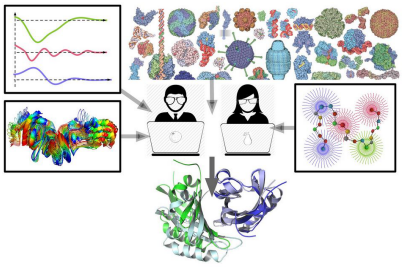

Contribution ID: **10** Type: **Oral**

Clustering algorithms for structural studies: insights on novel metrics and cluster stability assessment

Clustering conformations of molecules or molecular assemblies is a central problem faced in most studies, be they concerned with atomic models or coarse grain models. This talk will review recent contributions in this realm.

The first one is concerned with the newly designed molecular distance RMSDcomb [1]. Given the decomposition of a structure into subdomains, RMSDcomb provides a weighted average of the lRMSD observed between these subdomains, stressing the role of local similarities–as opposed to the lRMSD which suffers from the lack of homogeneity inherent to global structural comparisons.

The second one is concerned with clustering methods providing a simple read out for the number of clusters [2]. In this perspective, we will briefly review density based clustering and the associated stability assessment based on topological persistence.

Finally, we shall discuss the first technique finding a many-to-many correspondence between two sets of clusters delivered by two different methods or the same method under two sets of parameters [3], which is particularly useful to consolidate competing clusterings.

All the methods discussed are provided within the Structural Bioinformatics Library, in the following packages:
https://sbl.inria.fr/doc/Molecular_distances_flexible-user-manual.html
https://sbl.inria.fr/doc/Cluster_engines-user-manual.html#sec-Cluster_engines-programs
https://sbl.inria.fr/doc/D_family_matching-user-manual.html

[1] F. Cazals, T. Dreyfus, D. Mazauric, A. Roth and C.H. Robert. Conformational Ensembles and Sampled Energy Landscapes: Analysis and Comparison, J. Comp. Chem., 36, 2015.

[2] F. Cazals and R. Tetley. Characterizing molecular flexibility by combining lRMSD measures, Proteins, 87, 2019.

[3] F. Cazals, D. Mazauric, R. Tetley and R. Watrigant. Comparing two clusterings using matchings between clusters of clusters, ACM J. of Experimental Algorithms, 24, 2019.

Primary author: Dr CAZALS, Frederic (Inria)

Presenter: Dr CAZALS, Frederic (Inria)

Session Classification: Session 3