Trajectory Plots Guide

Introduction

In this guide, we will use the Trajectory Plots functionality within Maestro's Trajectory Player to explore molecular events occurring in the simulation of a protein-ligand complex. Trajectory Plots is a new, feature-rich tool within the Trajectory Player that allows interactive analysis of simulation trajectories written by Desmond – Schrodinger's molecular dynamics (MD) engine. Here we will give several examples of how to use Trajectory Plots for common analysis tasks. These tasks include: tracking measurements throughout the trajectory; setting up RMSD and RMSF calculations; doing the energetic decomposition of system components; exporting data and images.

We are also deprecating the Simulation Event Analysis (SEA) panel, the original simulation-analysis tool previously available through Maestro. This tool will be removed completely in future releases, but as of *Schrödinger Suite 2022-1*, the SEA panel is disabled. Please contact the support team for the command to enable it. We are excited about releasing Trajectory Plots and we would like to hear your feedback regarding this tool, please send it to help@schrodinger.com.

Preparing and Running MD Simulation

For this guide, the system we are using for illustration is 3C-like cysteine protease (CoV-2 $3CL^{pro}$) in a complex with a PF-00835231 covalent ligand. For more information about the system, see *Hoffman et al., J. Med. Chem.* ***2020*** *(63) 21*.

We prepared the 6XHL structure from the *RCSB* repository using the Protein Preparation Workflow, keeping chain A of the crystal complex. The prepared complex was then solvated with an SPC cubic water box with counterions and 150 mM of NaCl salt using the System Builder. The Molecular Dynamics panel was used to equilibrate the system followed by a 100-nanosecond NPT simulation production run. If you want to use this system as an example, you could prepare the structure in the same way, or use the input files located at X.

Trajectory Player

Inputs for this tutorial can be found here. To view the simulation of the production stage, load the output CMS file into Maestro and double-click on the **T** icon next to the entry name. This should open the Trajectory Player at the bottom of the workspace. After setting up the system representation (hiding water/ions and showing the ligand in *ball-and-stick* representation) the trajectory can be animated with the controls in the Trajectory Player.

The Trajectory Plots feature is accessible through the 'Plot' menu on the Trajectory player, see *Figure 1*.

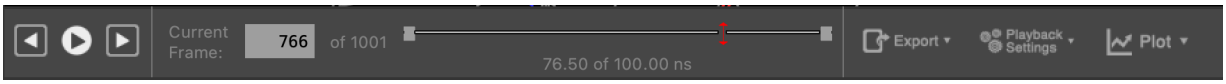
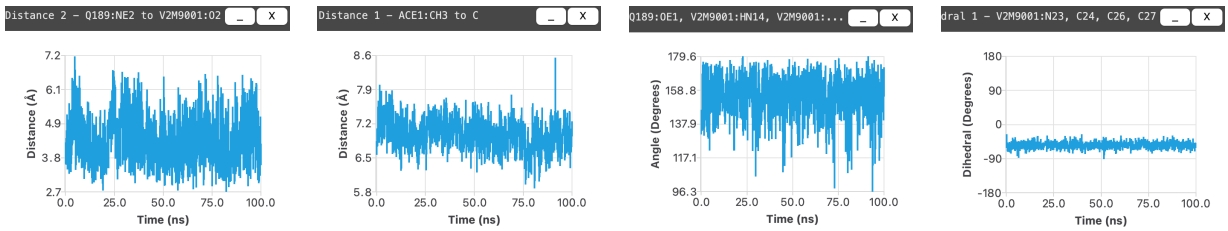


Figure 1. Trajectory player controls with Plot option (on the left)

Geometric Measurements

In the Maestro Workspace, it is easy to set up measurements such as distances, angles, dihedrals. Clicking the 'Measure' button on the 'Favorites Toolbar' allows setting up measurements in the Workspace. These measurements can then be tracked throughout the simulation by going to Plot → Measurements → "Currently in Workspace." Clicking on the plot updates the Workspace to the corresponding frame in the simulation, and all the Workspace measurements are updated.



A. Distance between two atoms

B. Distance between the centroids of two ligand rings

C. Angle between a key hydrogen bond

D. Dihedral angle within the ligand.

Figure 2. Tracking Workspace measurements throughout the simulation.

Interaction Counts

Types of Interactions

Within the “Interaction Counts” menu several molecular interactions can be tracked: Hydrogen Bonds, Halogen Bonds, Salt Bridges, π-π stacking, π-Cation interactions. By default, the interactions within the ligand and receptor selections are shown, but these selections can be changed within the Interactions Menu at the lower right corner of Maestro, see *Figure 3*.

