



université PARIS-SACLAY

# PRÉDICTION DE STRUCTURE TRIDIMENSIONNELLE DE MOLÉCULES D' ARN PAR MINIMISATION DE REGRET PAR MÉLANIE BOUDARD

**Présentée par : Melanie Boudard Discipline : informatique Laboratoire : DAVID**

## Résumé :

Les fonctions d'une molécule d'ARN dans les processus cellulaires sont très étroitement liées à sa structure tridimensionnelle. Il est donc essentiel de pouvoir prédire cette structure pour étudier sa fonction. Le repliement de l'ARN peut être vu comme un processus en deux étapes : le repliement en structure secondaire, grâce à des interactions fortes, puis le repliement en structure tridimensionnelle par des interactions tertiaires. Prédire la structure secondaire a donné lieu à de nombreuses avancées depuis plus de trente ans. Toutefois, la prédiction de la structure tridimensionnelle est un problème bien plus difficile. Nous nous intéressons ici au problème de prédiction de la structure 3D d'ARN sous la forme d'un jeu. Nous représentons la structure secondaire de l'ARN comme un graphe : cela correspond à une modélisation à gros grain de cette structure. Cette modélisation permet de réaliser un jeu de repliement dans l'espace. Notre hypothèse consiste à voir la structure 3D comme un équilibre en théorie des jeux. Pour atteindre cet équilibre, nous utiliserons des algorithmes de minimisation de regret.

Nous étudierons aussi différentes formalisations du jeu, basées sur des statistiques biologiques. L'objectif de ce travail est de développer une méthode de repliement d'ARN fonctionnant sur tous les types de molécule d'ARN et obtenant des structures similaires aux molécules réelles. Notre méthode, nommée GARN, a atteint les objectifs attendus et nous a permis d'approfondir l'impact de certains paramètres pour la prédiction de structure à gros grain des molécules.

#### Abstract :

The functions of RNA molecules in cellular processes are related very closely to its three dimensional structure. It is thus essential to predict the structure for understanding RNA functions. This folding can be seen as a two-step process: the formation of a secondary structure and the formation of three-dimensional structure. This first step is the results of strong interactions between nucleotides, and the second one is obtain by the tertiary interactions. Predicting the secondary structure is well-known and results in numerous advances since thirty years. However, predicting the three-dimensional structure is a more difficult problem due to the high number of possibility. To overcome this problem, we decided to see the folding of the RNA structure as a game. The secondary structure of the RNA is represented as a graph: its corresponds to a coarse-grained modeling of this structure. This modeling allows us to fold the RNA molecule in a discrete space. Our hypothesis is to understand the 3D structure like an equilibrium in game theory. To find this equilibrium, we will use regret minimization algorithms. We also study different formalizations of the game, based on biological statistics. The objective of this work is to develop a method of RNA folding which will work on all types of secondary structures and results more accurate than current approaches. Our method, called GARN, reached the expected objectives and allowed us to deepen the interesting factors for coarse-grained structure prediction on molecules.

## INFORMATIONS COMPLÉMENTAIRES

**Christophe DÜRR**, Directeur de Recherche, à l'Université Pierre et Marie Curie /Laboratoire d'Informatique de Paris 6 (LIP 6) - CNRS 7606 - Paris - Rapporteur

**Jérôme WALDISPÜHL**, Professeur à l'Université Mc Gill/Ecole d'Informatique - Montréal (Canada) - Rapporteur

**Johanne COHEN**, Chargée de Recherche, Habilitée à Diriger des Recherches, à l'Université Paris Sud 11/Laboratoire de Recherche en Informatique (LRI) - UMR 8623 - Orsay - Directrice de thèse

**Dominique BARTH**, Professeur des Universités, à l'Université de Versailles Saint-Quentin-en-Yvelines/Laboratoire Données et Algorithmes pour une Ville Intelligente et

Durable de l'infrastructure à l'individu (DAVID) - Versailles - Co-Directeur de thèse  
**Alain DENISE**, Professeur des Universités, à l'Université Paris Sud 11/Laboratoire de  
Recherche en Informatique (LRI) - UMR 8623 - Orsay - Co-Directeur de thèse  
**Raphaël GUEROIS**, Chercheur, au CEA/Institut de Biologie Intégrative de la Cellule  
(I2BC) - Gif/Yvette - Examineur

**Alessandra CARBONE**, Professeure des Universités, à l'Université Pierre et Marie  
Curie/Laboratoire de Biologie Computationnelle et Quantitative (CQB) - UMR 7238 -  
Paris - Examineur

**Stefano MORETTI**, Chargé de Recherche, à l'Université Paris Dauphine/Laboratoire d'  
Analyse et Modélisation de Systèmes pour l'Aide à la Décision (LAMSADE) - UMR 7243  
- Paris - Examineur