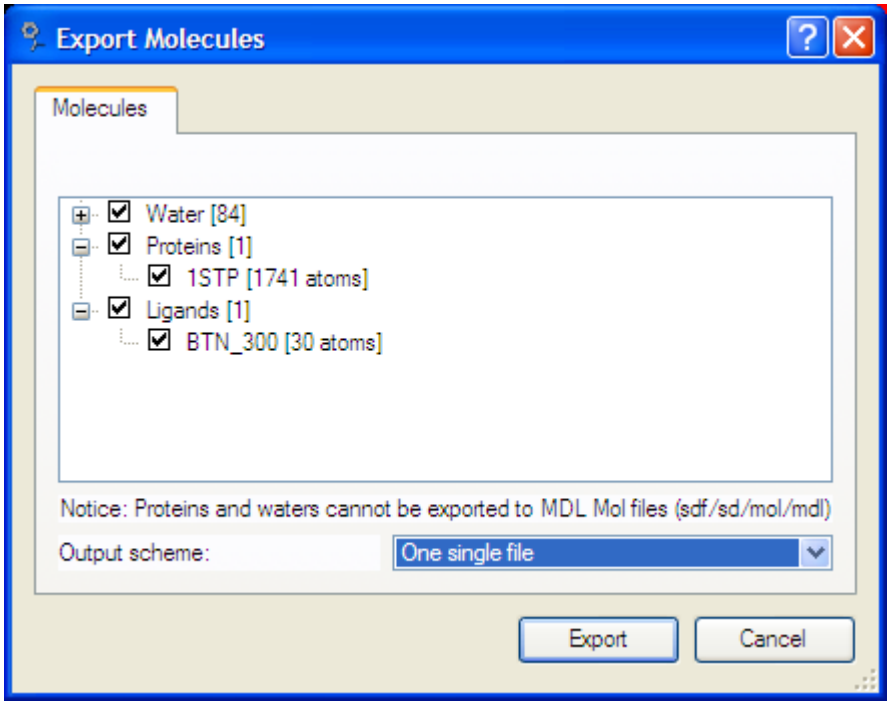


Figure 31: Export Molecules dialog: Select which molecules to export.

To export molecules, select **File** | **Export Molecules...** or **Export Molecules...** from the **Workspace** context menu in the **Workspace Explorer** (also available for proteins, ligands, cofactors, and poses).

Notice: Proteins and water molecules cannot be exported to SDF files.

Exporting Poses Found

To save the poses obtained from the docking runs, either use the **Export Molecules** dialog (described above) or save the poses from the **Pose Organizer** dialog.

5.3 Ligand Energy Inspector

The **Ligand Energy Inspector** allows you to get detailed information about the energy interactions for a given ligand or pose.

The **Ligand Energy Inspector** can be invoked in different ways. It can be started using the context menu in the **Workspace Explorer** by choosing **Open Energy Inspector** on any Ligand or Pose item. It can also be started from the **Pose Organizer** using the context menu on any pose entry or by selecting **Tools** | **Ligand Energy Inspector**.

Notice: the ligand energy inspector evaluates the energy of the ligand (or pose) when invoked. This means that the proteins, water molecules, and cofactors currently in the workspace are taken into account. If the workspace has been changed, the energy displayed here may not be the same as the one displayed in the Pose Organizer (since these were assigned by MVD during the docking evaluation).

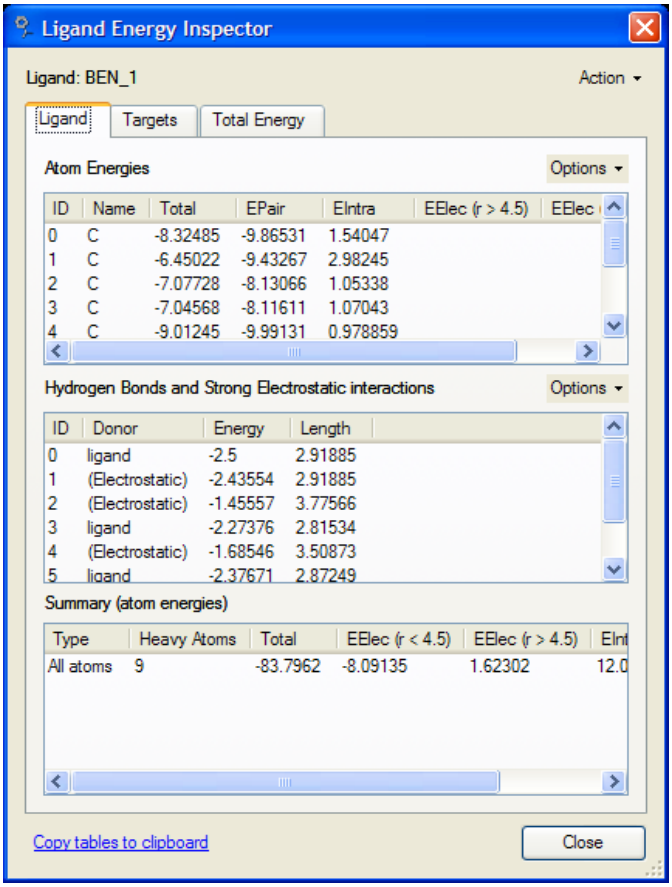


Figure 32: The Ligand Energy Inspector.

Besides inspecting the various energy contributions, it is possible to perform various actions, using the **Action** drop down menu:

- **Style Ligand Atoms by Energy**. This will scale the radius of the atoms proportionally to their energy contribution. Doing this makes it possible to get a visual overview of the important parts of the ligand.
- **Style Protein Atoms by Energy**. As above, this scales the protein atoms according to their energy contributions. Notice that protein atoms not interacting with the ligand are completely hidden. To make all protein atoms visible again, toggle the **Hide Residues** toolbar button.
- **Optimize Ligand and Protein Hydrogen Positions.** When docking with the Molegro Virtual Docker application the exact positions of the *rotatable* hydrogen atoms are not calculated. Instead it is assumed that the hydrogens are pointing in the optimal direction. In order to view the optimal direction of the rotatable hydrogens apply this option. Any rotatable hydrogens on the protein and ligand which are involved in hydrogen bonds will be oriented to the optimal direction.
- **Minimize Ligand.** This performs an energy minimization of the current molecule (with regard to its MolDock score energy).

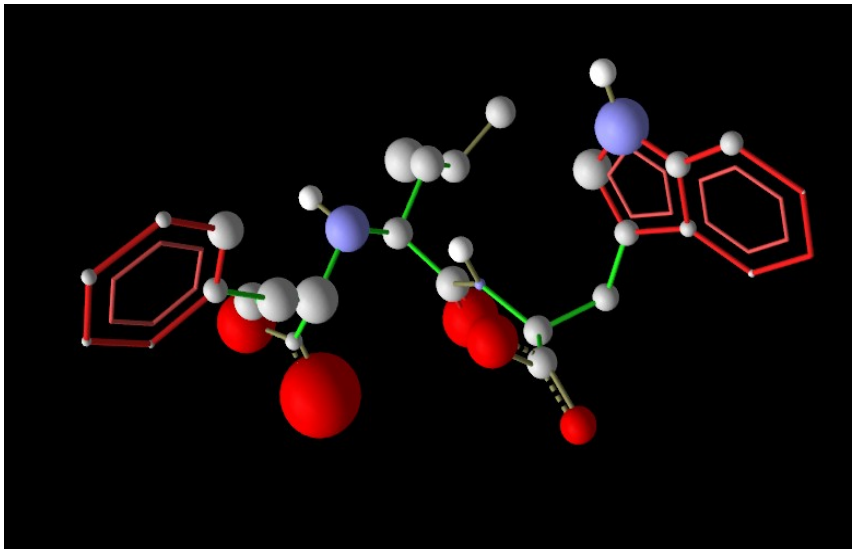


Figure 33: An example of the 'Style Ligand Atoms by Energy visualization', where atoms are scaled according to their energy contributions.

