Bridge
User's Manual

September 30, 2019

Malte Siemers
FU Berlin
Department of Physics

Contents

1 Description of the Tool

Bridge is a set of algorithms for the detection and analysis of networks comprised of hydrogen bonds or water wires. Bridgewas published in [Siemers and Lazaratos, 2019] and we recommend to read the article for a detailed study of the possible applications of the tool.

1.1 How to install

Bridge comes as a pymol plugin. It depends on MDAnalysis (at least ver. 0.19.2). To install pymol with MDAnalysis, we recommend using the miniconda python distribution (https://www.anaconda.com/). If miniconda is set up correctly, you can install pymol and MDAnalysis by typing:

```
conda install -c schrodinger pymol
conda install -c conda-forge mdanalysis
```

To open pymols plugin manager, click Plugin->Plugin Manager and in the plugin manager, go to the 'Install New Plugin' tab and click 'Choose file ...'. Select the .zip file containing Bridge and restart pymol. Now start Bridge by clicking Plugin->Bridge.

1.2 Reinstall

To reinstall, click on 'Plugin' in the pymol menu. Then on 'Plugin Manager' and find 'Bridge' in the list. click on uninstall. Restart pymol and click 'Plugin' again, then 'Plugin Manager' and go to the 'Install new plugin' tab. Click 'Choose file...' and select the zip I attached. After it says something about installing correctly, you can find 'Bridge' in the 'Plugin' menu of pymol.

1.3 Workflow

Bridgeis designed to work in a sequential way. The typical workflow involves the following steps in order:

1. Initialize
2. Search
3. Filter
4. Compute & Visualize

The rest of this document is structured to explain those four steps. Bridge can be used to analyse networks of hydrogen bonds OR water wires. The possible analyses for those two differ slightly, but the general workflow enumerated above remains the same. Here, a short summary of the whole process is presented. A detailed description of every step follows in the next sections.
Initialization sets necessary variables such as the filenames containing the data that is to be analyzed and the selection of atoms to be considered in the analysis. Initialization involves the same necessary and optional steps for water wires and hydrogen bonds. Search requires geometric parameters for the definition of hydrogen bonds and a choice for a specific algorithm for the detection of hydrogen bonds or water wires and their parameters. The selected algorithm will define the initial set of hydrogen bonds or water wires used in the later analyses and running the selected algorithm is the time consuming part of Bridge. The Filter section provides means for dissecting the initial results computed in Search. Filters are rules to reduce the number of hydrogen bonds or water wires, e.g. to find connected parts of the network or specific paths between amino acid residues. The Compute & Visualize section provides histograms and plots as well as their underlying data in ASCII format. Bridge can also visualize the initial and filtered results as a 2D projection. Several plots can compare data of different analyses. Bridge can save the sate of an analysis and can compare it to other states. This might be the same analysis for different proteins or different analyses for the same protein.

1.4 Naming Convention

Bridge addresses amino acid residues by concatenating their segment id, residue name and residue id with a minus sign (-). Nodes in the graph are labeled with this code and user input has to be structured in that way. To address glutamate 98 with the segment id SEG1 the user would have to type SEG1-GLU-98. If analyses are done on atom level, the atom name has to be added in the same fashion. SEG1-GLU-98-OE1 addresses the oxygen OE1 in glutamate 98 in segment SEG1. The word segment refers to the 'segid' selection command in MDAnalysis, residue names are the three-letter-codes for amino acids and residue ids or numbers refer to the 'resid' selection command in MDAnalysis. Atom names refer to the 'name' command in MDAnalysis.

2 Initialize

This part is the same for hydrogen bonds and water wires: In both cases the fields called

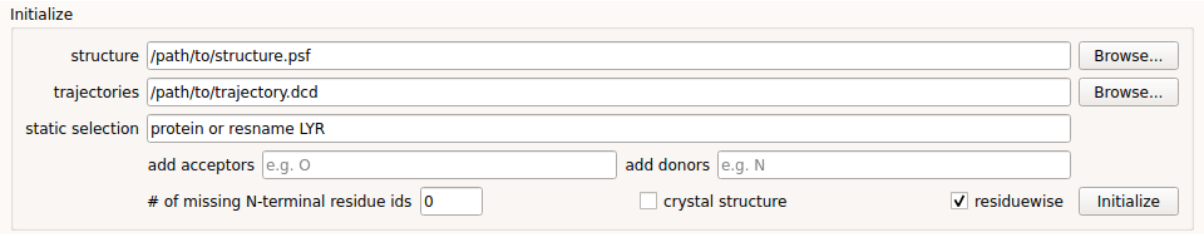

Figure 2.1: Screenshot of the Initialize section of Bridge.

