Molecular mechanisms of riboswitch activation using integrative, AI-guided molecular simulations

Riboswitches are regulatory elements of messenger RNA (mRNA) that control gene expression by binding to specific ligands, triggering large conformational changes that lead to their activation. Understanding the molecular mechanisms of riboswitch activation and developing rational strategies to modulate them require an atomistic description of how ligands reshape the complex conformational landscapes of these dynamic RNAs. Neither experimental techniques nor computational approaches alone can provide this information. In this project, we will develop an integrative computational-experimental approach to characterize the conformational changes that lead to the activation of a prototypical preQ$_1$ riboswitch upon ligand binding. Our approach integrates atomistic molecular simulations, enhanced by AI-guided sampling techniques, with solution experimental measurements to: (i) determine accurate conformational ensembles of preQ$_1$ in its free and bound states; (ii) elucidate the molecular mechanisms of activation; (iii) characterize variations across different species; and (iv) design novel preQ$_1$ inhibitors.

Context and state-of-the-art. Several X-ray crystallography and NMR structures of the preQ$_1$ riboswitch have been determined in its free form (*apo*) and bound to the cognate ligand preQ$_1$ (*holo*). These structures represent only static snapshots of complex conformational landscapes and therefore they do not really elucidate the mechanisms of ligand activation. Solution experiments like small angle X-ray scattering (SAXS), Förster resonance energy transfer (FRET) and chemical probing data (SHAPE), can, in principle, provide information on the entire conformational landscape of dynamic RNAs, although often at low resolution or as an ensemble-averaged measurement. On the other hand, *in silico* approaches like Molecular Dynamics (MD) can determine conformational landscapes of dynamic biological systems at atomistic resolution. However, their quality is limited by the simulation timescales accessible with current hardware and the accuracy of the underlying interatomic potentials (force field) used for sampling. Over the last decade, major advancements in AI techniques for accelerating MD, as well as integrative approaches for refining *in silico* ensembles with solution experimental data, have helped address these two challenges. The groups led by G. Stirnemann (CPCV lab, SU/ENS-PSL/CNRS) and M. Bonomi at Institut Pasteur have both made substantial contributions to the development and dissemination of these advanced methods [1-4].

Objectives. The main objectives of this project are to elucidate, at the atomistic level, the molecular mechanisms by which the preQ$_1$ riboswitch is activated by its cognate ligand, preQ$_1$; to determine how these mechanisms vary across species; and to leverage the structural and dynamic characterization of preQ$_1$ in its *apo* state to design novel compounds that inhibit riboswitch activation.

Work plan. The PhD will be organized around 5 aims.

Aim 1. Determination of structural ensembles of preQ$_1$ in apo form and bound to its cognate ligand. Atomistic, explicit-solvent MD simulations of preQ$_1$ from *Thermoanaerobacter tengcongensis* (Tte-preQ$_1$) will be initiated from the X-ray structures in *apo* form (PDB 6vuh) and bound to preQ$_1$ (PDB code 6vui). State-of-the art force fields for RNA, water, and ions in solution will be used. Sampling will be accelerated using Hamiltonian replica-exchange simulations augmented with generative AI models [1, 2, 5]. To address residual force field inaccuracies, the resulting ensembles will be refined using a MaxEnt reweighing approach [3] incorporating FRET [6] and SAXS [7] data. This refinement procedure will minimally adjust the *in silico* MD ensemble populations to improve agreement with solution experimental data (**Figure 1**). The refined ensembles will be validated using SHAPE data [8].

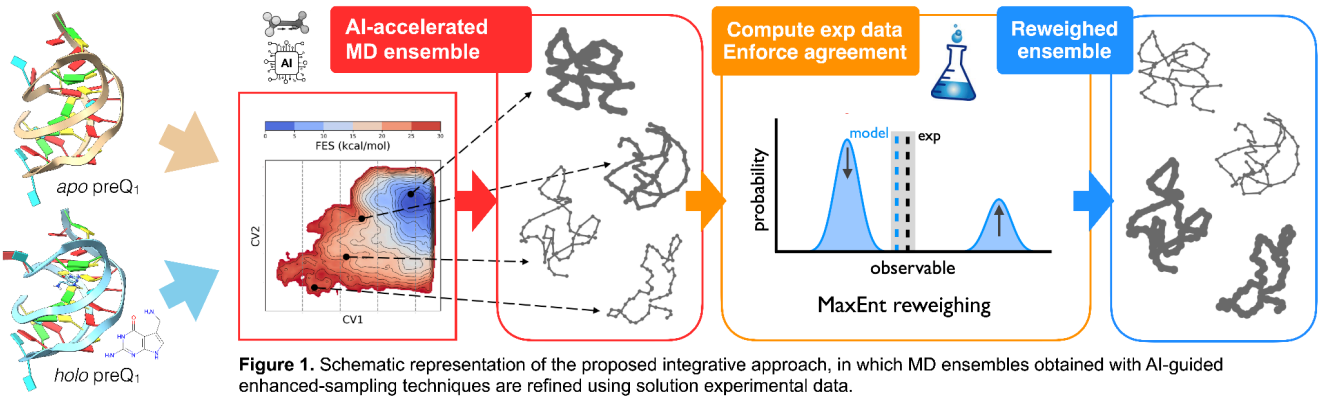

Figure 1. Schematic representation of the proposed integrative approach, in which MD ensembles obtained with AI-guided enhanced-sampling techniques are refined using solution experimental data.