The Ligand Tab

The **Ligand** tab page consists of three tables.

The **Atom Energies** table shows information about individual atoms in the ligand. When hovering the mouse over an atom in the 3D view, it will automatically be highlighted in the table. Similarly when selecting entries in the table, atoms will be selected in the 3D GUI. It is possible to show or hide this table using the **Options** drop-down menu.

The following types of energy contributions may be listed for a ligand atom:

- **EPair**. This is the pairwise (PLP) steric and hydrogen bonding energy between a ligand atom and a receptor atom. Pairwise interactions between a ligand and either cofactors or water molecules will show up as 'EPair (cofactor)' and 'EPair (water)'.
- **EIntra**. This is the internal ligand energy between a ligand atom and the other atoms in the ligand.
- **EElec**. This is the pairwise electrostatic interactions. For the protein they are divided into long-range and short-range interactions ('EElec (R < 4.5 Å)' and 'EElec (R > 4.5 Å)').

The second table (**Hydrogen Bonds and Strong Electrostatic Interactions**) shows a list of all hydrogen bond and strong electrostatic interactions between the ligand and the target atoms. From the **Options** drop-down menu it is possible to show or hide the table, but it is also possible to toggle the table to display covalent bonds instead (**Show Covalent Bond Energies**). Finally the **Options** menu also makes it possible to toggle whether

hydrogen bonds and strong electrostatic interactions should be visualized in the GUI: Hydrogen bonds are visualized as dashed lines (where strong hydrogen bonds appear more solid) and strong electrostatic interactions are visualized as partial spheres oriented in the direction of the interaction. Green partial spheres correspond to favorable interactions, while yellow spheres correspond to non-favorable interactions.

The bottom panel (**Summary (atom energies)**) displays the sum of all atom interactions. (Notice that this is not the full energy of the ligand. Some interactions, like covalent bonding energies and constraint energies, are not included. For a complete list of energy contributions, see the **Total Energy** tab).

The Target tab

The **Target tab** displays a list of all targets atoms (atoms in proteins, cofactors, and water molecules in the workspace) involved in an interaction with the ligand. Atoms are only displayed in the list if the interaction energy is greater then 0.3 (in MolDock Score units). As with the Ligand Atom Energy table, selecting atoms in the table will select them in the 3D view and vice versa. The energy contributions are also divided into the same categories as in the Ligand Atom Table (for instance EElec and EPair).

The Total Energy Tab

The **Total Energy** tab displays a hierarchical breakdown of the various energy contributions.

The **Value** column displays the various terms which the MolDock Score and the RerankScore are based on.

The **MolDock Score** column shows how the MolDock score energy is composed. The MolDock score is a sum of a subset of the Value terms (all terms are given the same weight).

The **Rerank Score** uses a weighted combination of the terms used by the MolDock score mixed with a few addition terms (the Rerank Score includes the *Steric (by LJ12-6)* terms which are Lennard-Jones approximations to the steric energy – the MolDock score uses a piecewise linear potential to approximate the steric energy). The coefficients for the weighted Rerank Score are given in the **Rerank Weight** column, and the weighted terms and their summations are given in the **Rerank Score** column.

The relation between the terms showed in the Ligand Energy Inspector and the terms found in a mvdresults file is shown in the table below:

Ligand Energy Inspector Term	MVDResults Term
Total Energy	
External Ligand interaction	
Protein - Ligand interactions	
Steric (by PLP)	Steric
Steric (by LJ12-6)	VdW (LJ12-6)
Hydrogen bonds	HBond
Hydrogen bonds (no directionality)	NoHBond90
Electrostatic (short range)	Electro
Electrostatic (long range)	ElectroLong
Cofactor - Ligand	E-Inter (cofactor - ligand)
Steric (by PLP)	Not present in the mvdresults file, but can be calculated as: E-Inter (cofactor - ligand) - Cofactor (hbond) - Cofactor (elec)
Steric (by LJ12-6)	Cofactor (VdW)
Hydrogen bonds	Cofactor (hbond)
Electrostatic	Cofactor (elec)
Water - Ligand interactions	E-Inter (water - ligand)
Internal Ligand interactions	E-Intra (tors, ligand atoms)
Torsional strain	E-Intra (tors)
Torsional strain (sp2-sp2)	E-Intra (sp2-sp2)
Hydrogen bonds	E-Intra (hbond)
Steric (by PLP)	E-Intra (steric)
Steric (by LJ12-6)	E-Intra (vdw)
Electrostatic	E-Intra (elec)
Search Space Penalty	E-Penal
Soft Constraint Penalty	E-Soft Constraint Penalty

The Settings Tab

On the settings tab, the ligand evaluation can be customized. This can be important when inspecting poses from a docking run: Since the Ligand Energy Inspector is not aware of which scoring function settings were used during the docking, it is necessary to match the settings here to those selected in the Docking Wizard.

Internal ES toggles whether internal electrostatic interactions should be calculated for a pose, **Internal Hbond (no directionality)** toggles whether a pose should be allowed to have internal hydrogen bonds (notice that hydrogen bond directionality is not taken into account for internal hydrogen bonds in ligands), and **Sp2-Sp2 Torsions** determines whether an additional dihedral term should be added for taking Sp2-Sp2 bonds into account (see Appendix I: Docking Scoring Function).

5.4 RMSD Matrix

The **RMSD Matrix** dialog can be used to quickly inspect deviations between molecules in the workspace. In addition to the standard measure **Pairwise Atom-Atom RMSD (by ID)**, two variants **Pairwise Atom-Atom RMSD (checking all automorphisms)** and **Pairwise Atom-Atom RMSD (by nearest unmatched neighbour)** of the RMSD measure tries to take intrinsic symmetries of the molecule into account when calculating RMSD. The recommended choice is **Pairwise Atom-Atom RMSD (checking all automorphisms)**, which is also used by default.

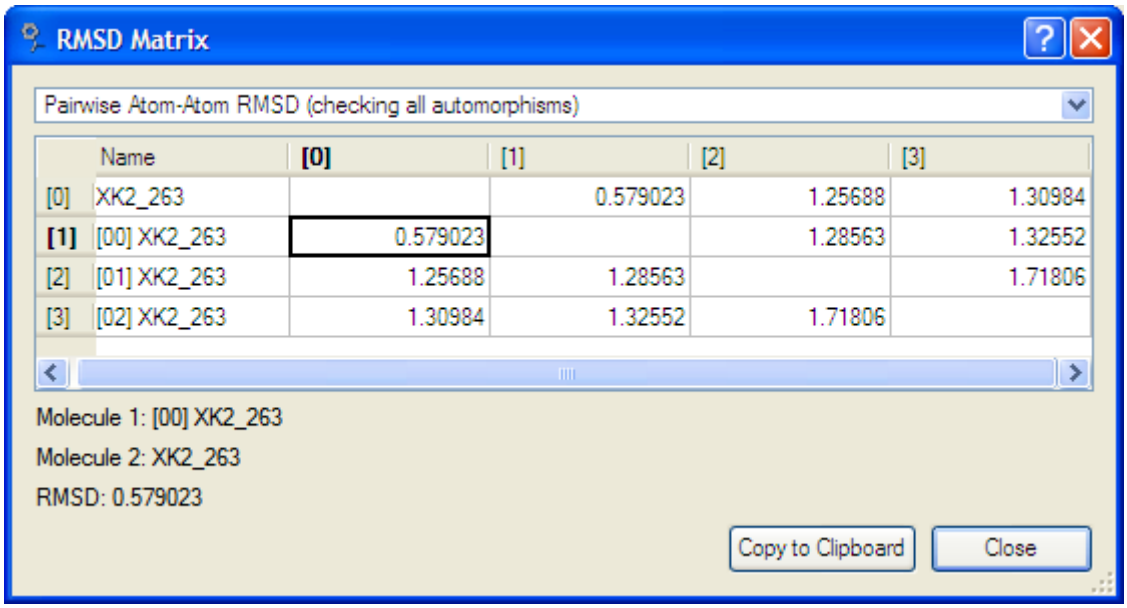


Figure 34: RMSD Matrix dialog.