structure and static selection are necessary options, the rest is optional.

structure A structure file has to be provided using the Browse button or by directly typing it. Bridge works with all formats that are compatible with MDAnalysis. We tested .psf, .pdb and .gro

trajectories One or more trajectory files can be specified. They are concatenated in the order in which they appear in the field. Using the Browse button, multiple files can be selected by pressing Ctrl while clicking on them. For manual input, filenames have to be separated by a comma.

static selection Bridge uses static selections that are compatible with the MDAnalysis selection syntax. Static selections do not contain time dependent parameters, such as a geometric criterion. A valid static selection would be "resname GLU or resname POPC". For further information on the syntax consult the MDAnalysis selection syntax documentation at `https://www.mdanalysis.org/docs/documentation_pages/selections.html`.

add acceptors Atom names defining acceptors can be defined in this field. By default, we exclude N and O, so if you want to analyse backbone-backbone or backbone-sidechain interactions, you have to include "O" in this field.

add donors Atom names defining acceptors can be defined in this field. By default, we exclude N and O, so if you want to analyse backbone-backbone or backbone-sidechain interactions, you have to include "N" in this field.

# of missing n-terminal residues Correct numbering if internal and pdb numbering differ. This only affects labels in plots and not the hydrogen bonds or water wires detected.

crystal structure Check this toolbox, if there are no hydrogen defined in your structure file. Bridge will then treat unprotonated donor atoms as hydrogen bond donors and will rely only on distance for hydrogen bond detection.

residuewise hydrogen bonds are always calculated per atom. This toolbox decides the format in which the hydrogen bonds or water wires are saved internally. If residuewise is checked, all later analyses are done on the level of residues. For water wires this will always save the shortest possible wire between two residues.

3 Search

Here the geometric criterion for hydrogen bonds as well as a specific algorithm for the detection of hydrogen bonds or water wires have to be defined. The geometric criteria apply for both water wires and hydrogen bonds, while the detection algorithms differ.

3.1 WATER WIRES

In this section we will describe every option and field for hydrogen bonds and water wires.

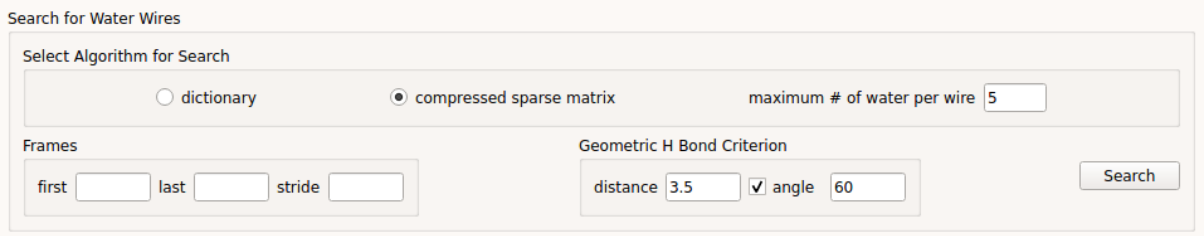

Figure 3.1: Screenshot of the Search section for water wires

For both, hydrogen bonds and water wires, the grouped options with the titles Frames and Geometric H Bond Criterion are the same:

first defines the first frame to be analyzed in the trajectory. If multiple trajectories are supplied, they are concatenated in the give order and the internal numbering is continuous.

last defines the last frame to be analyzed in the trajectory.

stride sets a stepping in the frames. A stride of 2 will analyze every second frame, 3 every third, etc.

distance sets the maximum distance between heavy atoms of hydrogen bonding partners.

angle enables or disables angle calculation for the detection of hydrogen bonds and sets the maximum angle. The angle is defined as the angle between the lines connecting

donor heavy atom and hydrogen and the line between hydrogen and acceptor heavy atom. An angle of zero corresponds to a straight line through the donor heavy atom, then the hydrogen and then the acceptor heavy atom.

