Dimensional reduction on a molecular dynamics trajectory of a small protein

December 16, 2022

At the link https://drive.google.com/file/d/1DV5w0HZ8F3IU8nY0lgg2HyT5-71dfuVf/view?usp=sharing you will be able to download a molecular dynamics trajectory of the miniprotein CNL025, one of the smallest peptides capable of forming a stable β hairpin. This part of the dynamics explores for most of the time this "folded" state.

1. Extract from the trajectory the coordinates of the Cα atoms, and perform PCA on these coordinates. Plot the spectrum of the eigenvaleus.

2. Align all the configurations to the first configuration of the trajectory, eliminating in this manner rotations and translations. To perform the alignment you can use VMD, or Gromacs, or a tool of your choice. Repeat point 1 with the aligned coordinates. Comment on the difference between the PCA spectra between the two cases. Plot the configurations in the space of the first two principal components. By using a molecular visualization program (for example VMD) visualize the conformational change corresponding to the first principal component.

3. Perform Multi Dimensional Scaling on the same dataset using as distance between the frame the so-called dRMSD. In particular, denoting by x_i^t the coordinates of the Cα atom i in at time t, and by $d_{i,j}^t = ||x_i^t - x_j^t||$ the cartesian distance between atom i and j at time t

$$dRMSD^{t,t'} = \sqrt{\frac{1}{N}\sum_{i,j>i}\left(d_{i,j}^t - d_{i,j}^{t'}\right)^2}$$

   where N is the number of Cα atoms. Also in this case, compute the spectrum, and plot the configurations in the space of the first two principal components.

4. by performing a kernel transformation of dRMSD attempt to find a projection in which the first two components describe the largest possible fraction of the variance.

In order to solve this exercise you are expected to write a code to perform PCA and MDS. You can use an external library for linear algebra operations.