The dialog can be invoked by choosing **RMSD Matrix** from the **Tools** menu. The **Copy to Clipboard button** can be used to copy the table to the clipboard for further inspection in an external text editor or spreadsheet.

6 Customizing Molegro Molecular Viewer

6.1 General Preferences

Molegro Molecular Viewer can be customized using the **Preferences** dialog, which can be invoked from the **Edit** menu or by pressing **F4**. Preference settings are categorized in **General**, **Graphics**, **Mouse**, and **Parsing** tabs.

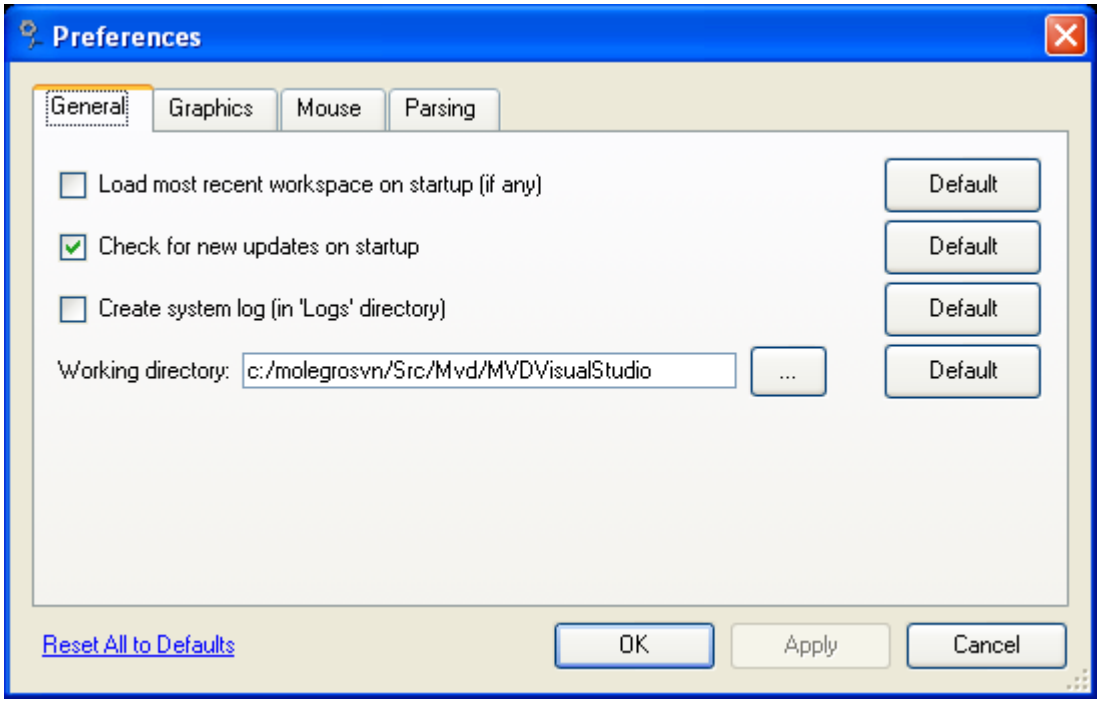


Figure 35: First tab of the preferences dialog.

In the **General** tab (see Figure 35), the following settings are available:

- The **Load most recent workspace on startup (if any)** option toggles automatic import of the last used workspace.
- The **Check for new updates on startup** option enables MMV to automatically check for new updates during startup.
- The **Create system log (in 'Logs' directory)** option is used to toggle whether a system log is created for each execution of MMV. The system log contains information about user actions conducted and is used to track potential bugs and performance problems. The log files are stored in the `Logs` directory located in the same directory as the `mmv` executable file. *Notice:* If you encounter problems with MMV please email the log file created before the crash to: bugs@molegro.com
- The **Working directory** setting is used to set the current **Working directory**, which is the root path for file related operators (e.g. when loading and saving molecular structure files and log files).

The **Graphics** tab (see Figure 36) contains settings related to the **Visualization Window**:

- The **Show pivot point (rotational center)** option toggles the visibility of the pivot point (small grayish ball).
- The **Show root atom** option toggles the visibility of the currently chosen root atom for each of the ligands in the workspace. The root atom is used as root in the torsion tree, which is used to construct the ligand conformation during the MVD docking simulation.
- The **Fade 3D labels when in background** option toggles fading of labels in the **Visualization Window.**
- The overall rendering quality can be specified using the **Quality** option. Modern computers with dedicated 3D hardware should be able to run at highest quality even when rendering relatively large molecules. It is easy to test new quality settings by selecting the level of quality and pressing the **Apply** button.

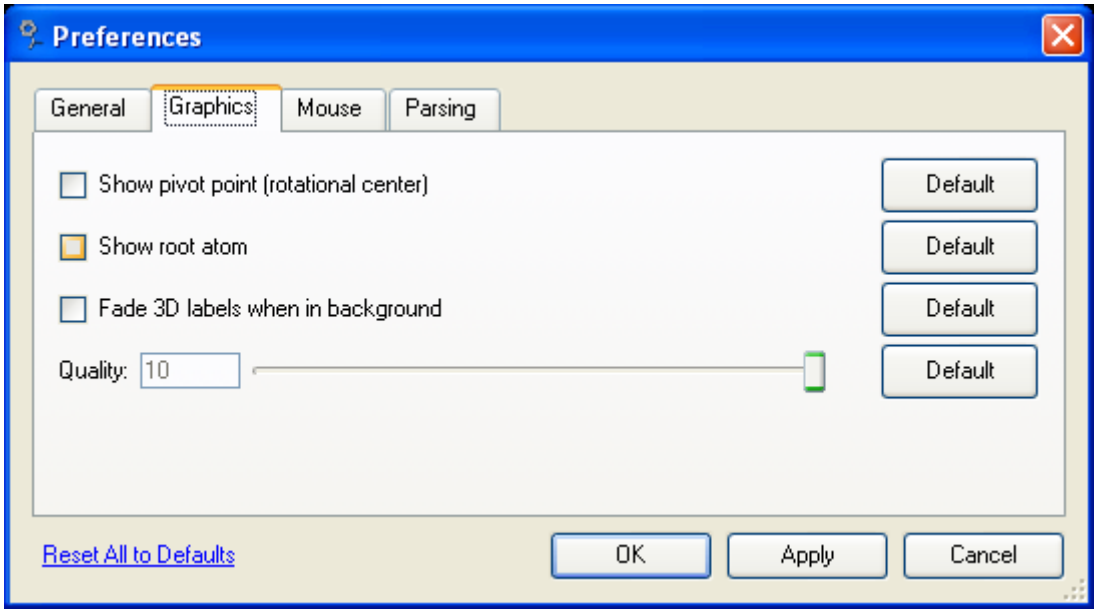


Figure 36: The graphics tab of the Preferences dialog.

