

PROPOSAL: a Java 2D tool for multiple local protein structure alignment

1 Introduction

This document is a complete user guide for PROPOSAL, a Java 2D application for local alignment of two or more protein structures. For the details of the algorithm, please refer to the corresponding paper.

PROPOSAL can quickly perform the alignment task, by setting few parameters, and proposes a user-friendly interface to visualize alignments in 3D space, using the JMol framework (<http://jmol.sourceforge.net/>). Moreover, users can export the alignments in .sif format as 2D contact maps alignment, for the visualization in Cytoscape.

We recommend to run PROPOSAL from command line, instead of opening the .jar file with the Java Platform Binaries, since the button icons used within the program may not be visualized in the second mode. From a command prompt, type:

```
java -jar Proposal_2D.jar
```

Fig. 1 depicts the main frame of PROPOSAL. It contains five panels:

- "Structures list" (Fig. 2, left panel): it contains the list of aligning structures in .pdb format and enables the visualization of one or more structures in 3D using JMol;
- "Log panel" (Fig. 6, center top panel), for logging all the operations performed by the tool;
- "Alignment list" (Fig. 7, center left panel), containing the list of all local alignments found by PROPOSAL;
- "Alignment details" (Fig. 7, center right panel), where the user can obtain more information about a specific alignment, i.e. aligned residues for each structure and final mapping;
- "Parameters" (Fig. 5, right panel), for setting input parameters, check the progress of the alignment task and handle it.

In what follows we details the main features of our tool.

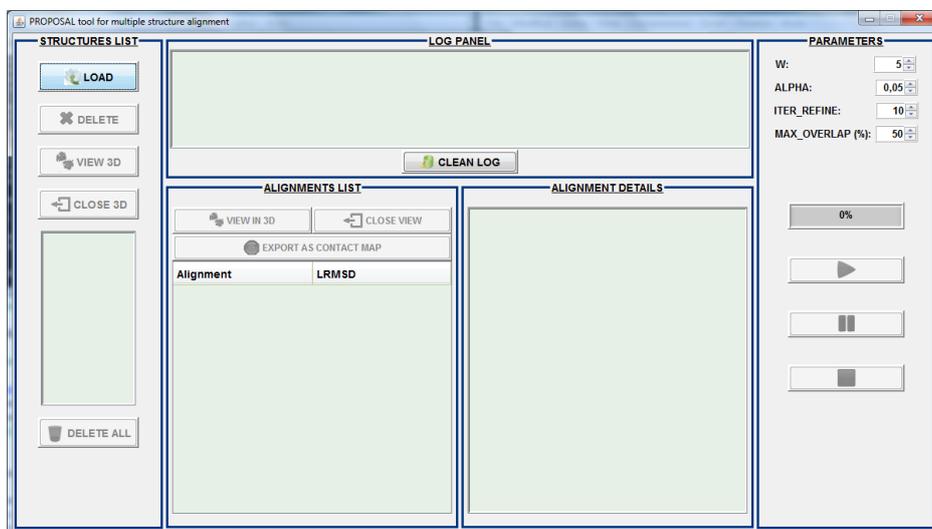


Figure 1: General view of PROPOSAL main frame.

2 Loading protein structures

One or more structures in .pdb format can be loaded by clicking on the "Load" button of "Structures List" panel (Fig. 2). If a protein is already present in the list of loaded structures, it will not be added. A protein structure can be deleted from the list, by selecting it from the list and clicking on "Delete" button. "Delete all" button empties the list and clear all loaded structures.

3 3D visualization of PDB structures

Loaded PDB structures can be visualized in 3D using Jmol, by selecting one or more proteins from the structure list and clicking on "View 3D" button. If at least two structures have been selected, users can choose among two kinds of visualizations:

- "In separate windows" (Fig. 3): selected PDB structures will be visualized in different Jmol panels;
- "In a single window" (Fig. 4): selected PDB structures will be visualized in a single frame, in a grid layout manner, with three Jmol panels for each row.

Different styles and color schemes can be chosen for the visualization of PDB structures. The "Cartoons" style and the "Amino" color scheme are selected by default. For details about the meaning of available styles and colors, please refer to Jmol documentation (<http://jmol.sourceforge.net/docs/>). Clicking on "Close 3D", all currently opened 3D windows will be closed.

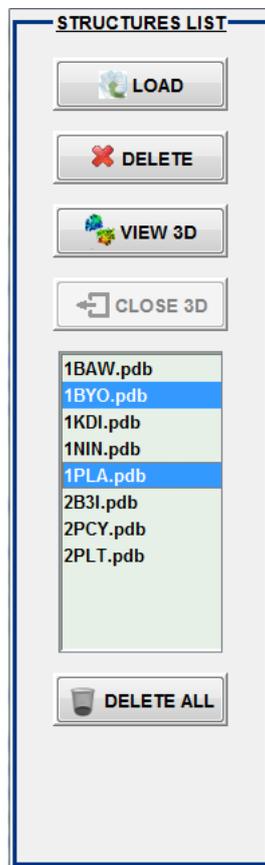


Figure 2: "Structures list" panel.

