RODRIGO ADOLFO REYES FEREGRINO

rodrigo.reyesferegrino@utoronto.ca ◇ LinkedIn ◇ Cel: 437-551-9980 ◇ rodri-rf.github.io

EDUCATION

Honours Bachelor of Science with Professional Experience Year (PEY), University of Toronto 2021 – 2026

- Double major: Computer Science & Chemistry | Minor: Philosophy | **Entrance Scholarship**
- Graduate-level coursework: Introduction to Quantum Algorithms (CSC2332), Deep Learning and Neural Networks (CSC413)
- Additional relevant coursework: Enriched Linear Algebra I, Enriched Data Structures, Algorithm Design & Analysis, Software Design, Software Tools & Systems Programming, Introduction to Computational Chemistry, Structural Biochemistry, Introduction to Medicinal Chemistry and Molecular Recognition

TECHNICAL SKILLS

ML & Deep Learning: PyTorch, Scikit-learn, Hugging Face Transformers, ESM3, Mamba2, Boltz-2; generative model training (language models, state-space models), BPE tokenization, MCMC optimization, Bayesian inference

Quantum Computing: Qiskit, PySCF, OpenFermion; Quantum Phase Estimation (QPE), Hamiltonian simulation, noise characterization, error mitigation, variational algorithms

Bioinformatics & Scientific Computing: Biopython, RDKit, COBRApy, NumPy, SciPy, Pandas, Matplotlib, Plotly; docking pipelines, string/k-mer algorithms, dynamic programming, graph-based assembly, Bayesian variant calling, FBA/FSEOF, sequence alignment (BLAST, MASH)

Computational Chemistry: Autodock, HADDOCK3, OpenBabel, Spartan (DFT and Hartree-Fock calculations)

Engineering & DevOps: Python (expert), C/C++, Java, SQL, Bash | Docker, Git, AWS (S3, EC2, RDS, Aurora), SLURM, CI/CD, Unix/Linux; cloud-distributed data pipelines, multi-threaded/multi-processed workflows

C++ Libraries: Boost, TBB, htslib, samtools, GROMACS

Data & Databases: PostgreSQL, AWS Aurora, Amazon RDS, SQLite, MySQL, FileMaker; ETL/ELT pipeline design, schema optimization, SQL query tuning

Languages: Spanish (native), English (C2), German (C1)

RESEARCH EXPERIENCE

Undergraduate Researcher: Quantum & Computational Chemistry Aug 2025 – Present
Chemical Physics Theory Group (Izmaylov Lab), University of Toronto *Toronto, Canada*

- Co-developed Sampling-based Quantum Diagonalization (SQD) method using Q-SENSE (Quantum SENiority Subspace Expansion) ansätze for NISQ-era electronic structure calculations.
- Designed numerical experiments, linking several parameters to sampling efficiency, to systematically optimize the Q-SENSE states and post-processing methods for best ground eigenstate estimation.
- Contributed theoretical insights connecting SQD, Q-SENSE, Adaptive Variational Optimization (AVO), and selected configuration interaction (CI) methods.
- Manuscript in preparation (see Publications).

Undergraduate Researcher: Quantum Computing May 2024 – Present
University of Toronto (Supervisor: Prof. Nathan Wiebe) *Toronto, Canada*

- Developed quantum simulation framework in Qiskit combining qDRIFT time evolution with QPE for Hamiltonian eigenvalue estimation; implemented both QFT-based and Kitaev's iterative QPE algorithms.
- Designed extrapolation protocol using Chebyshev node sampling to eliminate systematic $O(t^2)$ qDRIFT errors, approaching Heisenberg-limited precision scaling—directly applicable to near-term error mitigation pipelines.

- Investigated block-encoding methods (Linear Combination of Unitaries, Qubitization) and product formulas (Trotter-Suzuki, truncated Taylor series).
- Implemented parallelized parameter sweeps (joblib/loky) with CI-integrated automated test suites.

Dry Lab Lead & Researcher, iGEM Competition March 2022 – November 2025
iGEM Toronto, University of Toronto *Toronto, Canada*

- Led the computational modeling (Dry Lab) division of UofT's synthetic biology research team across 4 competition cycles. Managed a 10+ member cross-disciplinary team: set technical roadmaps, ran code reviews, mentored researchers, and translated stakeholder constraints into engineering requirements. Won **4× Gold Medal** and **2× Best Computational Model** out of 250+ international teams.
- Planned and organized weekly team meetings (60+ over tenure), set goals, conducted interviews, and coordinated with other leads and mentors across a 30+ member team.
- Presented research findings through poster presentations at conferences (e.g., Biozone, In Vitro 2023).

2025 — AI-Driven Phage Engineering (PHORAGER):

- Engineered bacteriophage receptor-binding proteins using AI-driven design integrating ESM3 generative protein language models, Boltz-2 structure prediction, MCMC optimization with simulated annealing, and HADDOCK3 molecular docking. All designs showed <40% identity to known proteins (BLASTp). Developed open-source platform reducing phage screening time from months to days.