The choice for algorithms is different for water wires and hydrogen bonds. The choice for water wires makes no difference in the resulting set of water wires. The algorithms only differ in the internal representation of the graphs used to find the shortest water wires and a corresponding speed difference.

dictionary should be used for small systems or systems with a low expected density of hydrogen bonds.

compressed sparse matrix should be used for large systems and systems with a high density of hydrogen bonds. Since the computation time is generally negligible for small systems with both options, this is the recommended default option.

maximum # of water per wire define the maximum number of water molecules a wire can be made of. Longer wires will not be detected. Bridge will always detect the shortest wire between two residues or atoms. This option has a large effect on the computation time.

3.2 HYDROGEN BONDS

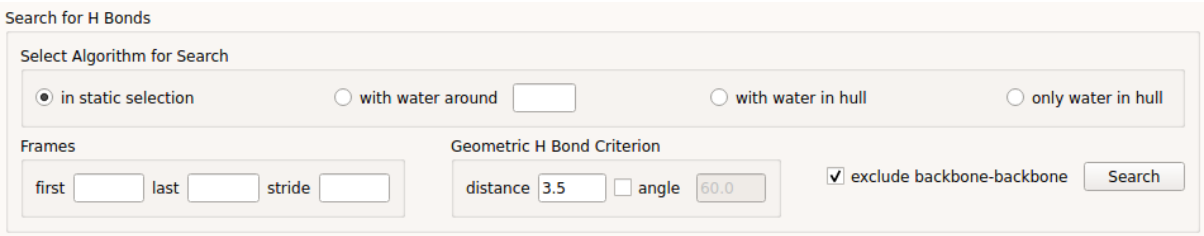

Figure 3.2: Screenshot of the Search section for hydrogen bonds

The following options are the same as for water wires:

first defines the first frame to be analyzed in the trajectory. If multiple trajectories are supplied, they are concatenated in the give order and the internal numbering is continuous.

last defines the last frame to be analyzed in the trajectory.

stride sets a stepping in the frames. A stride of 2 will analyze every second frame, 3 every third, etc.

distance sets the maximum distance between heavy atoms of hydrogen bonding partners.

angle enables or disables angle calculation for the detection of hydrogen bonds and sets the maximum angle. The angle is defined as the angle between the lines donor heavy atom - hydrogen and hydrogen - acceptor heavy atom. An angle of zero corresponds to a straight line through donor heavy atom - hydrogen - acceptor heavy atom.

The options for the choice of algorithm differ from the selection for water wires:

in static selection This algorithm will only detect hydrogen bonds within the static selection specified in the initialization. No water is considered.

with water around Here, a radius has to be defined. This radius defines a space around the atoms of the static selection that is dynamically updated every frame. Within this space water is considered as a possible hydrogen bonding partner. All possible hydrogen bonds are detected: water-water, water-residue and residue-residue.

with water in hull A convex hull is spanned around the point cloud defined by the positions of the possible hydrogen bonding heavy atoms every frame. Water within this space is considered as possible hydrogen bonding partners in addition to the static selection. All possible hydrogen bonds are detected: water-water, water-residue and residue-residue.

only water in hull A convex hull is spanned around the point cloud defined by the positions of the possible hydrogen bonding heavy atoms every frame. Only water within this space is considered as possible hydrogen bonding partners. Only water-water hydrogen bonds are detected. This is significantly faster than the above algorithm.

4 Filter

Bridge can apply a combination of filters to the initially found set of hydrogen bonds or water wires. This reduces the number of wires or bonds. The order in which the filters are applied is important, because it can affect the results. All later operations like plots or computations will be based on the filtered set of wires or bonds.