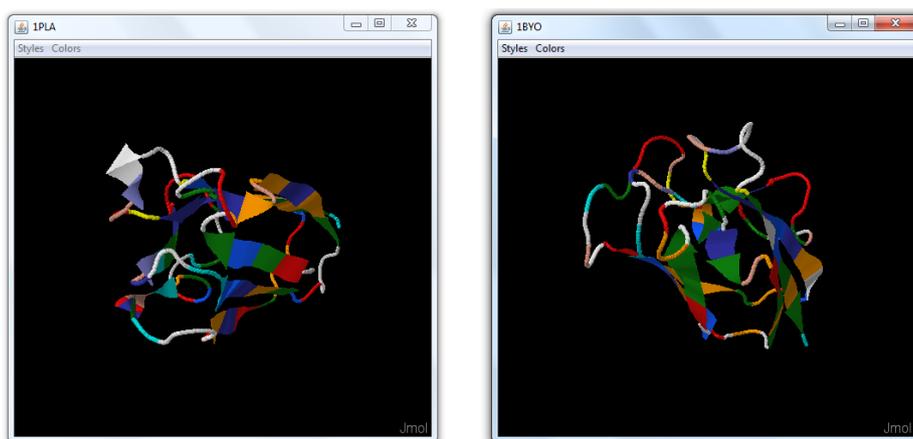


Figure 3: Visualization of PDB structures in separate windows.

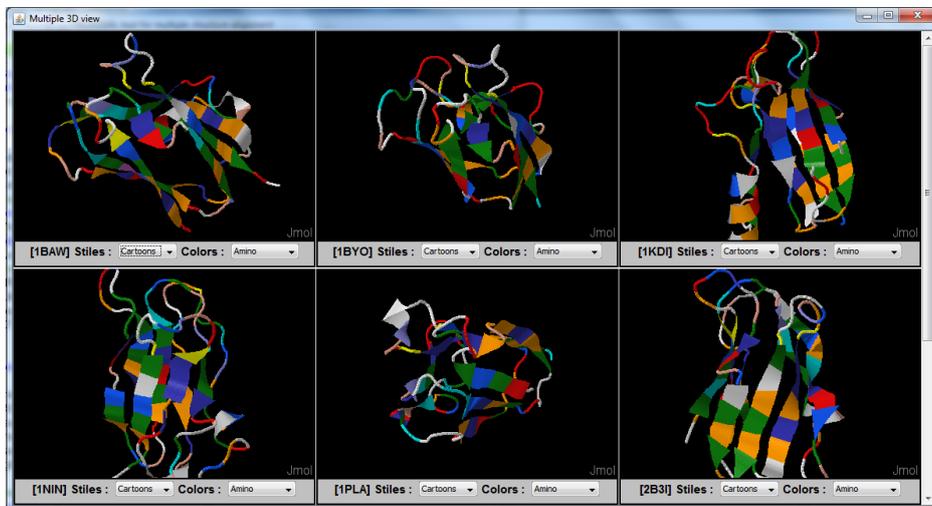


Figure 4: Visualization of PDB structures in a unique window.

4 Setting input parameters

The "Parameters" panel (Fig. 5) enables the setting of all the input parameters of PROPOSAL. By passing the mouse arrow on the name of a parameter, a tool tip string will indicate the meaning of that parameter. We suggest default values for *alpha* and *IterRefine*, in order to guarantee the best tradeoff between speed and accuracy. Anyway, to reach even better accuracy in some cases, the user may choose lower values of *alpha* and higher values of *IterRefine*.

5 Running proposal

After loading PDB structures and setting the input parameters, the user can run PROPOSAL by clicking on the "Play" button, below the progress bar of "Parameters" panel (Fig. 5). In any time, the alignment task can be paused, resumed or canceled by clicking on "Pause", "Play" and "Stop" button, respectively. The progress bar shows the advancement of the alignment task. Other information about the state of the process can be visualized in the "Log Panel" (Fig. 6).

