

Raphaël Rubrice

Incoming PhD student, from January 2027
Université de Montréal / Mila & TUM / Helmholtz Munich
Research Assistant, Theis Lab, Helmholtz Munich

raphael.rubrice@helmholtz-munich.de
raphaelrubrice.github.io
github.com/raphaelrubrice
linkedin.com/in/raphael-rubrice
Google Scholar
ORCID
Munich, then Montréal from 2027

RESEARCH INTERESTS

My main background is in computational biology and machine learning. I am interested in *AI-assisted experimental design*, helping decide which experiments to run next, and in how our models hold up *out of distribution*, in biological contexts they were not trained on. Both are steps toward virtual experimental biology, where much of the experimental work could start in silico.

EDUCATION

from Jan 2027 — **PhD, Applied Mathematics** — Université de Montréal / Mila & TUM / Helmholtz Munich
Joint PhD supervised by Guy Wolf (UdeM, Mila) and Fabian Theis (TUM, Helmholtz Munich); first year in Montréal.

2025 – 2026 — **MVA Master, Mathematics, Vision, Learning** — ENS Paris-Saclay & Université Paris-Cité
Highest Honors. Health track: machine learning for health and biology, generative modeling, protein science. Master's thesis at Helmholtz Munich (BANmyBED, below).

2024 – 2025 — **MSc Computational Biology** — Université Paris-Saclay
Highest Honors. Bioinformatics and AI: learning theory, deep learning, gene-expression analysis.

2022 – 2026 — **Engineering degree, life sciences** — AgroParisTech
National institute for life sciences and industries. Health-engineering specialization; biostatistics and machine learning.

2020 – 2022 — **Classe préparatoire BCPST** — Lycée Masséna, Nice
Competitive preparatory track for the grandes écoles: biology, chemistry, physics, mathematics.

RESEARCH EXPERIENCE

2026 – present — **Research Assistant & Master's thesis** — Theis Lab, Helmholtz Munich
- Thesis: **BANmyBED**, batched, amortized and non-myopic Bayesian experimental design for lab-in-the-loop CRISPR screens. A Bayesian correction adapts frozen design policies to a new screen, and inference-time steering trades off hits against diversity.
- Accepted as a poster at the NGDL workshop, ELLIS UnConference (Palais des Congrès de Paris, 8 Dec 2026).
- Supervised by Xiaotong Fu, Alessandro Palma, Lorenzo Consoli, Fabian Theis and Guy Wolf.

2025 — **Research internship, 6 months** — Dassault Systèmes, Clinical Decision team
- Explainable clustering for clinical data, with a focus on small cohorts.
- Surveyed (non-deep-learning) explainable methods and proposed a framework to interpret previously obtained clusters; delivered an internal Python package.
- Supervised by Aziliz Cottin and Marine Zulian.

2023 – 2025 — **Research project** — UMR MIA Paris-Saclay & Micalis Institute
- Predicting bacterial antagonism (*Bacillus* vs. pathogens) from biofilm morphology; homology-controlled cross-validation, SHAP and permutation importance. Led to a journal publication.
- Supervised by Antoine Cornuéjols and Vincent Guigue, with Romain Briandet and Virgile Guéneau.

2024 — **Research internship, 3 months** — Cristofari Lab, IRCAN, Nice
- TELLAM: an RNA-seq pipeline to quantify retrotransposon expression in senescent cells.
- Supervised by Sophie Lanciano and Gaël Cristofari.

AWARDS & FELLOWSHIPS

2026 – 27 — **Unaite Fellowship for AI in Biology** — Unaite, inaugural cohort, Paris

2026 — **Impact+ Doctoral Scholarship** — Canada Impact+ Research Training Awards

2025 — **AI Methodology Award** — Glioblastoma hackathon, Owkin & Servier

2025 — **PR[AI]RIE-PSAI Institute Scholar** — PR[AI]RIE, Paris School of AI

PUBLICATIONS & PRESENTATIONS

Dec 2026 — **BANmyBED: Batched, Amortized and Non-myopic Bayesian Experimental Design for Lab-in-the-Loop CRISPR Screens**
R. Rubrice, X. Fu, A. Palma, L. Consoli, F. J. Theis, G. Wolf. *Poster*, Workshop on Next Generation of Discovery Loops (NGDL), ELLIS UnConference, Palais des Congrès de Paris. Non-archival.

2025 — **A machine-learning framework for the prediction and analysis of bacterial antagonism in biofilms using morphological descriptors**
R. Rubrice, V. Guéneau, R. Briandet, A. Cornuéjols, V. Guigue. *AI in the Life Sciences*, Elsevier. Paper, Code, Data.

Jun 2025 — **Predicting forage nutritional value with AI**
G. Tran, R. Genin, A. Lauront, M. Petitet, **R. Rubrice**, V. Heuzé, V. Guigue, A. Cornuéjols. *Oral presentation*, 1st EAAP Conference on AI in Animal Science (AI4AS), ETH Zurich.

SELECTED COURSE PROJECTS

single-cell — **Extending scVAE: a hierarchical mixture-of-mixtures approach** — MVA, P. Latouche, P.-A. Mattei
New architecture enforcing a hierarchical bias on single-cell gene-expression modeling. Code

perturbation — **Label-informed optimal transport for cross-modality matching** — MVA, J. Feydy
Labeled Gromov-Wasserstein optimal transport for predicting perturbation responses (Ryu et al., 2024).

generative — **Ambient diffusion: clean distributions from corrupted data** — MVA, B. Galerne, A. Leclaire
Reimplemented Ambient Diffusion (Daras et al., 2023) from scratch.

protein — **Proteina: scaling flow-based protein structure generative models** — MVA, F. Cazals
Studied the model and investigated its representation space.

molecules — **Molecular captioning challenge** — MVA, M. Vazirgiannis
Among the top-scoring solutions; benchmarked a wide range of graph neural network architectures.

SERVICE

2026 — **Volunteer, Virtual Cell Workshop** — Helmholtz Munich

SKILLS & LANGUAGES

Programming — Python, PyTorch, scikit-learn, R, Bash

Tools — Git, Docker, Nextflow, Windows, Linux, macOS

Methods — Bayesian experimental design, generative models (VAEs, diffusion), single-cell and perturbation data analysis, graph neural networks, explainable AI

Languages — French (native), English (C2; IELTS 8/9, TOEIC 975/990)