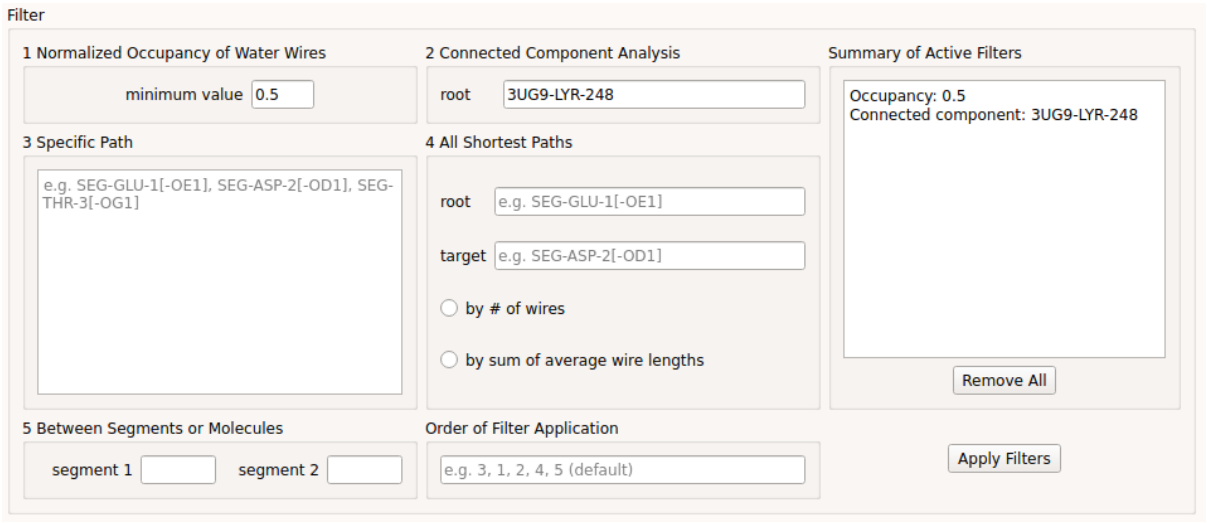

Figure 4.1: The Filter section of Bridge for water wires. The only differences for hydrogen bonds from water wires is no possibility to define how shortest paths are computed and the option to filter for backbone-backbone interactions.

4.1 1 Normalized Occupancy

minimum value Hydrogen bonds or water wires with a lower value are not considered in further computations.

4.2 2 Connected Component Analysis

root This is a graph based filter. The graph will be cropped to the connected component containing the root node.

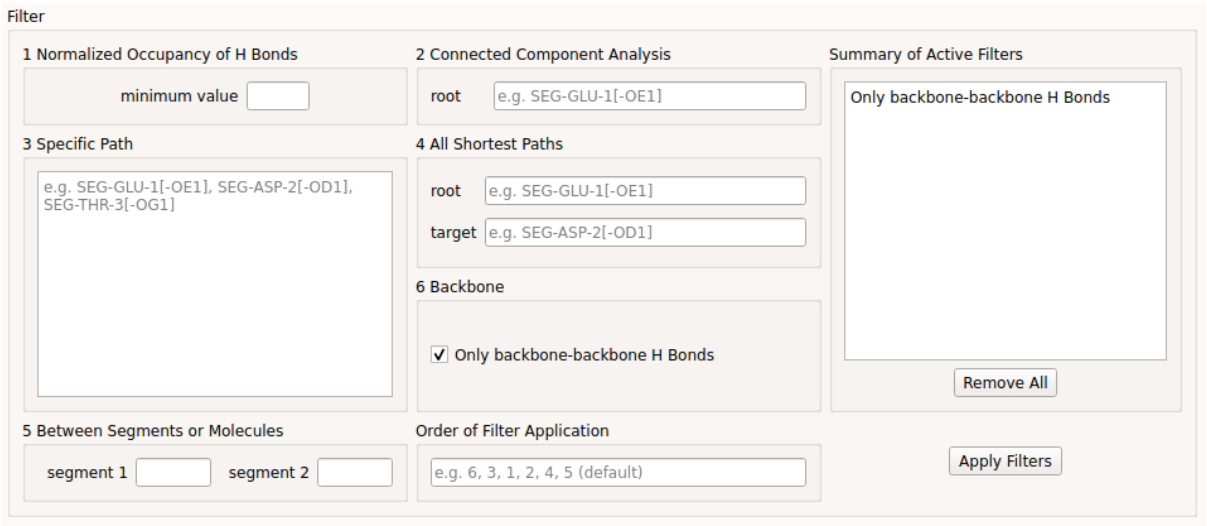

Figure 4.2: The Filter section of Bridge for hydrogen bonds.

4.3 3 Specific Path

The graph will be reduced to the specific path in the text field. The path has to exist in the graph in the order given in the text field. The nodes have to be separated by comma and in the convention described in section 1.4.

