

Investigating the conformational dynamics of protein-RNA complexes

Contact: yasaman.karami@inria.fr

Contexte et atouts du poste

The main goal of this PhD project is to elucidate the conformational dynamics of protein-RNA complexes and to develop a method for understanding their allosteric communications. To achieve this objective, the candidate will develop a deep learning model to predict communication pathways. They will investigate several facets of the algorithms and ultimately provide this application to the broad scientific community. The proposed deep learning model will be set up to address the characterization of dynamic properties of protein-RNA interactions at the 3D level. This approach is a prerequisite for the design of new therapeutics that target protein-RNA complexes.

The PhD candidate will be hosted in the CAPSID team within LORIA at the Inria, Nancy - Grand Est site. The candidate will be supervised by Yasaman Karami (Inria researcher) with expertise in analyzing proteins conformational dynamics and drug design [1,2], and Malika Smail-Tabonne (associate professor, University of Lorraine) expert in machine learning. CAPSID team (<https://capsid.loria.fr/>) is directed by Dr. Marie-Dominique Devignes and provides a multidisciplinary and international environment for students. The team benefits from experts in RNA-protein interactions and structural bioinformatics, as well as in computer science. The group is equipped with a computational platform, MBI-DS4H (<https://mbi-ds4h.loria.fr/>) composed of 8 nodes and 12 GPUs, and provides technical support to the users.

Mission confiee

Background: Biomolecules such as proteins and nucleic acids are at the heart of virtually all fundamental cellular processes. They adopt complex dynamic behavior and their functions are directly linked to the arrangement of atoms in 3D (structure) and dynamics. These systems undergo conformational changes in response to different environmental conditions, such as mutations, changes in temperature or electrostatic potential, and binding to other molecules. These conditions could induce changes at long distant sites according to the mechanism known as “allostery”. Moreover, such changes may alter the structural plasticity of a biomolecule, induce malfunctioning, thereby provoking diseases. Therefore, characterizing the structure, dynamics and conformational changes of biomolecules can help understand the molecular mechanisms of underlying diseases, and hopefully prevent them by designing drugs. Molecular modeling is the field of characterizing structure of biomolecules. While major improvements have been introduced with the breakthrough of artificial intelligence (AlphaFold2 [3]), predicting conformational dynamics of biomolecules is still challenging. At the same time, molecular dynamics (MD) simulations can provide a powerful, versatile, and accurate computational microscope to study the dynamic behavior of biological systems (Nobel Prize 2013 in chemistry <https://www.nobelprize.org/prizes/chemistry/2013/summary/>). These simulations generate conformational ensembles (trajectories) to capture the motion of systems.

Main task: In this PhD the goal is to design and implement a method for characterizing the conformational dynamics of protein-RNA complexes using the set of trajectories generated from MD simulations in a systematic way. This is of high importance specifically due to the complexity of these systems and large size of produced data from MD simulations. On the one hand, MD simulations have been largely employed to analyse protein-protein complexes, however their application in protein-RNA complexes has yet to be explored. The performance of MD simulations directly depends on the force-field parameters. The recent improvements of such parameters have made them suitable for the analysis of nucleic acids and for characterizing conformational changes in RNA molecules. On the other hand, computational approaches have been developed to investigate the conformational dynamics in proteins [4,5,6]. In the proposed PhD project, the main goal is to adapt such methods to analyse protein-RNA complexes. Moreover, by leveraging the large amounts of data generated by

MD simulations, experimental and computational techniques [3], deep learning can help uncover the underlying mechanisms that govern protein-RNA interactions.

Main activities:

- Literature review of the relevant studies
- Developing a method to extract communication networks in protein-RNA complexes
- Designing a deep learning model
- Implementing the method and preparing a software using Python
- Validating the method and analysing the results
- Writing dissertation, scientific articles and presenting the work in international conferences

Skills

- Master's degree in Computer Science, Bioinformatics, Chemoinformatics or a related master program
- Proficiency in programming languages (Python, PyTorch or R) and good coding practices
- Skills in algorithm design and graph theory
- Experience in machine learning and/or deep learning
- Ability to work independently and also to work in a team
- Excellent oral and written English skills

References

- [1] Temmam S, Vongphayloth K, Baquero E, Munier S, Bonomi M, Regnault B, Douangboubpha B, Karami Y, et al. Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature*. 2022; 604.7905: 330-336.
- [2] Karami Y, Lopez-Castilla A, Ori A, Thomassin JL, Bardiaux B, Malliavin T, Izadi-Pruneyre N, Francetic O, Nilges M. Computational and biochemical analysis of type IV pilus dynamics and stability. *Structure*. 2021; 29.12: 1397-1409.
- [3] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021 Aug 26;596(7873):583-9.
- [4] Gheeraert A, Vuillon L, Chaloin L, Moncorgé O, Very T, Perez S, Leroux V, Chauvot de Beauchêne I, Mias-Lucquin D, Devignes MD, Rivalta I. Singular Interface Dynamics of the SARS-CoV-2 Delta Variant Explained with Contact Perturbation Analysis. *Journal of Chemical Information and Modeling*. 2022 Jun 6;62(12):3107-22.
- [5] Karami Y, Laine E, Carbone A. Dissecting protein architecture with communication blocks and communicating segment pairs. *BMC bioinformatics*. 2016; 17.2: 133-148.
- [6] Karami Y, Bitard-Feildel T, Laine E, Carbone A. "Infostery" analysis of short molecular dynamics simulations identifies highly sensitive residues and predicts deleterious mutations. *Scientific reports* 2018; 8(1): 1-18.