6 Alignments visualization

When PROPOSAL ends, the final list of local alignments found by the algorithm will be visualized in the "Alignments list" panel (Fig. 7). For each alignment, the corresponding average LRMSD (Least Root Mean Square Deviation) across all pairs of aligned structures is reported. By selecting an alignment from the list, more info about the alignment will be visualized in the "Alignment details" panel (Fig. 7). Details include the list of all aligned residues and their 3D coordinates for each structure and the final mapping of the alignment, represented

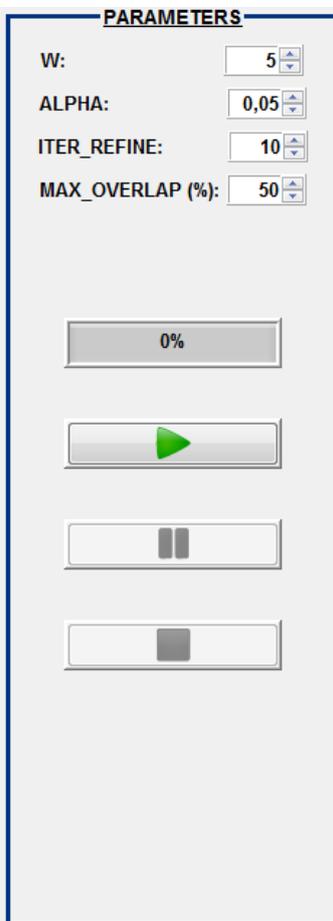


Figure 5: "Parameters" panel.

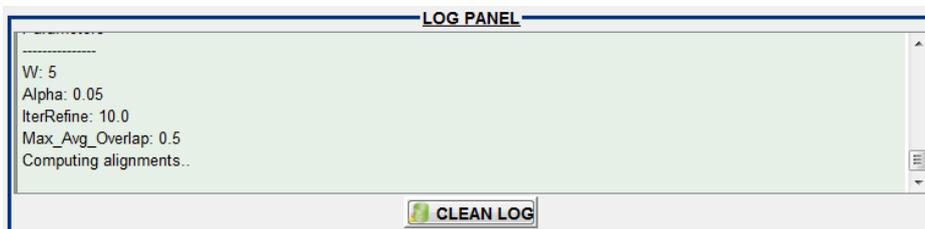


Figure 6: "Log Panel" panel.

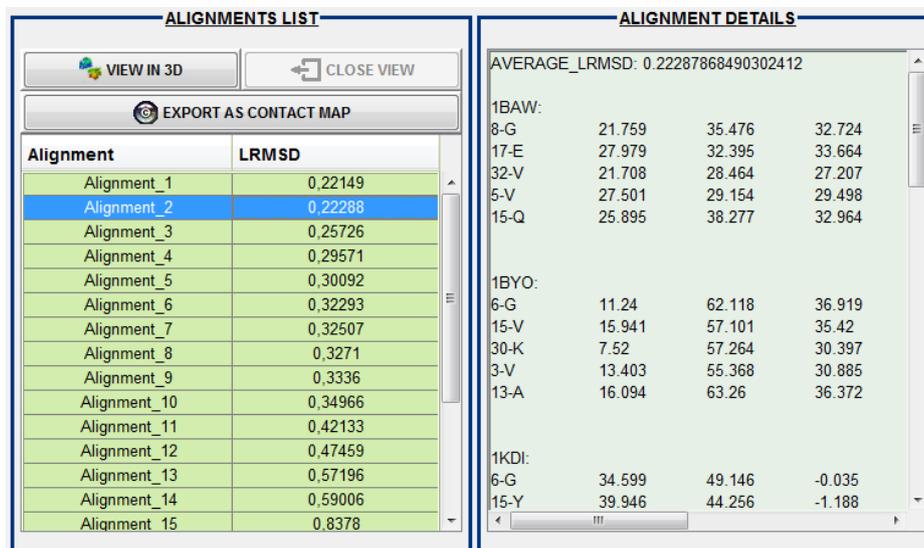


Figure 7: "Alignments list" and "Alignment details" panels.

as a matrix where each column contains all residues of a protein structure and each row contains the residues that match across the different proteins.

Alignments can be further analyzed for visualization. Two kinds of visualization are available:

- As 2D contact maps alignment, by selecting an alignment from the "Alignment list" panel and clicking on "Export as contact map" (Fig. 7);
- As 3D structure alignment in Jmol, by selecting an alignment from the "Alignment list" panel and clicking on "View in 3D" (Fig. 7).

In case a), an alignment graph will be created. The alignment graph is a graph where nodes are all aligned residues and two classes of edges are present:

- Intra-edges, connecting aligned residues of the same protein, that are at distance at most 10\AA one another;
- Inter-edges, connecting aligned residues of different proteins that match one another in the local alignment;

The alignment graph will be then saved in .sif format for the visualization on Cytoscape, where each row has the following format:

```
PDBid1_idRes1 edgeClass PDBid2_idRes2
```

where edgeClass can be "inter" or "intra", "idRes1" and "idRes2" are the ids of aligning residues in the corresponding PDB file and "PDBid1" and "PDBid2" can be equals (i.e. represent the same protein structure) or different.