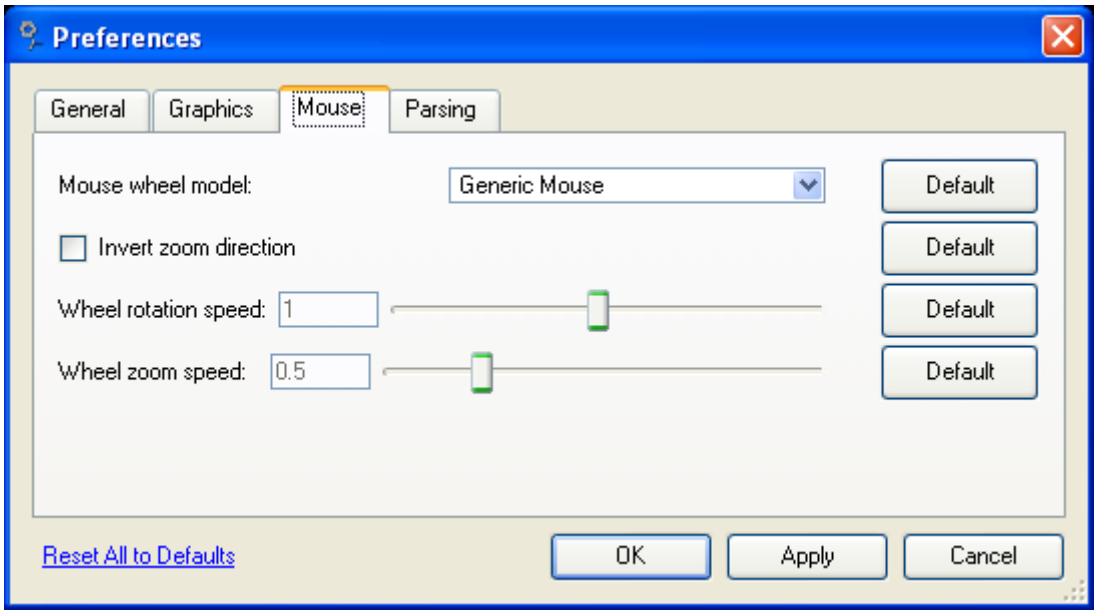


Figure 37: Mouse Preferences.

The **Mouse** tab customizes how the mouse interacts with the 3D world. MMV supports the 360 degrees scroll-ball on the *Apple Mighty Mouse*. Currently, the 360 degrees scroll-bar feature is only supported on Mac OS X (since no mouse drivers are available for other platforms), but the mouse still works as a generic mouse on Windows and Linux.

To enable *Apple Mighty Mouse* support select it under **Mouse wheel model**. When **Apple Mighty Mouse** mode is selected, the scroll-ball can be used to rotate the 3D world. Additionally, the scroll-ball button can be used to zoom in

the 3D world by pressing the button while using the scroll-ball as a standard mouse-wheel. However, to enable the zoom option, the scroll-ball button should be set to **Button 3** in the *Mac OS X Mouse preferences* dialog (see Figure 38).

Invert zoom direction toggles how the the 3D worlds zooms – rotating the scroll wheel towards the user will normally make the 3D objects appear larger, but this behavior can be inverted by toggling this option on. The setting also applies to zooming using both mouse buttons.

It is also possible to adjust the mouse wheel sensitivity (by using the **Wheel rotation speed** and **Wheel zoom speed** sliders).

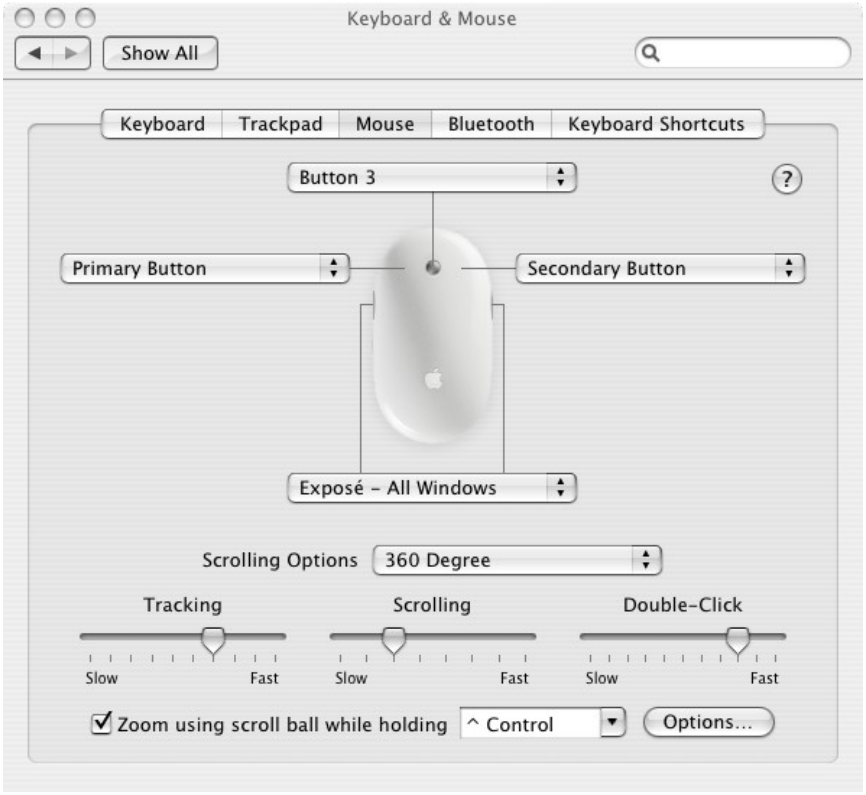


Figure 38: Mighty Mouse preferences on Mac OS X.

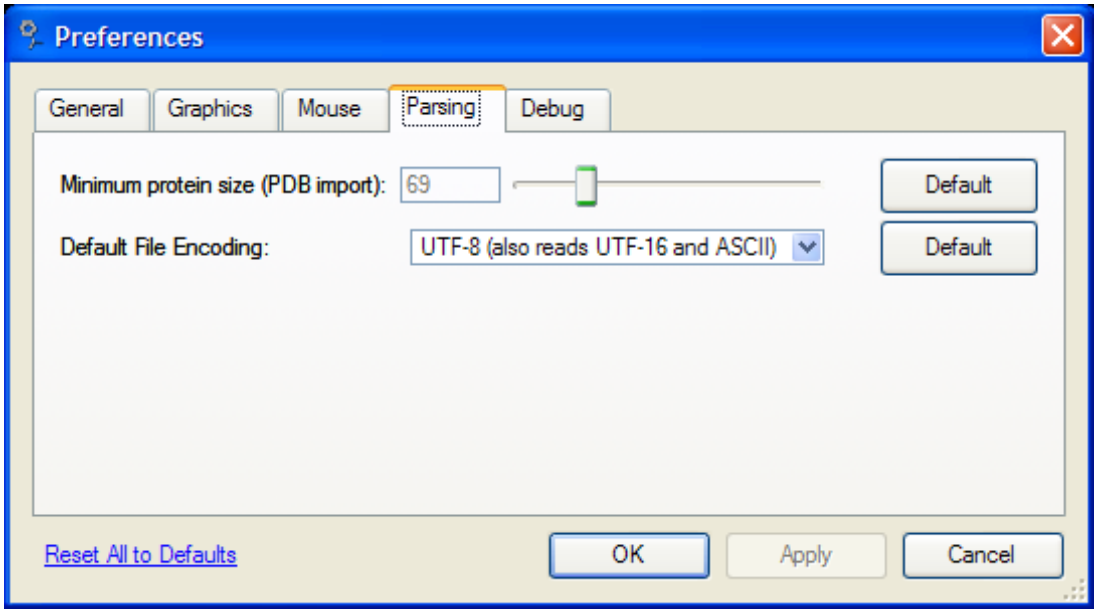


Figure 39: Parsing preferences.

