Responsable du stage

- **Coordinateur du stage: Sergei Grudinin + Clovis Galiez**
- Nom du laboratoire et de l'équipe de recherche: LJK CNRS (DAO / SVH)
- Email : Sergei Grudinin (sergei.grudinin@univ-grenoble-alpes.fr), Clovis Galiez (clovis.galiez@univ-grenoble-alpes.fr)

Objet du stage

- **Intitulé du stage : Generative models for pore-forming antimicrobial peptides**
- **Durée du stage: 6m**
- **Lieu du stage : LJK, IMAG, Grenoble Alpes University, 38401 St Martin d'Hères (Grenoble)**
- **Résumé du stage:**

Introduction:

A promising strategy to fight against antimicrobial resistance can be designing novel antimicrobial peptides (AMPs). These are short amino acid sequences (typically ranging from 6 to 50 residues in length) that kill various bacteria, viruses and fungi mainly through membrane disruption, but also a few other mechanisms. Recently, protein design (proteins are long peptides, longer than 50 residues) has been transformed by AI tools such as RF-Diffusion, ProteinMPNN, and AlphaFold (awarded the 2024 Nobel Prize in Chemistry). Building on these advances, this project aims to transfer the protein-specific AI developments to peptides by reusing structural and sequence patterns in proteins with tandem repeats[9].

Methods:

We will use precomputed data from TMVisDB, AlphaFoldDB, and UniProt databases. TMVisDB contains transmembrane region predictions by the TMbed algorithm for approximately 46 million proteins. These proteins are all the proteins predicted to have transmembrane properties from the AlphaFoldDB database (~200 million proteins), as accessed in July 2026. Protein structures are sourced from the AlphaFoldDB database, which contains more than 200 million structures predicted by the AlphaFold algorithm, based on sequences from the UniProt database. Our main training data will be a dataset of α-helical proteins with annotated repeated transmembrane segments.

Then, the goal of the trainee we be to extend the proteinMPNN algorithm (both encoder and decoder parts) to the sequences with internal non-exact repeats. Several methods will be tested, based on the MSA transformer, structural repeats, and different graph representations of the initial data. Successful designs will be tested in-silico by AlphaFold refoldability scores and then, if successful, synthesized and tested in cell and liposome models by the DPM project partner.

Requirements:

We are looking for creative, passionate and hard-working individuals (of M1 or M2 level) from applied math / computer science background with exceptional talent for computer science and mathematics and interest in computational biology / chemistry / physics. Excellent oral, written and interpersonal communication skills are essential (working language will be English – knowledge of French is a plus). Good knowledge of PyTorch / C++ / signal processing / machine learning / structural bioinformatics will be an asset. The trainee will implement different architectures using PyTorch and test them on a local GPU-accelerated cluster.

Remuneration:

€4.50/hour i.e., approximately €630/month for a 35-hour week.

To learn more about the internship setting:

Nestled in the heart of the French Alps, **Grenoble** is home to several renowned universities and research institutes, including Université Grenoble Alpes. In addition to its academic offerings, Grenoble is surrounded by magnificent natural landscapes, with easy access to the mountains for skiing, hiking, and other outdoor activities. Grenoble has a diverse and international population.

Université Grenoble Alpes (UGA) has 60,000 students, 6,000 researchers, and 3,000 doctoral candidates. It consistently ranks in the 100–250 bracket of the world's best universities in major international rankings and is among the top 10 French higher education institutions.

UGA comprises 71 research units, including the Jean Kuntzmann Laboratory (LJK), a joint research unit of CNRS, Grenoble Alpes University (UGA), Grenoble Institute of Engineering (INP), and National Institute for Research in Digital Science and Technology (Inria). The LJK focuses on research in computer science and applied mathematics. The intern will join the DATA department of LJK (which stands for Données et Aléatoire : Théorie et Applications / Data and Stochastic: Theory and Applications), a core scientific division focusing on both the theoretical and applied aspects of data analysis and random phenomena..

The MIAI (Multidisciplinary Institute in Artificial Intelligence) AI cluster aims to conduct cutting-edge research in artificial intelligence, offer engaging courses for students and professionals at all levels, support innovation within large companies, SMEs, and startups, and inform and engage with the public regarding all aspects of AI. It brings together partners from the public and private sectors and aims to develop AI for the benefit ofpeople and the environment. https://miai-cluster.univ-grenoble-alpes.fr/miai-institute/