PhD Project

Fragment-based modeling of protein-RNA complexes for protein design

Domains: Molecular modeling, Structural bioinformatics, Computational biology, methods development

Keywords: Combinatorial optimization; protein-RNA interaction; integrative modeling; data-driven docking

Promotor/Supervisor:	Dave Ritchie / Isaure Chauvot de Beauchene
Location:	LORIA (CNRS – INRIA – Lorraine University) at Nancy, France
Starting date:	Between March and December 2019
Duration:	36 months
Funding:	H2020-MSC-ITN project RNAct
Salary:	~2100 euro/month net (+ family allowance if relevant)

Overall aim: The PhD project is part of a multi-disciplinary European ITN project called RNAct, involving 8 inter-connected PhD projects in computational and experimental molecular biology, biophysics and system biology. The overarching aim of RNAct is to integrate experiment and computation to design novel RNA-binding proteins, for synthetic biology and bio-analytics. RNAct will create and characterize novel functional RNA-Recognition motifs in proteins, with customized recognition of specific single-stranded RNA (ssRNA).

Context: The CAPSID team is developing methods to model the 3D structure of biological macromolecules and their assemblages. Docking consists of modeling a molecular assembly from the 3D structure of each molecule, by sampling the possible relative positions of the two molecules, and evaluating the interaction energy to identify the most stable positioning . We have recently developed a method for modeling protein-bound RNAs by **combinatorial assembly** of fragments [see 1, 2, 3, 4, 5].

Aims : The proposed PhD project focuses on two main axis:

- Optimize algorithms for single-stranded **RNA docking methodology**
- Create a RNA-RRM **model generation pipeline**
- Incorporate new sources of information for **data-driven docking**: multiple conformations, low resolution biophysical data, conserved amino-acid – nucleotide contacts [6]
- Assist with developing a **framework of protein design** for RRMs

Strategy : The assembly of poses into chains corresponds to the search of lower energy path in a graph representing the pairwise connectivity of the poses. One optimisation step will be to switch from a brute force and exact search to a probabilistic search, by adding weights to the vertices and edges of the graph.

The PhD student will spend few months of secondments at Dynamic Biosensors GmbH (Germany) to relate in silico work to the experimental determination of RNA-RRM binding with biochips, and to learn about commercial software development.

Eligibility : European funding require an international mobility: The candidates must not have resided or carried out their main activity in France for more than 12 months in the 3 years prior to the recruitment.

Candidates must have a masters degree in any of the relevant disciplines: (bio-)physics, structural biology, bio-informatics or computer science.

The project is highly interdisciplinary: the day-to-day work involves a lot of programming on atomic representations of proteins and nucleic acids. Good programming skills (preferentially Python and/or C++) are essential. Knowledge of structural biology is very desirable, skills in discrete mathematics and statistics would be appreciated. Most importantly, candidates must be motivated to learn about all disciplines relevant to the project.

Candidates must be fluent either in French or in English.

Applications should be sent on the RNAct centralized application web site:
on https://wvranken.wufoo.com/forms/zjepab507unqzd/, select project ESR4