Aim 2. Analysis of molecular mechanisms of preQ$_1$ activation upon ligand binding. Comparison of the *apo* and *holo* ensembles determined in **Aim 1** will provide insight into the mechanism of preQ$_1$ activation. Different dimensionality reduction techniques mostly based on AI (variational autoencoders, information bottleneck) and RNA-specific structural metrics, such as eRMSD, will be used to classify the large conformational space sampled in our simulations. Notably, these techniques will help determine whether preQ$_1$ binds to a pre-formed *holo*-like structure via a conformational selection mechanism or, alternatively, induces significant conformational changes upon binding to an intermediate state (induced-fit mechanism).

Aim 3. Molecular mechanisms of preQ$_1$ activation across different species. Structural ensembles of preQ$_1$ from *Bacillus subtilis* (Bsu-preQ$_1$) in both *apo* and *holo* forms will be determined using the approach introduced in **Aim 1** and analyzed following the protocol described in **Aim 2**. Comparison with the results obtained for Tte-preQ$_1$ will provide insights into the differences in activation mechanisms across species.

Aim 4. Identification of novel preQ$_1$ inhibitors. The integrative ensembles of *apo* preQ$_1$ determined in **Aim 1/3** provide an opportunity to design novel compounds that inhibit riboswitch activation by binding to and stabilizing non-active conformations. To explore this, we will use SHAMAN, a small-molecule binding site detector for RNA ensembles recently developed in the Bonomi lab [9]. Multiple SHAMAN runs will be performed starting from significantly populated, non-active-like states determined in **Aim 1/3**. The binding sites detected by SHAMAN will guide ensemble virtual screening campaigns to identify potential inhibitors of preQ$_1$ activation. Our predictions will be tested by the "Chemogenomic and Biological Screening Platform" at Institut Pasteur.

Aim 5. Methods dissemination, data sharing, and educational material. All the methods developed in this project will be distributed to the community in the open-source PLUMED library (www.plumed.org) [10], of which M. Bonomi is founder and core-developer. All the data needed to reproduce our study will be deposited in PLUMED-NEST (www.plumed-nest.org). Tutorials to train researchers will be shared in PLUMED-TUTORIALS (www.plumed-tutorials.org) [11], an educational resource co-created by the two PIs.

Outcomes and impact. This project will not only provide atomic-level insights into the activation mechanism of the preQ$_1$ riboswitch but will also establish a versatile modeling framework for studying other dynamic riboswitches and RNA molecules. Additionally, our structure-based ligand discovery pipeline could be expanded into a general strategy for identifying novel compounds that inhibit riboswitch activation in bacteria, offering a promising approach to fight antibiotic resistance in bacterial pathogens. All results will be published in high-impact open-access journals and presented at national and international conferences.

Risks and mitigation. The main challenges of this project are (i) achieving an exhaustive exploration of the preQ$_1$ conformational space and (ii) obtaining conformational ensembles that are minimally impacted by the inaccuracies of the molecular mechanics force field used. The two PIs bring complementary, recognized expertise in developing and applying advanced simulation techniques, based on AI-guided enhanced sampling and molecular simulations of reactivity (Stirnemann [1, 2, 12]) and integration with experimental data (Bonomi [3, 4]), which will mitigate these issues. Additionally, the computational resources available in both host labs will contribute to the successful execution of this proposal.

Research environment, thesis supervision, and complementary expertise. The PhD student will conduct their research in a scientifically-vibrant atmosphere in two leading institutions located in the heart of Paris. These establishments offer state-of-the-art computational and experimental resources, which will be instrumental in the successful realization of this proposal. The PhD student will be directed by G. Stirnemann and co-supervised by M. Bonomi. Both PIs and their labs have proven expertise, complementarity, and international recognition in the fields of computational structural biology and chemistry, as demonstrated by the two ERC grants recently awarded to the PIs. Finally, both labs are committed to cultivating a diverse, inclusive, and internationally collaborative research environment.

[1] Forget, …, **Stirnemann** J Chem Theory Comput 20 (2024) 6263

[2] Benayad, **Stirnemann** Biophys J 123 (2024) 296a

[3] Borthakur, …, **Bonomi**, Robustell bioRxiv (2024)

[4] **Bonomi** *et al.* Sci Adv 2 (2016) e1501177

[5] Gabrié *et al.* PNAS 119 (2022) e2109420119

[6] Suddala *et al.* NAR 41 (2013) 10462

[7] Jenkins *et al.* J Biol Chem 286 (2011) 24626

[8] Wang *et al.* ACS Cent Sci 8 (2022) 741

[9] Panei, Gkeka, **Bonomi** Nat Commun 15 (2024) 5725

[10] **Bonomi** *et al.* Nat Methods 16 (2019) 670

[11] Tribello, **Bonomi**, …, **Stirnemann** *et al.* J Chem Phys (2025)

[12] Benayad, …, **Stirnemann** PNAS 121 (2024) e2322040121