PhD PROJECT PROPOSAL - Computational and Quantitative Biology

RHAPSODY

RNA Hybrid AI Platform for Structural Observables and Dynamics

Principal Investigator: Dr Davide Mercadante
Host Institute: Dipartimento di Farmacia, Universitá Federico II

Duration: 3 years (fully funded)

Background to the Proposed Research

RNA molecules play fundamental roles in biology, regulating gene expression while controlling numerous cellular processes. Unlike DNA, RNA is highly dynamic, constantly interconverting between multiple conformational states underpinning its biological function. Capturing these structural states remains one of the major challenges in understanding and designing RNA molecules with specific therapeutic functions. Molecular (MD) simulations provide an invaluable tool for studying RNA dynamics, but atomistic simulations are often too computationally expensive to reach the timescales required to observe biologically relevant conformational changes. Coarse-grained (CG) simulations overcome this limitation by grouping atoms into spheres and enabling simulations to be orders of magnitude faster. However, this increased efficiency comes at the cost of reduced resolution. Recently, in collaboration with the De Simone's group we developed the first machine learning framework capable of predicting atomistic NMR chemical shifts directly from coarse-grained protein simulations, allowing experimental NMR data to guide molecular simulations while retaining the efficiency of coarse-grained models [1, 2]. This project will extend that paradigm to RNA, creating the first differentiable framework that directly links coarse-grained RNA simulations with experimentally measured NMR observables.

Proposed Research

The aim of this project is to develop a new generation of AI models capable of translating coarse-grained RNA conformations into experimentally measurable NMR observables[3]. By combining molecular simulations, neural networks and statistical learning, the project will enable RNA simulations to be directly restrained by experimental data while preserving the computational efficiency of coarse-grained models. The research will focus initially on NMR chemical shifts, one of the richest sources of structural information available for RNA. The developed methodology will then be expanded to incorporate additional NMR observables, creating a unified framework for experimentally driven coarse-grained RNA simulations. Ultimately, the project will establish a powerful computational platform capable of studying RNA structural ensembles across timescales that remain inaccessible to conventional atomistic simulations.

Research tasks and proposed timeline

Year 1: developing ML-based models for RNA chemical shift prediction.

Develop graph neural network architectures capable of predicting atomistic RNA chemical shifts directly from coarse-grained models. Curate large structural and NMR datasets, investigate suitable molecular representations (coarse-grained mapping), and benchmark different machine learning architectures against existing atomistic prediction methods.

Year 2: integrating experimental restraints into molecular simulations.

Integrate the learned predictor into differentiable molecular dynamics simulations, allowing experimental observables to bias coarse-grained RNA simulations. Develop differentiable restraint functions and validate the methodology using RNA systems with experimentally determined structures and NMR datasets.

Year 3: moving towards a unified framework for RNA structural refinement.

Extend the framework beyond chemical shifts by incorporating additional NMR observables. Apply the methodology to biologically important RNAs exhibiting substantial conformational heterogeneity and benchmark its performance against conventional atomistic simulation approaches.

Preferred Background and Skillset of the Candidate

This project sits at the interface of structural biology, molecular simulation and artificial intelligence. The successful candidate will develop expertise in coarse-grained modelling, molecular dynamics simulations, machine learning, scientific programming and computational biophysics while working closely with collaborators in biomolecular NMR spectroscopy. Suitable applicants may come from a range of quantitative disciplines, including chemistry, physics, computer science, or mathematics.

Applicants with a stronger chemistry, biology or physics background should ideally have:

- strong quantitative and analytical skills.
- programming experience (preferably Python).
- an interest in molecular simulations and biomolecular structure.
- enthusiasm and committment in learning modern machine learning techniques.

Applicants with a stronger computer science or mathematical background should ideally have:

- strong programming (C++ and CUDA preferred) and algorithmic skills.
- an interest in applying artificial intelligence to biological problems.
- willingness to develop knowledge in molecular biophysics and structural biology.
- willingness to work at the interface between disciplines.

Project's wider relevance

Artificial intelligence is transforming structural biology, from protein structure prediction to biomolecular design. However, the integration of experimental data into coarse-grained molecular simulations remains largely unexplored, particularly for RNA or generally for nucleic acids. This project builds directly upon our group's recent breakthrough in developing differentiable ML-trained models that connect coarse-grained protein simulations with experimental NMR measurements. By extending these ideas to RNA, the successful candidate will contribute to the development of an entirely new computational methodology with broad applications in RNA biology, biomolecular engineering and integrative structural biology. The project combines cutting-edge machine learning, molecular simulations and experimental structural biology, providing opportunities to collaborate with experts in NMR spectroscopy, CG modelling and RNA structural biology. The resulting software will enable researchers worldwide to perform experimentally guided coarse-grained simulations of RNA, for the first time.

References

[1] M. Cullen, C. Biancaniello, K. Taškova, V. Miletić, D. Mercadante*, and A. De Simone*. "Integrating NMR Restraints into Coarse-Grained Simulations: Toward Accurate Conformational Ensembles of Complex Protein Systems". In: *Journal of the American Chemical Society* 148.12 (2026).

[2] M. Cullen, C. Biancaniello, K. Taškova, V. Miletić, A. D. Simone, and D. Mercadante*. "Capturing secondary structure in coarse grained intrinsically disordered proteins with simulations driven by chemical shifts". In: *bioRxiv.org* (2026). eprint: `https://www.biorxiv.org/content/early/2026/01/06/2026.01.06.697719.full.pdf`.

[3] R. P. Barnwal, F. Yang, and G. Varani. "Applications of NMR to structure determination of RNAs large and small". In: *Archives of Biochemistry and Biophysics* 628 (2017).