4.4 4 All Shortest Paths

root and target The graph will be reduced to all nodes and edges belonging to a shortest path between the root and the target. There can be multiple shortest paths and a reduction to one path has to be made with the specific path filter. Both root and target have to follow the nameing convention described in section 1.4.

by # of wires Specific to water wires. This defines "shortest" in shortest paths. Checking this option counts the number of wires to be taken to reach target from goal.

by sum of average wire lengths Specific to water wires. This defines "shortest" in shortest paths. Checking this option computes the average number of waters between root and target. This can lead to different results than "by # of waters". E.g., If three wires with a length of 1 connect root and target and one wire with the length of 5, toggeling between "by # of waters" and "by sum of average wire

lengths" will switch between those two.

4.5 5 Between Segments or Molecules

segment 1 If only segment 1 is specified, the results are reduced to water wires or hydrogen bonds with at least one partner in the specified segment.

segment 2 If segment 1 and segment 2 are specified, only bonds or wires remain, that comprise of one partner in segment 1 and the other partner in segment 2.

4.6 6 Backbone

This filter is specific to hydrogen bonds analysis only. It removes all hydrogen bonds that involve side chain atoms.

4.7 Order of Filter Application

The order has to be supplied as 1 to 5 distinct numbers from {1,2,3,4,5} separated by comma. The numbers relate to the numbers shown in front of every filter title.

4.8 Active Filters

The summary of active filters only shows filters that are active on the data. They can be removed by pressing the 'Remove All' button. Further computations will rely on the filtered data using the active filters. If a combination of filters leads to an empty selection, the user will be informed. Filters can be activated by typing valid parameters into the respective fields of the filter and pressing the 'Apply Filters' button. Pressing the 'Remove All' button deactivates all filters and restores the initial results from the Search section.

5 Visualize and Options

Bridge can visualize the initial and filtered results as a 2D projection as a graph. The projection is computed using Principal Component Analysis and is deterministic for a specific set of coordinates and not modifiable. This provides a direct visual feedback to the filters and is meant to guide the analysis. The graphs are made in the same way for hydrogen bonds and water wires and have the same options:

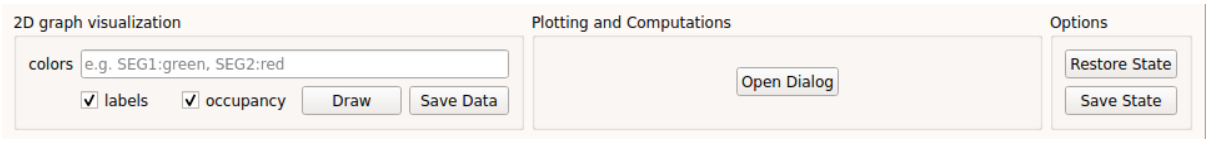

Figure 5.1: Visualization, computations and plotting section of Bridge. "2D graph visualization" creates a graphical representation of the initial or filtered data by showing hydrogen bonding or water wire connected partners as nodes connected by edges in a graph using a 2D projection of their coordinates. "Plotting and Computations" opens a new dialog. "Options" can be used to save and restore analysis states.

5.1 2D GRAPH VISUALIZATION

"

colors The visualization automatically detects segments and colors nodes with respect to the segment they belong to. Custom coloring can be defined in this field with comma separated pairs in the form of segmentid:color. A valid option would be "SEG1:green, SEG2:red"

labels If checked, nodes are labeled with respect to the residues or atoms they represent.

occupancy If checked, edges are labeled with the occupancy of the water wire or hydrogen bond they represent.

Draw Show the graph.

Save Data Save the list of water wires or hydrogen bonds as a text file. The first column contains the water wired or hydrogen bonding partners coded depending

on the state of "residuewise". If "residuewise" is checked, the format is segmentid-residuename-residueid, otherwise segmentid-residuename-residueid-atomname.

