Contact Map Trajectory Viewing Tool

Corban Rivera, and Alexey Onufriev*

February 14, 2005

*Department of Computer Science, 660 McBryde, Virginia Tech, Blacksburg, VA 24061

Sophisticated *in silico* biological recreation tools like Amber exist to simulate the movement of microscopic biological entities in a controlled medium. These tools construct sub angstrum detail of biological phenominon. However, additional tools are needed to describe the interacting constituants of biological components. Molecular docking and binding drive biological mechanisms. Biologists demand tools which explicitely show the persistant contacts in a simulation.

Keywords: Contact Map, Molecular Trajectory, Folding

The Contact Map Viewer gives scientists and researchers the ability to observe the progression of contacts made during protein folding. As highlights, the viewer allows both distinct trajectory overlays and independent control of contact sequence animation. The Contact Map Viewer differs from other contact mapping tools because it allows multiple contact maps to be viewed simultaneously in the same context.

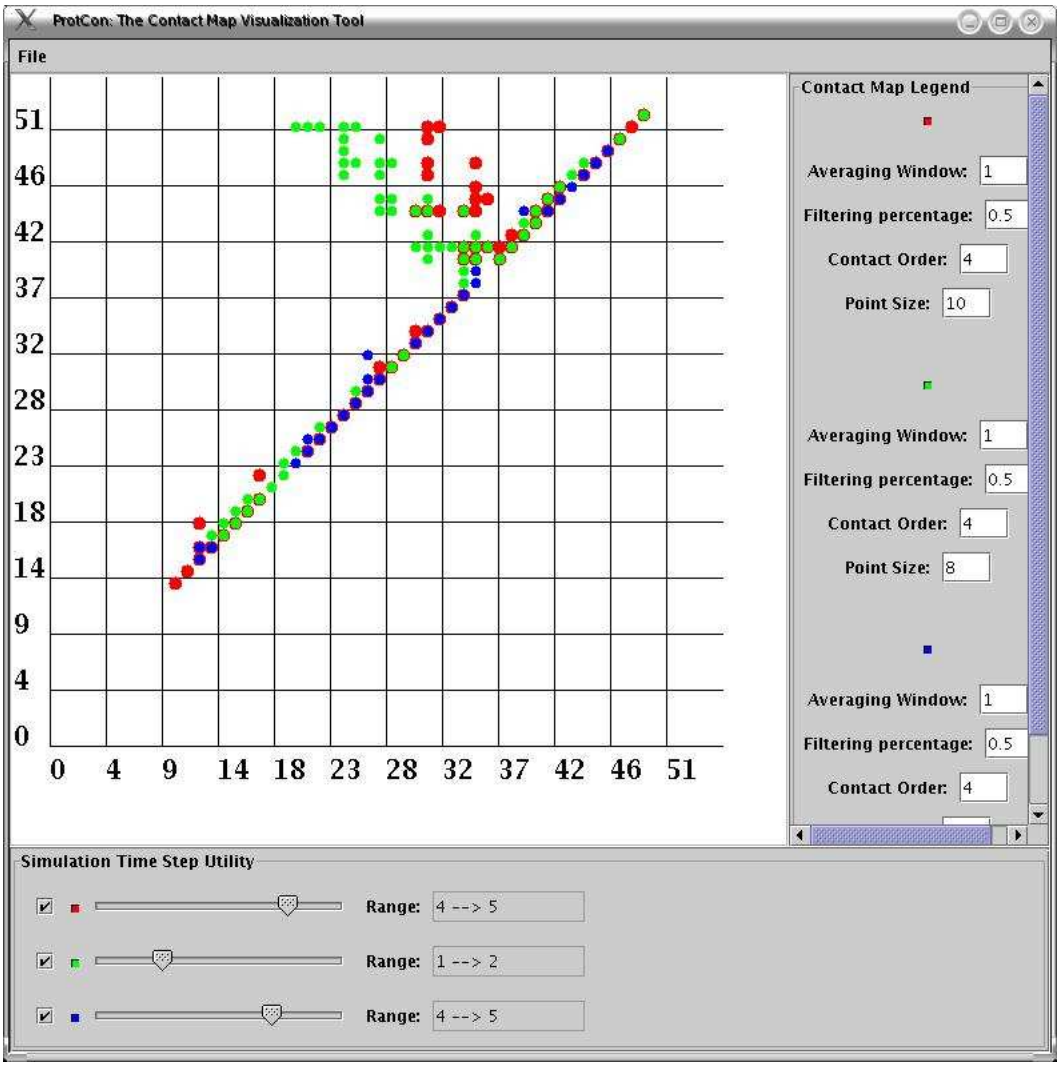

Figure 1: Folding Trajectory Comparison and Visualization Utility

ectionAdding a Trajectory The process for adding a trajectory is now fully incorporated into the GUI. To add a new trajectory to the contact map, sel ect the menu option File-¿Open. A file selection window will appear. The file selection window allows many files to be sel ected at the same time. Use Ctrl or Shift to select all of the PDB files that you would like to be included in a single tra jectory. Please use PDB files with only one model for each file.

Use the File-¿Save option in the menu to open s save file selection window. Select the location where the file is to be sav ed by navagating in the window. Add a file name in the lower text field. Finally, press save to save your contact map in . eps (Extended Post Script) format.

3 Feature Descriptions

A contact map is a graph indicating the residues in contact in a protein. To qualify as a valid contact, an atom from each residue must be the at the proper bond distance. The color of the points corresponds to the color assigned to the trajector y. Grid numbering and spacing is automatically calculated.

When new trajectories are loaded. The size of the contact map will automatically adjust to accomidate the largest protein r esidue count or the largest sequence amino acid count.

3.1 Contact Map Legend

The contact map legend serves as a property list for each trajectory. The trajectories name and color is indicated on the s ame line.

