**Methods**

In the Meshi (Meshi reference) framework I have written a program that finds hydrogen bond pair couples (HBPCs) and creates a weighted graph where the hydrogen bond pairs (HBPs) are the nodes. These nodes are connected by edges if they are coupled and the weights are how many times they are coupled in the database. In addition to the weights, each edge also contains the locations of the hydrogen bonds (HBs), in every protein in the database, involved in the HBPC.

To add more proteins or structures to the graph, I have written a program that adds the information of any new graph derived from the added structures, to the main graph derived from the database. This happens without changing the weights of the main graph since the database is non redundant.

I then use Ctyoscape (Ctyoscape reference) to visualize and analyze this graph of HBPCs. In Cytoscape, I am able to create subgraphs of any protein in the database as well as any CATH representative structure that has been added to the graph. CyToStruct is a plugin for Cytoscape that can open external applications and run scripts while using data from any edge or node but only one at a time. I have written a script for CyToStruct that opens a pdb file in PyMOL that contains the selected edge. The script then colors the residues involved in the HBPC and shows the HBs involved. This plugin cannot use data from multiple edges, so visualizing subgraphs on protein structures was not possible.

In order to view subgraphs on protein structures, I have written a script file for PyMOL that imports a saved subgraph edge table. It then assigns each protein containing an edge in the subgraph, the percentage of the subgraph it contains and sorts them by this value. The script then opens the first protein in this sorted list and colors all the HBPCs in the subgraph.