Together with the .sif file, a node attribute file (in .attrs format) will be created, specifying for each aligned residue the corresponding symbol (in one letter code). Each row of the .attrs file has the following format:

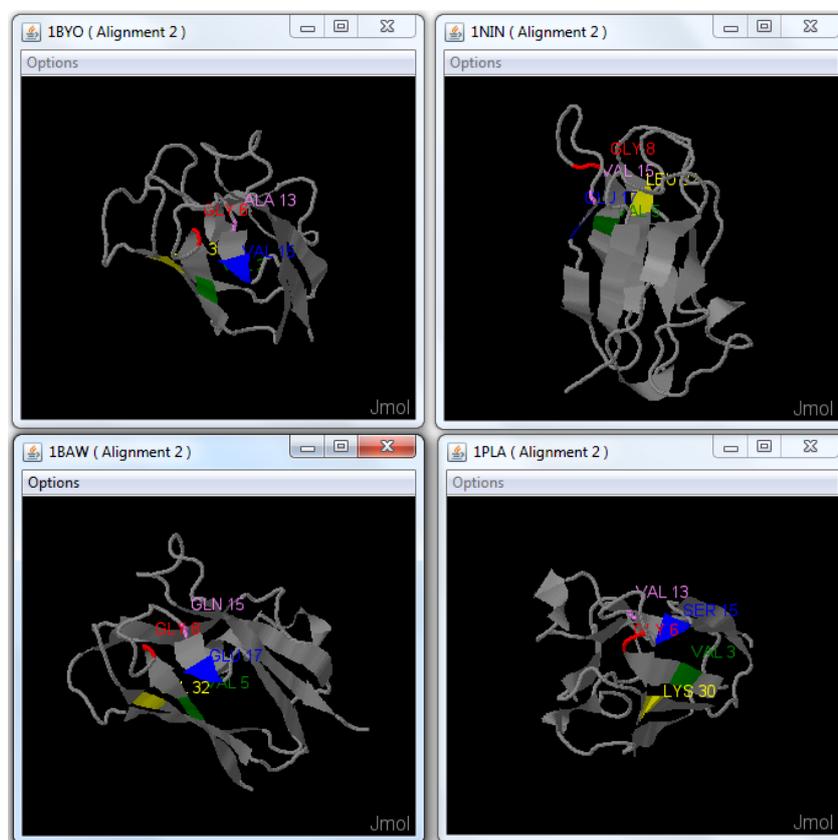


Figure 8: Visualization of a structural local alignment in separate windows.

PDBid_idRes = symbol

By clicking on "Export as contact map", the user can choose a name and a location for the .sif file. The .attrs file will be saved in the same folder with the same name of .sif file, followed by the ".attr" suffix.

In case b) the structures will be visualized in 3D using Jmol. The aligned residues within the structure will be emphasized with different colors and labeled with the corresponding symbol (as a three letter code). Users can choose among two kinds of visualization:

- "In separate windows" (Fig. 8): PDB structures will be visualized in different Jmol panels;
- "In a single window" (Fig. 9): PDB structures will be visualized in a single frame, in a grid layout manner, with three Jmol panels for each row.

Structures will be visualized in the classical "Cartoons" style. Clicking on "Close view" in the "Alignment list" panel, all currently opened 3D windows will be closed.

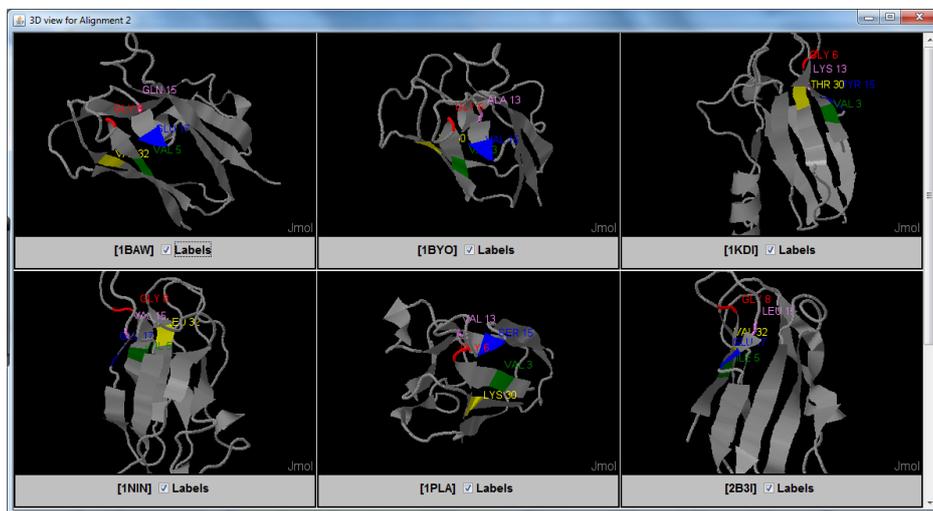


Figure 9: Visualization of a structural local alignment in a unique window.