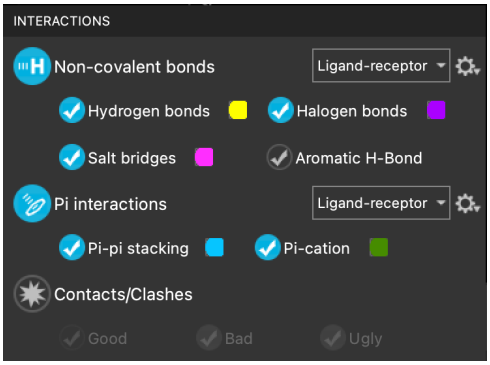


Figure 3. Changing Interactions selections

Counting number of Hydrogen Bonds

The commentary on how the compound was developed mentions the importance of optimizing the hydrogen-bond network between the ligand and the receptor. Molecular interactions between the protein receptor and the ligand can be viewed by toggling on the interactions button  . You can then calculate the number of hydrogen bonds for the same selection by choosing Plot → Interaction Counts → “Hydrogen Bonds.” The calculated HB count should appear on the left panel within a few seconds, see *Figure 4*. Clicking on the plot updates the

Workspace to the corresponding frame in the trajectory. Similarly, other interaction type counts can be calculated, or all types by selecting Plot → “Interaction Counts” → “All Types.”

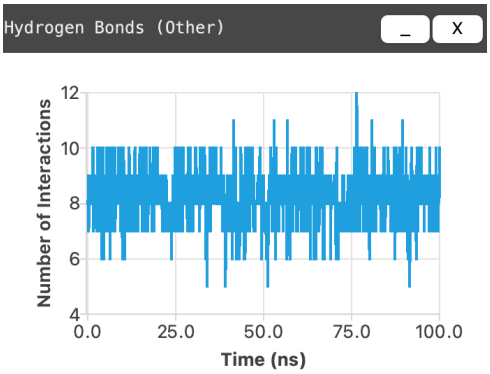


Figure 4. Hydrogen Bond Count

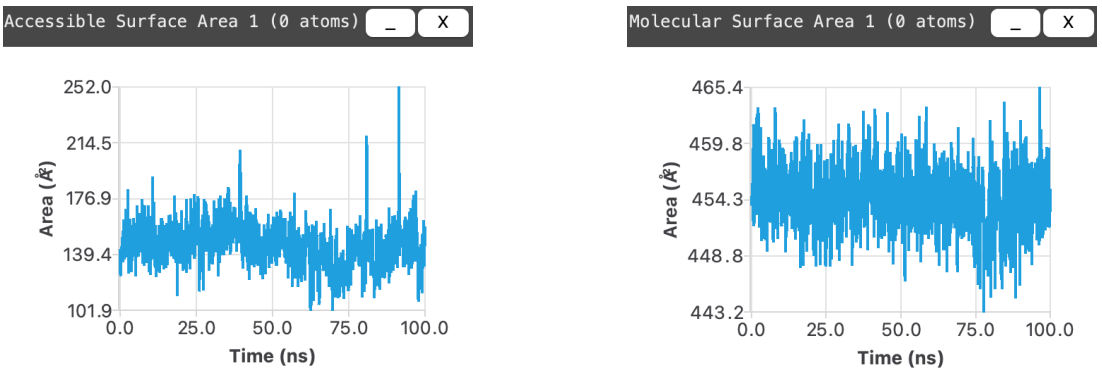
Molecular Descriptors

You can calculate multiple molecular descriptors with Trajectory Plots. The supported analyses within this category are RMSD, Radius of Gyration, Solvent Accessible Surface Area (SASA), Polar Surface Area (PSA), Molecular Surface Area (MSA). To calculate these descriptors, first, a selection needs to be made in the Workspace, followed by a selection of the analyzer.

Calculating Solvent Accessible Surface Area

In our example system, we would like to explore how solvent-exposed the ligand is within the receptor. This can be done by calculating the SASA of the ligand over the course of the simulation. To do this, we need to first select the ligand and then choose Plot → Descriptors → Solvent Accessible Surface Area. The time series for Ligand SASA throughout the simulation appears within a minute on the right side of the Workspace, see *Figure 5A*. Clicking on the plot updates the Workspace to the corresponding frame in the trajectory.

To get a sense of how exposed the ligand is in terms of the fraction of its total surface area, the steps above can be repeated, selecting MSA instead. By exporting the data of these plots and taking the ratio values by using a spreadsheet or scripting tools, you can obtain the relative solvent-accessible SA to the total molecular SA.