The final settings tab, **Parsing**, contains the **Minimum protein size (PDB import)** option. This option is used for setting the minimum number of heavy atoms required for parsing a molecule as a protein during PDB import (default is 69 heavy atoms). If the parsed molecule contains less heavy atoms than the specified threshold value it is parsed as a ligand (and residue information is ignored).

The **Parsing** tab also determines how MMV handles non-standard characters (such as special national characters). This setting is used when importing and exporting molecular structures in text file format (such as SDF,Mol2,PDB files) and when working with other text files (e.g. 'mvdresults' files). XML files (such as the MVDML file format) are always stored as UTF-8.

The **Default File Encoding** drop-down box allows you to choose which encoding should be used. It is recommended to use the default setting, UTF-8 Unicode. Using the UTF-8 encoding all Unicode characters can be encoded and since molecular data files rarely contain special characters, it is more space-efficient than UTF-16 (where each character always uses at least 2 bytes). Files stored as 8-bit ANSI/ASCII files will also be imported correctly as Unicode if they do not contain any special national characters, and UTF-16 will also be automatically recognized in this mode. It is also possible to store data as Locale 8-bit. In this encoding all characters are stored as a single byte, meaning only 256 characters can be represented. The actual characters included in this set depends on the current national codepage settings on the machine. This option should only be used when exporting data to older software products not capable of parsing Unicode text.

The preference settings are stored when exiting the MMV application. The location of the saved settings depends on the operating system used:

- Windows: the settings are stored in the `system registry`.
- Mac OS X: the settings are stored in a `com.molegro.MMV.plist` file located in the `<user folder>/Library/Preferences/` folder.
- Linux: the settings are stored in a `mmvrc` file located in a hidden folder named `<user folder>/.molegro`.

6.2 Command Line Parameters

Currently, the following command line parameters are available:

```
<filename>
-currentPath
```

The `<filename>` parameter can be used to import molecular files during MMV startup. If more than one file is listed (separated by spaces), each file will be imported.

Example: `/Molegro/MMV/bin/mmv 1stp.pdb`

The `-currentPath` parameter can be used to override the working directory specified in the general preference settings with the current path. This is particularly useful when running MMV from different working directories (using a terminal window) or when using a script to start up MMV.

Example: `/Molegro/MMV/bin/mmv -currentPath`

7 Appendix I: Supported File Formats

MMV accepts the following molecular structure formats:

- PDB (Protein Data Bank). Supported file extensions: *pdb/ent*.
- Mol2 (Sybyl Mol2 format). Supported file extensions: *mol2*.
- SDF (MDL format). Supported file extensions: *sdf/sd* (for multiple structures) and *mol/mdl* (for a single molecular structure).

Currently, the following information is ignored during import of molecular structures:

- Lone pairs and dummy atoms (all file formats).
- When alternative atoms are reported, only the first alternative is used. The remainder is ignored (all file formats). If one of the other alternatives should be used, change the order of occurrence in the the file before import.
- CONNECT records (PDB format).
- SUBSTRUCTURE records are ignored during import but created when structures are exported (Mol2 format).

Notice: Although extensive testing and validation of the import and export of these file formats have been conducted, parsing errors may occur. Compliance with the file format standards/protocols will reduce parsing problems significantly. The import/export routines used have been extended to handle deviations from the file format protocols, but parsing errors may still occur. Found parsing errors can be reported (send email to bugs@molegro.com).

Additionally, Molegro Molecular Viewer and Molegro Virtual Docker uses their own MVDML file format. MVDML is a shorthand notation for *Molegro Virtual Docker Markup Language* and is an XML-based file format. In general, MVDML

can be used to store the following information:

- Molecular structures (atom coordinates, atom types, partial charges, bond orders, hybridization states, ...)
- Constraints (location, type, and constraint parameters)
- Search space (center and radius)
- State information (workspace properties, ...)
- Cavities (location, cavity grid points)

Notice: Purely graphical objects (e.g. labels, interactions, annotations, backbones, and surfaces) are not saved.

8 Appendix II: Automatic Preparation

The principles behind automatic preparation in MMV are described below.

Aromaticity

- All rings (closed loops) are identified.
- These rings are 'weeded out', until a 'smallest subset' (capable of covering all ring bonds) remains.
- These rings are considered aromatic if:
  1) For 5-cycles: the mean torsion angle is less then 9.5°
  2) For 6-cycles: the mean torsion angle is less then 12°
- If the aromatic ring contains an atom which has out-of-plane bonds, it is degraded to be non-aromatic.

Notice that this is only a geometrical check for aromacity. It does not include more advanced checks such as Hückel's rule, and may fail on overlapping ring systems.

Assign Hybridization

- All atoms with average bond angles > 155°, are marked as SP1
- All atoms with average bond angles > 115°, are marked as SP2
- All remaining atoms are marked SP3.
- All atoms part of aromatic rings are marked as SP2.
- Ensure that if an atom is SP2 or SP, it must be connected to another SP or SP2 or a terminal atom. Otherwise the atom is degraded (i.e. SP2 -> SP3)

- Lastly the geometry surrounding a SP2 atom should be planar, otherwise it is degraded to SP3.

Bond Order

- All atom bonds are set to 'unknown'. All implicit hydrogens are set to '-1'.
- All bonds to SP3 atoms are set to 'single' order.
- Next, a template file containing standard chemical motifs (-POO-, C(NH2)(NH2), ...) is processed. The templates are located in the file: `\misc\data\preparationTemplates.xml`
- All unset SP2-SP2 bonds involved in a planar geometry (less than 10 degrees) are set to 'double'.
- Next all SP2 atoms are checked to see if a double bond to a neighbour atom is possible. If several atom bonds are possible, the atom with highest electro negativity is chosen. If this still results in several possibilities, the atom closest to the current one will be chosen.

