Proposition de recherche doctorale

Machine learning for determining continuous conformational transitions of biomolecular complexes from single-particle cryo-electron microscopy images

Mots clés : Array

- Directeur de thèse : Slavica JONIC
- Co-encadrant(s) :
- Unité de recherche : Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie IMPMC
- Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal: Sciences et technologies de l'information et de la communication

Résumé du projet de recherche (Langue 1)

This PhD thesis will be focused on the development of machine learning methods for single-particle cryo-electron microscopy (cryo-EM) image analysis of biomolecular complexes. These methods should allow determining full biomolecular conformational variability and continuous biomolecular conformational transitions with high precision. Characterizing the different conformations that can coexist is essential for understanding how the complexes function and addressing their dynamics. Indeed, high-resolution cryo-EM is currently well established but efficient and user-friendly cryo-EM image analysis methods to fully determine continuous conformational heterogeneity (a continuum of states instead of a few discrete states) still need to be developed (https://doi.org/10.1016/j.sbi.2016.12.011). In this context, we have previously developed a method referred to as HEMNMA, which combines cryo-EM image analysis with molecular mechanics simulation (https://doi.org/10.1016/j.str.2014.01.004). In this PhD thesis, machine learning approaches such as deep learning neural networks (NNs) will be explored for their combination with HEMNMA in order to speed up obtaining the full conformational landscape and processing large data sets, which should allow a more precise determination of paths of continuous conformational transitions. Most often, artificial NNs are used to “predict” a discrete quantity output for a given input such as class labels in the classification problems like the automatic particle identification in cryo-EM images. Here, NNs will be used in the so-called regression problem i.e. to “predict” a continuous quantity output for the given input, which will be a first such approach in the cryo-EM domain. In the same context of continuous conformational heterogeneity analysis from cryo-EM images, besides the supervised learning, non-supervised learning approaches may also be explored. The methods developed during this thesis work will be valorized using experimental cryo-EM data of the human 80S ribosome and the human ATPase p97, in collaboration with Dr Bruno Klaholz, IGBMC, Strasbourg, France (80S ribosome) and Dr Isabelle Rouiller, Université de Melbourne, Melbourne, Australia (ATPase p97).