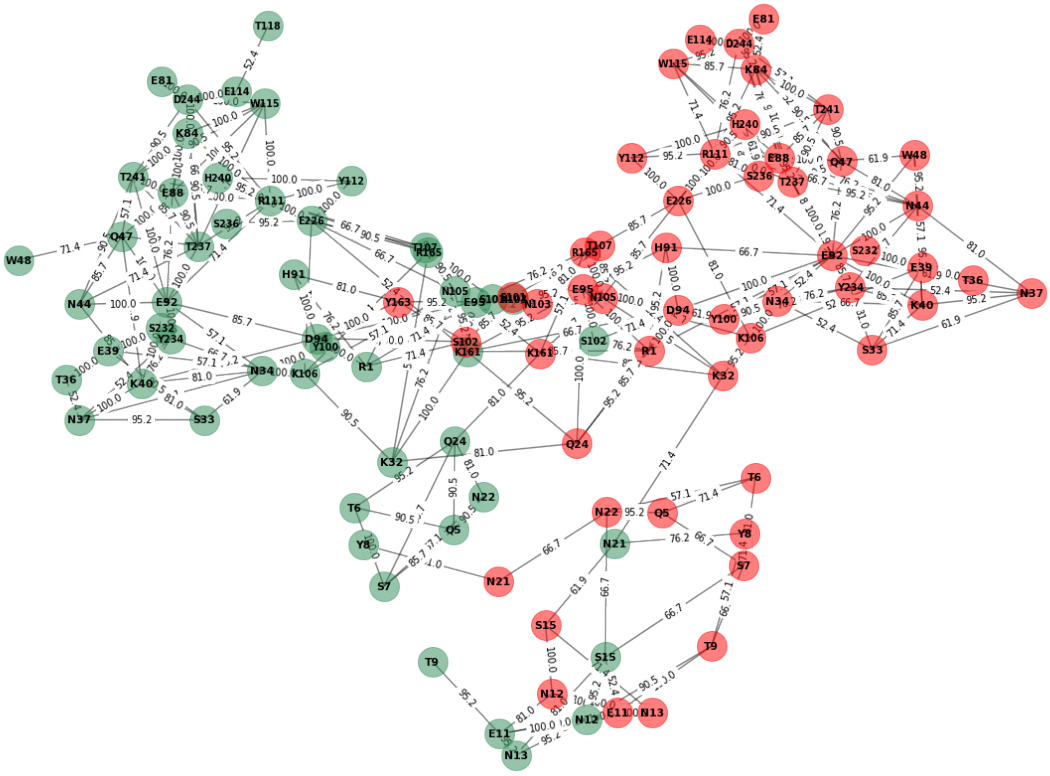

Figure 5.2: Graph representation of the filterd set of all water wires through a simulation of C1C2 representing a connected component in the graph of all water wires with an occupancy above 50%. Different colors represent different segment ids and edges show the occupancy of a connection.

5.2 The graph viewer

The graph viewer opened by pressing "Draw" in the "2d graph visualization" box comes with some tools for manipulating the view.

Figure 5.3: Toolbar for manipulation of the graph view.

"Home" Reset the view to the initial view.

"Arrows" Redo/Undo a step

"Cross" Click the into the figure to move the view with the cursor.

"Reading-glass" Click and form a rectangle with the cursor to zoom into the rectangle.

"Dots on lines" Manually set borders around the view.

"Increasing arrow" Axis labeling is turned of in the graph browser.

5.3 Plotting and Computations

Bridge plots statistics and values derived from graph properties and the choices are mostly the same for water wires and hydrogen bonds. The plotting dialog comes with a preview panel and the possibility to save the plot or the underlying data:

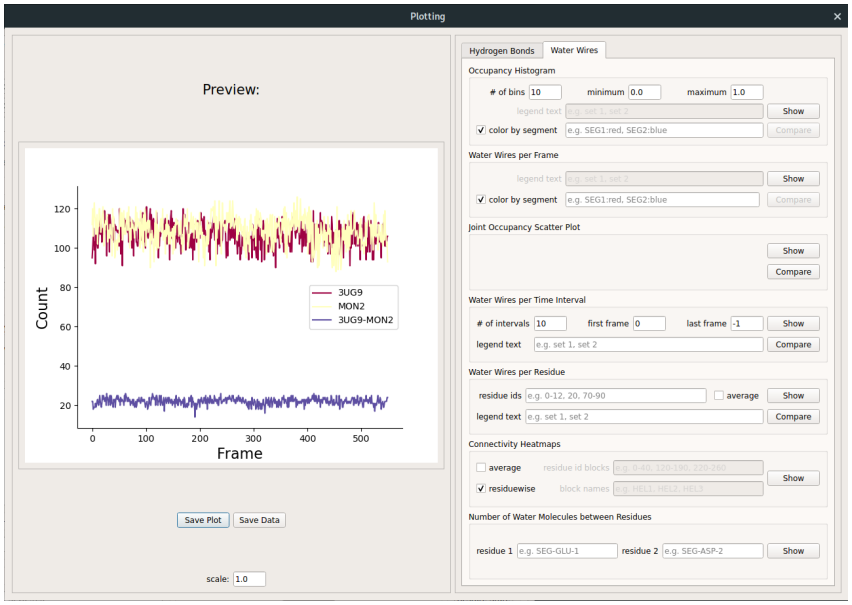

Figure 5.4: Plotting dialog for water wires with a preview panel on the left and all analyses with parameters on the right.

Only the last plot on the paes for hydrogen bonds and water wires differ. The 'Show' buttons show a preview on the left side of the window. If there is a 'Compare' button, the respective plot can be compared to a saved analysis. Clicking it will open a file dialog to find a saved (.hba for hydrogen bond analysis and .wwa for water wire analysis) analysis

file. After selecting the file a preview will be shown in the left panel. The 'Save Plot' button will save the plot in a desired format and the 'Save Data' button will save the underlying data as a text file.

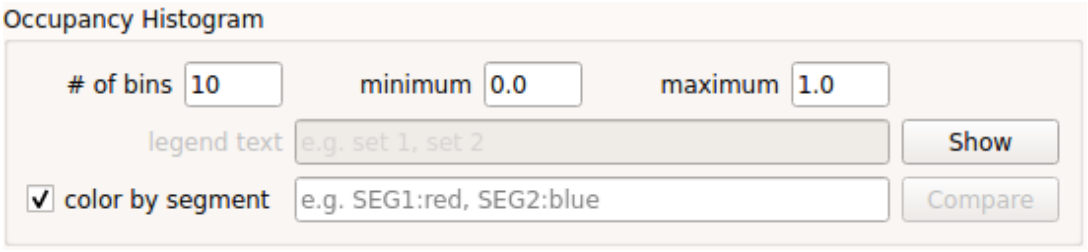

Figure 5.5: Options for the histogram plot.

# of bins The number of bins to be set between 'minimum' and 'maximum'.

minimum and maximum Define the upper and lower bounds of occupancy for the histogram.

legend text Only possible for a 'Compare' plot. Can be used to differentiate between the two states of analysis in a legend.

color by segment If true, the bars of the histogram are colored according to the segment the bonds or wires belong to. While this is true, No 'Compare' plots can be made.

Figure 5.6: Options for the hydrogen bonds per frame plot.

legend text Only possible for a 'Compare' plot. Can be used to differentiate between the two states of analysis in a legend.

color by segment If true, the bars of the histogram are colored according to the segment the bonds or wires belong to. While this is true, No 'Compare' plots can be made.

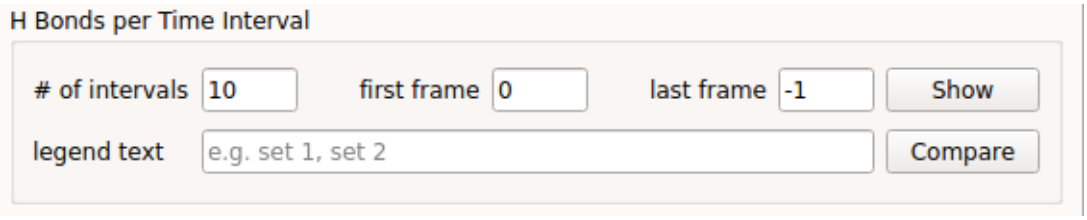

Figure 5.7: Options for the time histogram plot.

# of intervals The number of bins to be set between 'minimum' and 'maximum'.

first and last frame Define the first and last frame for the histogram.

legend text Only possible for a 'Compare' plot. Can be used to differentiate between the two states of analysis in a legend.

Figure 5.8: Options for the hydrogen bonds per residue plot.

residue ids Define the residues that will be plotted. Single numbers or ranges (a pair if residue ids separated by a minus (-)) separated by comma are valid.

average Count the absolute number of connections, if unchecked. Otherwise, the average number over the whole trajectory is counted.

legend text Only possible for a 'Compare' plot. Can be used to differentiate between the two states of analysis in a legend.