9 Appendix III: MolDock Score

The MolDock scoring function (MolDock Score) used by MVD [THOMSEN 2006] is derived from the PLP scoring functions originally proposed by Gehlhaar et al. [GEHLHAAR 1995,1998] and later extended by Yang et al. [YANG 2004]. The MolDock scoring function further improves these scoring functions with a new hydrogen bonding term and new charge schemes. The docking scoring function, E_{score}, is defined by the following energy terms:

$$E_{score} = E_{inter} + E_{intra}$$

where E_{inter} is the ligand-protein interaction energy:

$$E_{inter} = \sum_{i \in ligand} \sum_{j \in protein} \left[E_{PLP}(r_{ij}) + 332.0 \frac{q_i q_j}{4 r_{ij}^2} \right]$$

The summation runs over all heavy atoms in the ligand and all heavy atoms in the protein including any cofactor atoms and water molecule atoms that might be present. The E_{PLP} term is a piecewise linear potential described below. The second term describes the electrostatic interactions between charged atoms. It is a Coulomb potential with a distance-dependent dielectric constant given by: $D(r) = 4r$. The numerical value of 332.0 fixes the units of the electrostatic energy to kilocalories per mole. To ensure that no energy contribution can be higher than the clash penalty the electrostatic energy is cut-off at the level corresponding to a distance of 2.0 Å for distances less than 2.0 Å. Notice that although the electrostatic energy contribution has the theoretically predicted

magnitude, the other energy terms are empirically motivated and the total energy does not necessarily correlate with the true binding affinity. The charges are set according to the scheme listed in Table 2. Metal ions are assigned a charge of +1 (e.g. *Na*) or +2 (e.g. *Zn*, *Ca*, *Fe*).

charge	ligand atoms	protein atoms
0.5	N atoms in $-C(NH_2)_2$	His (ND1/NE2)
		Arg (NH1/NH2)
1.0	N atoms in $-N(CH_3)_2$, $-(NH_3)$	Lys (N)
-0.5	O atoms in –COO, $-SO_4$,	Asp (OD1/OD2)
	$-PO_2$, $-PO_2-$	Glu (OE1/OE2)
-0.66	O atoms in $-PO_3$	
-0.33	O atoms in $-SO_3$	
-1.0	N atoms in $-SO_2NH$	

Table 2: Charge templates.

E_{PLP} is a "piecewise linear potential" using two different sets of parameters: One set for approximating the steric (Van der Waals) term between atoms, and another stronger potential for hydrogen bonds. The linear potential is defined by the following functional form:

$$E_{PLP}(0) = A_0,\ E_{PLP}(R_1) = 0,\ E_{PLP}(R_2) = E_{PLP}(R_3) = A_1,\ E_{PLP}(r) = 0 \text{ for } r \geq R_4$$

and is linearly interpolated between these values. The parameters used here (see Table 3) were adopted from GEMDOCK [YANG 2004].

	A_0	A_1	R_1	R_2	R_3	R_4
hydrogen bond	20.0	-2.5	2.3	2.6	3.1	3.6
steric	20.0	-0.4	3.3	3.6	4.5	6.0

Table 3: PLP parameters.

A bond is considered a hydrogen bond if one of the atoms can *donate* a hydrogen atom and the other atom can *accept* it. The atom types are assigned according to the scheme shown in Table 4.

type	atoms
acceptor	N and O (with no Hs attached)
donor	N and S (with one or more Hs attached)
both	O (with one H attached) or O in water molecule
nonpolar	all other atoms

Table 4: Hydrogen bond types.

The PLP hydrogen bond term mentioned above only depends on the distance between atoms. In order to take into account the directionality of the hydrogen bonding, the geometry of the hydrogen bond is examined and the following factor H_{factor} is multiplied to the PLP hydrogen bond strength:

$$H_{factor} = \Phi(\angle_{D\text{-}H\text{-}A};90°;150°)\cdot\Phi\ (\angle_{H\text{-}A\text{-}AA};90°;100°)\cdot\ \Phi\ (\angle_{D\text{-}A\text{-}AA};90°;100°)$$

Here *AA* (Acceptor Antecedent) denotes a heavy atom connected to the acceptor (A), *D* denotes the donor and *H* is the donated hydrogen atom. The ramp function Φ is defined as $\Phi(A;A_{min};A_{max}) = 0$ for $A \le A_{min}$ and $\Phi(A;A_{min};A_{max}) = 1$ for $A \ge A_{max}$ and is linearly interpolated between these values for $A_{min} < A < A_{max}$. If it is not possible to calculate one of these factors it is omitted. This is for example the case for hydroxyl rotors where the exact location of the hydrogen is not investigated during docking, and the two first factors cannot be calculated. The angle checks above were motivated by the approach taken by McDonald and Thornton [MCDONALD 1994].

E_{intra} is the internal energy of the ligand:

$$E_{intra} = \sum_{i \in ligand} \sum_{j \in ligand} E_{PLP}(r_{ij}) + \sum_{flexible\ bonds} A[1 - \cos(m \cdot \theta - \theta_0)] + E_{clash}$$

The double summation is between all atom pairs in the ligand excluding atom pairs which are connected by two bonds or less. The second term is a torsional energy term, parameterized according to the hybridization types of the bonded atoms (see Table 5). θ is the torsional angle of the bond. Notice that this angle is not necessarily uniquely determined. The average of the torsional energy bond contribution was used if several torsions could be determined. The last

term, E_{clash}, assigns a penalty of 1000 if the distance between two heavy atoms (more than two bonds apart) is less than 2.0 Å. Thus, E_{clash} term punishes infeasible ligand conformations.

	θ_0	m	A
sp^2-sp^3	0.0	6	1.5
sp^3-sp^3	π	3	3.0
sp^2-sp^2	0.0	2	3.0

Table 5: Torsional parameters.

(the sp^2-sp^2 term is not enabled by default)*

Terms in the '.mvdresults' file

After MVD has predicted one or more promising poses using the MolDock score, it calculates several additional energy terms. All of these terms are stored in the 'DockingResults.mvdresults' file at the end of the docking run.

The 'rerank score' is a linear combination of these terms, weighted by the coefficients given in the 'RerankingCoefficients.txt'.

A '.mvdresults' file is not meant to be interpreted or inspected manually. Instead it should be opened in MMV or MVD (either by dragging it onto the workspace or by selecting 'File | Import Docking Results (*.mvdresults)...'.

The following table explains the different terms in a '.mvdresults' file:

Textual Information	
Ligand	The name of the ligand the pose was created from.
Name	The internal name of the pose (a concatenation of the pose id and ligand name).
Filename	The file containing the pose.
Workspace	The workspace (.mvdml-file) containing the protein. (Notice: This entry appears in the header of the mvdresults file)
Run	When running multiple docking runs for each ligand, this field contains the docking run number.
Energy terms (total)	
Energy	The MolDock score (arbitrary units). Notice that this value is always calculated using the non-optimized MolDock score (and hence may differ from the PoseEnergy below which may use interpolation on precalculated grids).
RerankScore	The reranking score (arbitrary units).

PoseEnergy	The score actually assigned to the pose during the docking. Notice that since the score is calculated by the scoring function chosen in the Docking Wizard, there may be small differences to the MolDock score reported in the 'Energy' entry (for instance when using the grid-based version of the MolDock score the grid interpolation may result in slighty different energies as compared to the non-grid MolDock score version)
SimilarityScore	Similarity Score (if docking templates are enabled).
LE1	Ligand Efficiency 1: MolDock Score divided by Heavy Atoms count.
LE3	Ligand Efficiency 3: Rerank Score divided by Heavy Atoms count.
Energy terms (contributions)	
E-Total	The total MolDock Score energy is the sum of internal ligand energies, protein interaction energies and soft penalties.
E-Inter total	The total MolDock Score interaction energy between the pose and the target molecule(s).
E-Inter (cofactor - ligand)	The total MolDock Score interaction energy between the pose and the cofactors. (The sum of the steric interaction energies calculated by PLP, and the electric and hydrogen bonding terms below)
Cofactor (VdW)	The steric interaction energy between the pose and the cofactors calculated using a LJ12-6 approximation. *Notice: This term is not used by the MolDock score*
Cofactor (elec)	The electrostatic interaction energy between the pose and the cofactors.
Cofactor (hbond)	The hydrogen bonding interaction energy between the pose and the cofactors (calculated by PLP).
E-Inter (protein - ligand)	The MolDock Score interaction energy between the pose and the protein. (Equal to Steric+HBond+Electro+ElectroLong below)
Steric	Steric interaction energy between the protein and the ligand (calculated by PLP).
HBond	Hydrogen bonding energy between protein and ligand (calculated by PLP).
Electro	The short-range (r<4.5Å) electrostatic protein-ligand interaction energy.
ElectroLong	The long-range (r>4.5Å) electrostatic protein-ligand interaction energy.
NoHBond90	This is the hydrogen bonding energy (protein-ligand) as calculated if the directionality of the hbond was not taken into account. *Notice: This term is not used by the MolDock score*
VdW (LJ12-6)	Protein steric interaction energy from a LJ 12-6 VdW potential approximation. *Notice: This term is not used by the MolDock score*
E-Inter (water - ligand)	The MolDockScore interaction energy between the pose and the water molecules.
E-Intra (tors, ligand atoms)	The total internal MolDockScore energy of the pose.
E-Intra (steric)	Steric self-interaction energy for the pose (calculated by PLP).

