Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Material: Clustering Ananlysis

We used clustering analysis tool of GROMACS (g_cluster)[1] to explore the conformation heterogeneity in the ensemble of protein structures generated by computer simulation. The GROMOS clustering algorithm described by Daura *et al*.[2] with a C^{α} RMSD cut-off was used to determine the structurally similar clusters. To choose a reasonable RMSD cut-off, first, we varied the C^{α} RMSD cut-off between 0.10 to 0.15 in steps of 0.01 and performed clustering analysis for each RMSD cut-off value. For 0.10 nm RMSD cut-off, we found three dominant clusters consisting ~50 % of the total protein structures, and for 0.15 nm RMSD cut-off, only a single dominant cluster consisting ~75% of total protein structures. In this work, we chose a C^{α} RMSD cut-off of 0.12 nm. For this choice of cut-off, we found three dominant clusters holding ~73% of total protein structures. Ramachandran maps of the central protein structure (structure with the maximum number of neighbors) of these clusters are shown in **Fig. 1S**. It is evident from the figure that all the three clusters share the characteristics of both PPII helix and β-sheet.

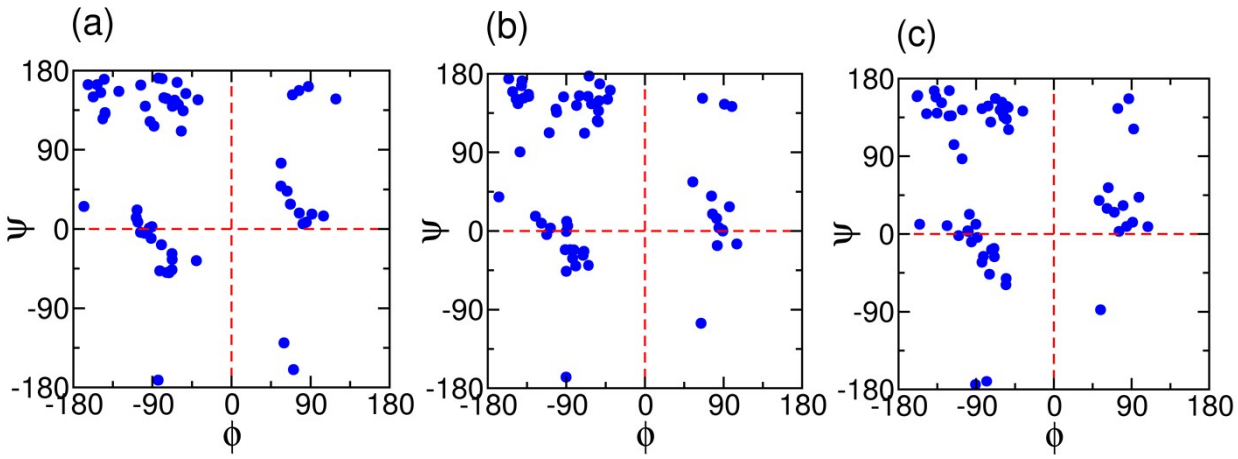

Figure -1S. Ramachandran map of the central protein structure of the (a) largest (b) second largest and (c) third largest clusters. These clusters consist approximately 30, 23 and 20 % of the total protein structures, respectively.

References -

1. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput*. 2008, **4**, 435.
2. X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, *Angew. Chem. Int. Ed. Engl*., 1999, **38**, 236