A. Solvent Accessible Surface Area

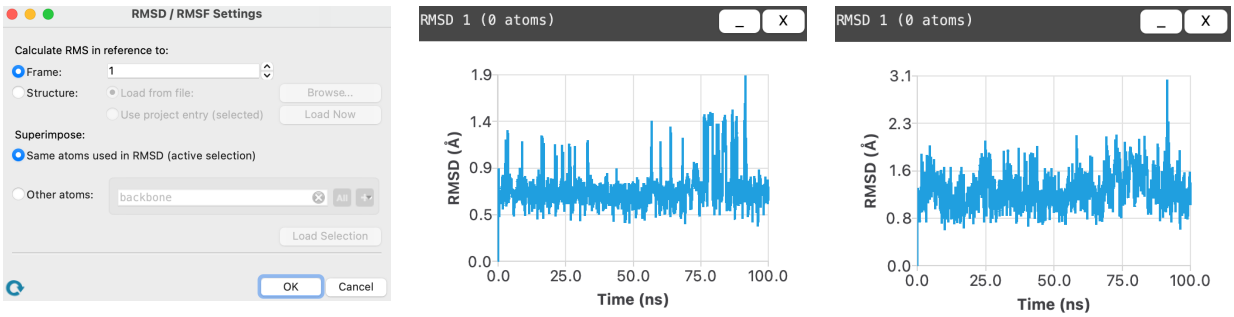
B. Polar Surface Area

Figure 5. Molecular Descriptors

Calculating RMSD of the Ligand

Throughout the analyses of molecular simulation data, one common task is to calculate the Root-Mean-Square Deviation of the ligand. For molecules of 150 heavy atoms or less, the RMSD calculation takes into account the symmetric substructures, this is to prevent the rotation of symmetric rings like *phenyl* or other groups like *trifluoro* and *methyl* groups. In larger biomolecules, symmetry is ignored during RMSD calculations.

To calculate ligand RMDS, we first need to set up the selections for reference coordinates and alignment selections. To open the RMSD Settings panel, choose Plot → Descriptors → View RMSD Settings. To measure ligand RMSD in reference to its initial conformation, select the “Same atoms used in RMSD (active selections)” option. To set up the calculation of RMSD to first align the ligand on the receptor, select “Other atoms" and enter “backbone” in the text box. Finally, to generate the RSMD plot select Plot → Descriptors → RMSD.



A. RMSD Settings panel

B. RMSD of the ligand with respect to the first frame conformation

C. RMSD of the ligand aligned on the first frame of protein backbone

Figure 6. RMSD settings and plots of the ligand RMSD

Root-Mean-Square Fluctuations

Another common task when analyzing simulations containing protein is to look at the fluctuation of the protein residues by calculating Root-Mean-Square Fluctuation (RMSF). Trajectory Plots offers two types of RSMF analyzers, one where the atoms in the selections are grouped by residues (“per residue”), and one where no grouping is done (“per atom”). Per-residue analysis works well for proteins, while the per-atom analysis works well for smaller molecules.

To generate an RMSF analysis of the protein receptor, select the protein backbone atoms with Quick Select on the toolbar and then launch the analysis by choosing Plot → RMSF → Per Residue. This generates an “advanced” plot, which shows up as an icon at the bottom of the plots. Double-clicking that icon opens the Protein RMSF plot. This plot packs in several computed properties: per residue fluctuations of the protein selection, experimental temperature B-factors that are stored within the structure, and the computed secondary structure of all residues (alpha-helix and beta-sheets are shown in red and blue, respectively). Hovering over the plot shows the residue information plus the RMSF/B-factor values associated with it. Clicking on the plot shows the residue in the Workspace.

Figure 7. RMSF plot of the protein receptor

Energy Analysis

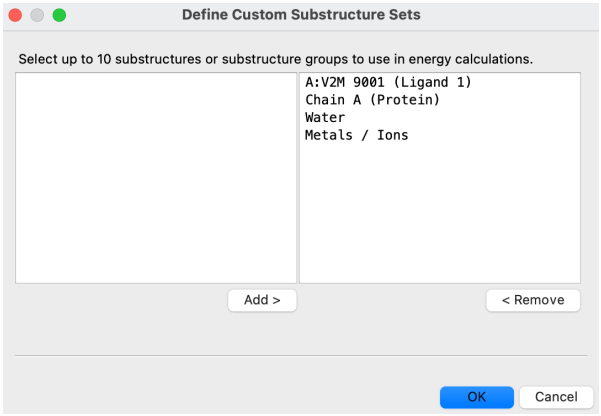
Looking at different energy components of a system is often a very useful analysis to determine the energies and forces that dominate and therefore drive the simulation to equilibrium. Trajectory Plot provides a simple interface to set up and analyze such analyses. Suppose you want to look at the internal energy of the ligand and correlate that to conformational changes

detected with another analyzer, or to explore the interaction energy between a protein and the receptor. This can all be done with energy analysis within Trajectory Plots.