E-Intra (hbond)	Hydrogen bonding self-interaction energy for the pose (calculated by PLP). *Notice: This is a non-standard term and is zero by default – it must be enabled by specifying the ''internalhbond=true' option to the EVALUATOR initializer list in a MVDScript file or by enabling the 'Internal HBond' option in the Docking Wizard.*
E-Intra (elec)	Electrostatic self-interaction energy for the pose. *Notice: This is a non-standard term and is zero by default – it must be enabled by specifying the 'ligandes=true' option to the EVALUATOR initializer list in a MVDScript file or by enabling the 'Internal ES' option in the Docking Wizard.*
E-Intra (tors)	Torsional energy for the pose.
E-Intra (sp2-sp2)	Additional sp2-sp2 torsional term for the pose . *Notice: This is a non-standard term and is zero by default – it must be enabled by specifying the 'sp2sp2bond=true' option to the EVALUATOR initializer list in a MVDScript file or by enabling the 'Sp2-Sp2 Torsions' option in the Docking Wizard. Also notice that only bonds that are chosen rotatable are taken into account when calculating the torsional terms for the ligand – and sp2-sp2 bonds are most often double bonds which per default are held fixed in the docking simulation.*
E-Intra (vdw)	Steric self-interaction energy for the pose (calculated by a LJ12-6 VdW approximation). *Notice: This term is not used by the MolDock score*
E-Solvation	The energy calculated from the implicit solvation model. *Notice: This energy term is considered to be an* experimental *feature only. Per default it is NOT calculated. In order to try this feature, the protein must be prepared by calling the 'prep solvation' command from the console. As of now, we recommend not to use it.*
E-Soft Constraint Penalty	The energy contributions from soft constraints.
Static terms	
Torsions	The number of (chosen) rotatable bonds in the ligand.
HeavyAtoms	Number of heavy atoms.
MW	Molecular weight (in dalton).
C0	Obsolete constant term. This value is always 1. (Older versions of the Data Analyser required an explicit constant column, in order to include a constant term in the fit – it is only included for backward compatibility)
CO2minus	Number of Carboxyl groups in ligand.
Csp2	Number of Sp2 hybridized carbon atoms in ligand.
Csp3	Number of Sp3 hybridized carbon atoms in ligand.
DOF	Degrees of internal rotational freedom. As of now this is the number of chosen rotatable bonds in the ligand and is thus equal to the 'Torsions' term. It is supposed to reflect how many rotational degrees of freedom are lost upon binding. Future work may include a more advanced model where the actual conformation is inspected in order to determine whether rotational degrees of freedom are lost.
N	Number of nitrogen atoms in ligand.
Nplus	Number of positively charged nitrogen atoms in ligand.
OH	Number of hydroxyl groups in ligand.

OPO32minus	Number of PO_4^{2-} groups in ligand.
OS	Number of ethers and thioethers in ligand.
carbonyl	Number of Carbonyl groups in ligand.
halogen	Number of Halogen groups in ligand.
Other terms	
RMSD	The RMS deviation from a reference ligand (if available).

10 Appendix IV: Keyboard Shortcuts

The following list contains the keyboard shortcuts available in MMV. On Mac OS X, the CTRL key is replaced by the command key.

- CTRL-O Import Molecules
- CTRL-SHIFT-O Open Workspace
- CTRL-SHIFT-C Clear Workspace
- CTRL-S Save Workspace
- CTRL-F Toggle full screen
- CTRL-H Toggle dockable windows
- CTRL-C Toggle Cofactors category on/off
- CTRL-L Toggle Ligands category on/off
- CTRL-P Toggle Proteins category on/off
- CTRL-W Toggle Water category on/off
- CTRL-Z Undo
- CTRL-Y Redo
- CTRL-Q Quit MMV
- CTRL-1 to 8 Invoke misc. visualization views
- F1 to F9 Invoke misc. dialogs

11 Appendix V: Console Commands

When entering commands in the console, the following commands can be used.

Notice: Some commands require a *molecule target*: these can be described using the following syntax:

Ligand[0] – the ligand with ID 0.

Ligand[4,5,6] – the Ligands with IDs 4,5 and 6. Multiple IDs are separated by comma.

Ligands – All ligands. By using the plural form of a category, all molecules in it are selected. The categories are: Pose, Cofactor, Protein, Water, Ligand.

Poses;Cofactors;Proteins;Ligands;Water[0]– All Poses, Cofactors, Proteins, Ligands and the first Water molecule. Multiple targets can be concatenated using a semi-colon.

Notice: The IDs of molecules are based on the order of occurrence in the corresponding **Workspace Explorer** category. For instance, ligand molecules listed in the **Ligands** category, begins with index 0 with increments of 1 (i.e. 0,1,2,3,...). If molecules are removed from the workspace, the IDs of the molecules are changed to follow the new order of occurrence in the list.

Command	Description
SET [active|reference] [targetligand]	Set active or reference ligand. A lot of operations (e.g. some surfaces) are only performed on the active ligand. The reference ligand is used to calculate RMSDs while docking.
EXPORT [moleculetarget]	Export as Mol2 or PDB. A File export dialog is opened for selection of a filename.
SURFACEDIALOG	Shows the Surface dialog.
PREPAREDIALOG	Shows the Preparation wizard.
LABELDIALOG	Shows the Label dialog.
GETPDB <key>	Downloads PDB with 'key' (4 letter code) from the Protein Data Bank.
ALIGN [MoleculeTarget1] [id1] [id2] [id3] [MoleculeTarget2] [id1] [id2] [id3]	Aligns atom id1,id2,id3 in MoleculeTarget1 with atom id1,id2,id3 in MoleculeTarget2.
SHOW CATEGORY <category> HIDE CATEGORY <category>	Shows or hides Workspace Explorer category with given name: i.e. SHOW CATEGORY water
REMOVE OBJECT [id]	Removes a 3D object from the world.
REMOVE [moleculetarget]	Removes an object from the workspace.
CD	Print current directory.
DIR	Shows dir of MVDML files in current directory.
PREV	Loads previous MVDML file in current directory.
NEXT	Loads next MVDML file in current directory.
RMSD	Invokes RMSD dialog.

