**Contacts**

Javier GONZÁLEZ-DELGADO
ENSAI-CREST, Rennes
javier.gonzalez-delgado@ensai.fr

Juan CORTÉS
LAAS-CNRS, Toulouse
juan.cortes@laas.fr

**M2 biostatistics internship: Statistical analysis of the local structure of flexible RNA molecules**

**Context** For decades, RNA molecules were considered as simple intermediates necessary for the translation of the DNA genetic information into proteins. However, recent works have revealed a surprisingly large range of function for non-coding RNA, fundamentally changing our view on these essential molecules: this includes ribozymes that can catalyze reactions like enzymes, riboswitches that can regulate gene expression through large conformational changes or micro-RNA that provide an extremely important pathway for gene silencing [1]. The diversity of RNA function is achieved by exploiting its complex energy landscape. Understanding those dynamic processes is thus crucial for both the fundamental biophysics of RNA but also to develop rational drug design or bioengineering. In this context, a major open challenge is to characterize the behavior of long flexible or unstructured stretches of RNA, such as large loops involved in key interactions for biological function or messenger RNA strands translated into proteins. However, there is actually a lack of methods to provide realistic atomic-resolution sampling of those flexible regions.

**Objective** The goal of this project is to investigate structural properties of RNA molecules at a local level. The main questions to be answered are: (i) is the conformation of a nucleotide dependent on the neighbors? (ii) are there correlations between the backbone dihedral angles and the sugar ring puckering or with the glycosidic angle connecting the sugar to the base? (iii) and if so, how to characterize these correlations? During the first part of the project, we will review the literature looking for existing answers to these questions. Then, we will perform our own statistical analyses from a large database of experimentally-determined high-resolution RNA structures. The overall results of this work will be made available and presented to the scientific community.

**Methods** The analyses will be based on statistical tests of independence and goodness-of-fit. The complexity of the data, defined in non-Euclidean spaces, will offer an opportunity for theoretical work involving the adaptation of existing methods or the development of new techniques, potentially useful from a more general perspective.

**Location** The internship will be co-supervised by Javier González-Delgado and Juan Cortés, and will take place primarily at ENSAI, with the possibility of visits or a short stay at LAAS. The project is part of an ongoing collaboration with Loïc Salmon (CRMN-CNRS) and Isaure Chauvot de Beauchene (LORIA-CNRS).

**References**

[1] L. R. Ganser, M. L. Kelly, D. Herschlag, H. M. Al-Hashimi, (2019). The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Bio*, 20: 474–489.