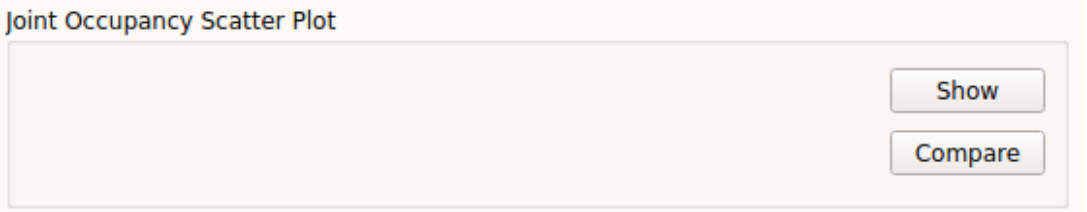

Figure 5.9: Options for the joint occupancy plot.

In the file dialog of the 'Compare' button, multiple other analysis files can be selected by holding down Ctrl while clicking on them. This will compare several joint occupancies in one plot.

Figure 5.10: Options for the heatmap plot.

average Count the absolute number of connections, if unchecked. Otherwise, the average number over the whole trajectory is counted.

residuewise If checked, all hydrogen bonds or water wires between every pair of residues will be shown. Else, the options "residue id blocks" and "block names" have to be specified.

residue id blocks Define ranges of residues (a pair of residue ids separated by a minus (-)). All connections within this range and all connections between the different sets defined by the ranges are shown.

block names There have to be as many names as residue id blocks. Names have to be separated by comma.

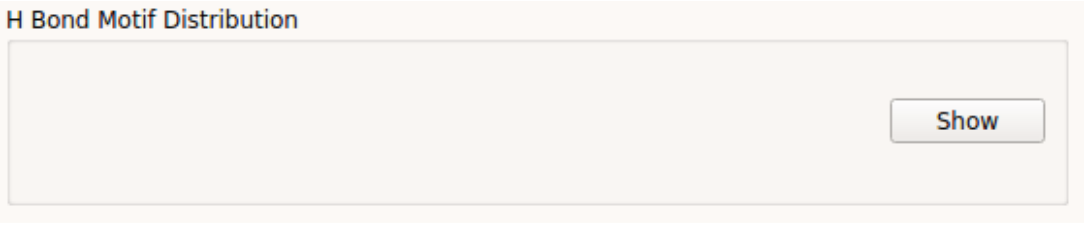

Figure 5.11: Options for the motif distribution plot.

This plot is only possible in a hydrogen bond analysis. The option 'residuewise' has to be unchecked during initialization for this to be enabled. To get intrahelical bonds, backbone atoms have to be specified in the 'add donors' and 'add acceptors' fields before initializing.

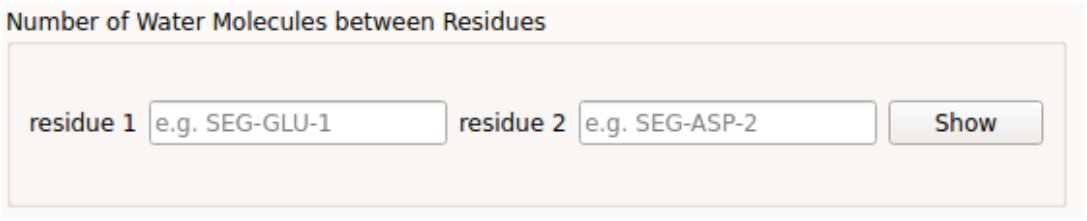

Figure 5.12: Options for the water in wire plot.

This plot is only possible in a water wire analysis.

residue 1 and residue 2 Show the number of water molecules in the water wire connecting residue 1 and residue 2 for every frame.

5.4 Options

The Options section allows saving and restoring of analyses. Saving stores all initial results from the Search section and all applied filters into a file (.hba for hydrogen bond analyses and .wwa for water wire analyses). After restoring such a saved analysis, all functionality is available and all parameters are restored to the interface. Saved analysis files can then be used to compare different analyses with each other from within the plotting dialog or with the graph viewer.

References

Malte Siemers, Michalis Lazaratos, Konstantina Karathanou, Federico Guerra, Leonid Brown, and Ana-Nicoleta Bondar. *Bridge: A graph-based algorithm to analyze dynamic H-bond networks in membrane proteins*, Journal of Chemical Theory and Computation, (2019).