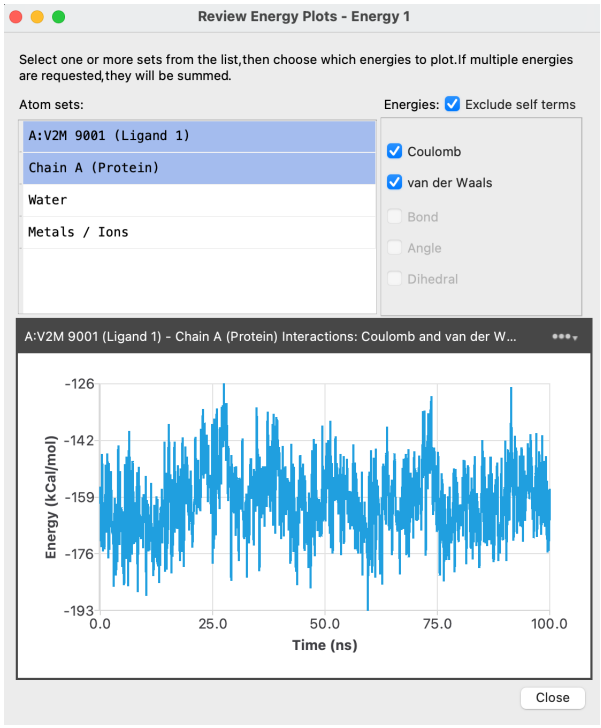
Energetic analysis requires access to a host machine with Linux and a supported GPU card. The host can be selected by choosing Plot → Energy Calculation → Job Settings.

For energetic analysis, we need to define substructure groups, which can be done in Plot → Energy Calculation → Define Custom Substructure Sets. For this example, we select four components: ligand, receptor, waters, and ions (see *Figure 8A*). Clicking OK submits the calculations to the previously selected host. *Note: these calculations take some time to complete as this analysis requires transferring large files to the selected host machine and launching Desmond to calculate the energies for the selected subgroup sets.*

The results are incorporated into the Advanced Plots area of the Trajectory Plots, and double-clicking the Energy icon opens the “Review Energy Plots” panel. Through this panel, various terms can be explored for selected sets. Within a single set, five energy terms can be explored: Coulomb (electrostatic), van der Waals (vdW), bond, angle, dihedral. Selecting all terms will show the total energy for the selected sets. To explore interactions between two sets, like the ligand and the receptor, selecting ‘Exclude self terms’ shows just the interaction energy between these terms (*Figure 8B*). Clicking on the plot automatically changes the frame number corresponding to that point in the Workspace.



A. Define custom sets



B. Explore energies of selected subgroups

Figure 8. Energy Analysis of the receptor-ligand complex

Exporting Trajectory Plot Data

The underlying data generated by Trajectory Plots can be exported in either CSV or Excel format by right-clicking on the time series plot (or on the three-dot menu in the Advanced plots like RMSF and Energy Analysis). Similarly, the export of the plot to PNG format in high-resolution is available as well through the Save Image selection.

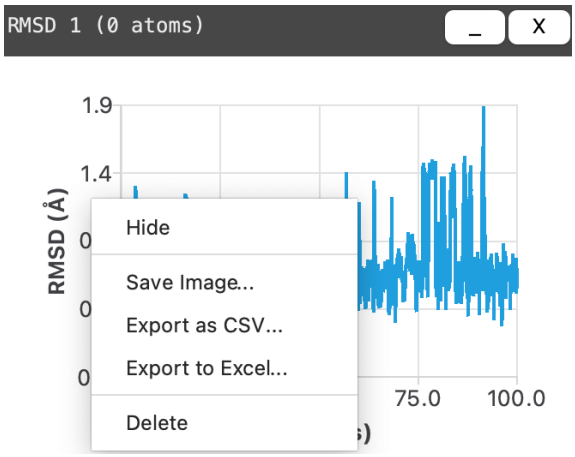


Figure 9. Exporting data and images in Trajectory Plots