Next, a text field allows the user to select the window size. The window size indicates the number of frames to be included in a contact map. A greater window size allows the user to identify persistent contacts in the protein. To be a non-volati le contact, the contact should appear in many frames in the window under consideration. By default, the window size is only a single frame.

The filtering probability will dictate how in how many frames a contact must occur to be considered in contact. By default, the value is 0.50. The value 0.50 indicates that a contact must appear in greater than half of the frames in the window to be shown. The resonable range for this value is 0 to 1.

The contact order text field allows the user to extract short range contacts from the contact graph. A greater contact dist ance allows users to focus on long range contacts in the protein. Indicating that residues in contact with adjacent residue s are included in the contact map, the default contact distance is 1.

trols the size of points for a entire trajectory. This feature will allow you to compare trajectories to each other simutaniously.

3.2 Controlling the Time Step

The simulation time step utility allows users to step through their trajectory in an animated fashion. For each trajectory, a slide bar gives the user dynamic control over the contact map in view. Allowing the user to adjust contact map for each trajectory independently, the slide bars for each trajectory are unlinked. To the left of the trajectory title, a check box allows the user to hide a particular trajectory from view. Additionally, the label to the right of the slider indicates the trajectory frames currently visible.

4 Usage Scenario

A researcher generated two protein folding trajectories by MD simulation using Amber. One protein folded into the correct topology under permissive experimental conditions; however, the same protein folded incorrectly under high salinity. The researcher will use the Contact Map Viewer to find the root cause for the deviation. The researcher can now use the Contact Map Viewer to examine the two folding trajectories. The researcher finds that an initial missing contact in the salination experiment caused the improper folding.

4.1 Requisite Packages

Java JRE 1.4

4.2 Command Format

java -jar ProtCon.jar

output: folding trajectory visualization in contact maps

compatibility: All platforms

For our purposes, a folding trajectory is physically a directory filled with PDB files recorded at regular intervals of time. The PDB files should contain only a single model. As naming conventions are diverse, we cannot rely upon file naming for time step ordering. However, we attempt to construct an ordering of pdb files by looking for the first number in the file name. Please note that this ordering may not be lexographic.