

Weekly Report(Up until 10th September, 2015)

Anirudh Tiwari

Work Done

1. Last time around I was testing K-Means on only two domain proteins. I created a set of proteins having 1-5 domains and ran my script on it.
2. I was stitching the scattered fragments based on their sequence. Now, I tried stitching based on their distance from the centroids(as calculated by k-means).
3. Partially read Stella's paper.

Results

1. The average overlap reported on testing(using stitching based on sequence) on 1-5 domain proteins is around 70-80%, but since I am experimenting with different methods to stitch the scattered fragments, this comparison is to be done for all the different methods. And only then I can come up with the best overlap score.

Next Steps

1. Try to stitch fragments based on secondary structure information and by calculating their interaction energy with other domains.
2. Complete reading Stella's paper.