SELECT ID <id> SELECT ATOM <x y z> SELECT RESIDUE <id> SELECT RESIDUEID <id>	Selection of objects: 'SELECT ID' selects all atoms with id = 'id'. 'SELECT ATOM' selects closest atom to specified x, y, z position. 'SELECT RESIDUE' selects residue with residue index = 'id'. 'SELECT RESIDUEID' selects residue with internal residue index = 'id'.
STATUS	Shows info about the objects in the workspace and Visualization Window. Loaded modules are also listed.
SAVE [filename]	Saves a MVDML file. Do not include extension in filename.
LOAD [filename]	Loads a MVDML file. Do not include extension in filename.
ADD {P}SURF{*}{+} {radius} {resolution} {probesize} ADD {P}SURF{*}{+} LIGAND {resolution} {probesize}	Adds a molecular surface. *Notice: It is much easier to use the Surface dialog in the GUI*. If prepended by 'p' the surface will be colored by electrostatic potential. If followed by '*' the surface will *carved* (a Connolly surface). We recommend the carved surfaces for best visualization. If not followed by '+' existing surfaces will be removed. If no radius is given, the surface will cover the protein. If a radius is given, the surface will cover the *protein* in the given radius, but centered at the current active ligand! If the variant with 'LIGAND' argument is used, the surface will cover the currently active ligand. Resolution is typically 0.4 - 0.9. Don't choose higher resolutions (i.e. lower) than 0.4 unless you are prepared to wait for a long time! The default value of the probesize (1.2) should be fine for most purposes. Examples: ADD PSURF*+ 10 (Electrostatic carved protein surface with radius 10Å centered at the active ligand) ADD SURF* ligand 0.4 (Carved ligand surface with resolution 0.4)

DELETE [Charges| HBOND|Labels|Poses]	Deletes specified objects.
LIST POSES	Shows all poses with info about parameters.
CLS	Clears console log.
CLEAR [workspace| selection]	'CLEAR workspace' removes all items in the current workspace. 'CLEAR selection' clears current selection.
HIDE [hydrogens | labels]	Hides either hydrogens or labels.
SHOW [hydrogens | labels]	Shows either hydrogens or labels.
FITTOSCREEN	Fit all molecules in the visualization window.
LABEL	Used for labeling objects. This command is described in detail in the paragraph below. *Notice: It is much easier to use the Label dialog in the GUI*
GUI Commands	
SLAB [near] [far]	Creates a slab (slicing) of the 3D world. *Notice: The Clipping Planes dialog is easier to use.*
QUALITY [value]	Sets OpenGL rendering quality from 0 to 10.
LIGHT [number] [on|off] [ambient] [diffuse] [specular] {[x] [y] [z]}	Sets OpenGL light sources.
FOG LINEAR [near] [far] FOG [EXP|EXP2] [exponent] FOG OFF	Sets OpenGL fog.
COLOR [protein|pose| ligand|water|cofactor] [fixed|cpk|hbond| hbond2|interaction| interaction2] {r g b}	Sets the color style of specified object. For more information about color styles, see the 'Visualization Settings' dialog section.
STYLE [protein|pose| ligand|water|cofactor] [vdw, fixed, stick, wireframe, none] atomScale bondScale	Sets the visualization style of specified object. The last parameter lineWidth is only used in wireframe mode, and is the line width in pixels. For more information about graphical styles, see the

lineWidth	'Visualization Settings' dialog section.
PROJECTION [perspective| orthogonal] angle	Determines perspective projection mode. Angle is the field-of-view angle for perspective projection. For more information see the 'Visualization Settings' dialog section.
BACKGROUNDCOLOR r g b	Sets the background color
LABELCOLOR r g b	Sets the labelling color
CAVITYCOLOR r g b	Sets the cavity color
REBUILD	Rebuilds all objects in the Visualizer Window. This command is necessary to call after the visualization styles or coloring schemes have been updated. Otherwise graphical changes will not be reflected in the GUI.

The label command works in the following way: it scans the input-string for known variables (like ID, HYB, ELE - see below) and replaces them with their value. That is, the command 'label bond bond_number:id' will add a label of type 'bond number x' to every bond (underscores are replaced with spaces). To clear all labels use 'label' without any argument.

Variable	Description
Atom labels. Syntax: '**Label *string***'	
ID	Internal atom index
Type	Hydrogen bond type: non-polar, acceptor, donor, both. The HBOND variable below is probably of more use.
PC PC!	Partial Charge. PC! ignores atoms with no partial charge.
HYB HYB!	Hybridization. HYB! only displays hybridization for atoms with other hybridizations than SP3 or unknown.
SP2	Labels SP2 hybridized atoms
SYM	Element symbol. (H, C, N, ...)
ELE	Element number.
IH	Number of implicit hydrogens.

HBOND HBOND!	Hydrogen bond type shown as : D, A, D+A, - (non-polar) HBOND! ignores non-polar atoms.
ETOT	Shows the total energy of the atom. This requires that the energy has been evaluated using the 'eval' command.
Bond labels. Syntax: '**Label bond *string***'	
ID	Internal bond index.
Type	Bond order: single, double, triple, aromatic, ...
ETOT	Shows the total energy of the bond. This requires that the energy has been evaluated using the 'eval' command.
Torsion Tree Labels. Syntax: '**Label tree *string***'. Torsion Tree labels are evaluated at each rigid component of the ligand.	
EA, *EA	EA shows the Atom Energy of the rigid component . EA* shows the difference between this energy for the pose and for the ligand. *It is only well-defined on poses.*
EB, *EB	EB shows the Bond Energy of the rigid component. EB* shows the difference between this energy for the pose and for the ligand. *It is only well-defined on poses.*
ETOT, *ETOT	Shows the total energy of the rigid component of the tree. This requires that the energy has been evaluated using the 'eval' command. ETOT* shows the difference between this energy for the pose and for the ligand. *It is only well-defined on poses.*
DEPTH	The depth of the torsion tree. Higher values indicate that the rigid component is built later. The rigid root component will have depth 0. Optimal placement of the rigid root will result in a lower maximum depth.
BOND	Shows the bond index that this rigid component rotates about.
ID	Internal index
ATOMS	Number of atoms in rigid component.

Residue Labels. Syntax: '**Label residue *string***'	
ID	Internal residue index
LONGNAME	Full residue name ('histidine', 'cysteine', ...)
NAME	3-letter abbreviation ('HIS', 'CYS', ...)
LETTER	1-letter abbreviation.

12 Appendix VI: Third Party Copyrights

Icons

The icon set used in MMV is taken from:

The Tango Icon Library: http://tango-project.org

They are released under the 'Creative Commons Share-Alike license':
http://creativecommons.org/licenses/by-sa/2.5/

13 Appendix VII: References

[THOMSEN 2006] Thomsen, R.; Christensen, M. H. MolDock: A New Technique for High-Accuracy Molecular Docking. *J. Med. Chem.*, 2006, 49(11), 3315-3321.

[GEHLHAAR 1995] Gehlhaar, D. K.; Verkhivker, G.; Rejto, P. A.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Docking Conformationally Flexible Small Molecules Into a Protein Binding Site Through Evolutionary Programming. *Proceedings of the Fourth International Conference on Evolutionary Programming,* 1995, 615-627.

[GEHLHAAR 1998] Gehlhaar, D. K.; Bouzida, D.; Rejto, P. A. Fully Automated And Rapid Flexible Docking of Inhibitors Covalently Bound to Serine Proteases. *Proceedings of the Seventh International Conference on Evolutionary Programming* 1998, 449-461.

[YANG 2004] Yang, J-M.; Chen, C-C. GEMDOCK: A Generic Evolutionary Method for Molecular Docking. *Proteins,* 2004, *55*, 288-304.

[MCDONALD 1994] McDonald, I. K.; Thornton, J. M. Satisfying Hydrogen Bonding Potential in Proteins. *J. Mol. Biol.,* 1994, *238*, 777-793.