Weekly Report(Up until 20thth August, 2015)

Anirudh Tiwari

Work Done

Visually analyzed the results of k-means on two domain proteins and compared it with the results from CATH database. Also, worked on calculating the overlap between the output of k-means with CATH.

Results

Following are the results of from my visual analysis of two domain proteins. Kindly have a look at the detailed report sent earlier for the complete analysis with images.

1. The domain boundaries need not contain a secondary structure in its entirety, it is possible that a secondary structure(say a Helix) may be divided amongst the two domains.

2. It is difficult to correctly judge the boundaries of a domain, as many of the domains have a separating boundary at a turn, with part of the turn on one side and the other part on the other.

3. K-means showed some scattered fragments because those chunks of the domains were spatially nearer to the other domain or were surrounded by the other domain.

Next Steps

Compare the performance of k-means with CATH by calculating the overlap between their results.