Understanding the Free Energy Landscape of Phase Separation in Lipid Bilayers using Molecular Dynamics

Ashlin J. Poruthoor, Alan Grossfield

February 6, 2023

Simulation details

Auxiliary Variables

For each lipid species, X in the system, we calculated the following,

1. Number of X clusters in the system under study.
2. Fraction of X_i lipids in X clusters
3. Fraction of X_i lipids in X core lipids.
4. Mean Silhouette Coefficient (MSC) of X_i Clusters, as implemented in scikit-learn.

Silhouette Coefficient is a method used to evaluate the clustering done by any technique, especially if ground truth labels are unknown. Here, for a X_i lipid in the cluster, mean intra-cluster distance (a) from other X_i lipids in the cluster is found. Similarly, for a X_i lipid in the cluster, the mean nearest-cluster distance (b) is also calculated. While the former assesses the 'cohesion' of a given X_i lipids with other X_1 lipids in a cluster, the latter assesses the 'separation' from the nearest cluster. Thus, Silhouette Coefficient for a X_i lipid, s, is defined as below,

$$\mathrm{s} = \frac{b - a}{max(a, b)} \tag{1}$$

The Mean Silhouette Coefficient of X_i Clusters is given by the mean s over all non-outlier X_i lipids. Here, we have omitted the MSC calculations for cases when there are no clusters or just one cluster detected by DBSCAN. MSC is bound between and -1 and 1. A high positive value corresponds to well segregated dense clusters, while a low negative value implies that lipids are assigned to clusters incorrectly.