2024 — Plasmid Foundation Model (Plasmid.AI):

- Developed machine learning pipeline for generating biologically viable, de novo plasmid sequences. Trained Mamba2-based generative model from scratch on 137K+ sequences with BPE tokenization, circular crop and reverse-complement augmentation, and two-phase pre-train/fine-tune schedule. Optimized for specific features including origins of replication (ORIs) and antimicrobial proteins.
- Conducted *in silico* validation leveraging advanced sequence alignment tools (BLAST, MASH) and motif discovery techniques to identify key plasmid features. Created the largest open-source toolkit for plasmid foundation models.

2023 — Metabolic Engineering:

- Led a team of 8 members in the design and implementation of a metabolic engineering project. Applied genome-scale metabolic models and linear optimization techniques (Flux Balance Analysis, Phenotypic Phase Plane analysis, FSEOF algorithm) to predict genetic modifications enhancing bacterial metabolic network efficiency. Evaluated over 12,000 genetic variations.

2022 — Oak Wilt Diagnostics:

- Used tools such as Primer3 and PrimerExplorer to develop software that optimized the loop-mediated isothermal amplification (LAMP) reaction for detecting the fungus causing Oak Wilt disease. Project won **Gold Medal** and nomination for Best Conservation Project.

Volunteer Researcher: Drug Discovery February 2024 – Present
Gene2lead *Mississauga, Canada*

- Leveraged unsupervised learning algorithms to identify candidates for drugs that could serve as MCHR1 antagonists. Worked with molecular structures in SMILES, SELFIES, and molecular fingerprint representations.
- Developed and automated pipelines for high-throughput screening of ligands that bind to the active site of the target enzyme through molecular docking simulations.

INDUSTRY EXPERIENCE

Software Developer (Bioinformatics) May 2025 – September 2025
Roche — Sequencing by Expansion (SBX) Technology R&D *Mississauga, Canada*

- Designed computational infrastructure for systematic benchmarking of Roche's novel SBX nanopore sequencing platform, evaluating basecalling quality, throughput, and accuracy across experimental conditions.
- Architected an end-to-end ETL pipeline for high-volume time-series sequencing performance data; migrated a legacy SQLite/S3 system to AWS Aurora (PostgreSQL), tuning schema and queries to enable reliable longitu-

dinal data tracking for 50+ engineers.
- Built a centralized performance dashboard and backing data services, translating semi-structured experimental output into actionable, queryable metrics for cross-functional R&D squads.
- Shipped production code under CI/CD, automated unit testing, and peer code review.

Data Scientist & Bioinformatician May 2024 – April 2025
Sanofi — Analytical Sciences R&D *Toronto, Canada*

- Extended production-grade genomics pipelines in Python and C++ processing 4.5B+ DNA reads per batch on HPC/cloud-distributed infrastructure, directly supporting vaccine quality workflows across 150+ countries.
- Implemented Bayesian inference methods for viral minority variant detection; developed novel heuristics (strand-bias correction, read-position mismatch-distribution analysis) to distinguish biological mutations from RT-PCR and sequencing artifacts, reducing false positives by 15%.
- Built graph-based algorithms using k-mer spectrum analysis and local sequence alignment to detect and quantify cross-sample contamination events arising from barcode misassignment and library preparation artifacts.
- Deployed optimized multi-threaded workflows on hybrid cloud/HPC clusters for large-scale genomic analyses.

Database Developer October 2023 – April 2024
Canadian Statistical Sciences Institute (CANSSI) Ontario *Toronto, Canada*

- Captured, entered, and analyzed data related to CANSSI Ontario operations, events, and programs. Harnessed existing software (FileMaker, Alchemer, MySQL) as well as developed custom tools with diverse Python libraries.
- Designed systems and pipelines for database management and automation, as well as for data cleaning and curation.

TEACHING EXPERIENCE

Teaching Assistant September 2024 – Present
University of Toronto *Mississauga, Canada*

- Lead tutorials, grade exams, host office hours, and help design course content for both the mathematics and computer science departments.
- Courses: CSC373 (Algorithm Design & Analysis), CSC363 (Computational Complexity & Computability), MAT135 (Differential Calculus)—led tutorial sections for 80+ students.

Private Tutor July 2019 – Present
Freelance and Tutorat Pro *Mexico City, Mexico and Mississauga, Canada*

- Tutored over 12 students in IB Chemistry, IB Physics, IB Mathematics, and German.

OPEN-SOURCE SOFTWARE

qDRIFT-QPE Extrapolation Framework (GitHub)

- Qiskit-based quantum simulation framework: qDRIFT + QPE (QFT-based & Kitaev's iterative), Chebyshev zero-noise extrapolation, parallelized parameter sweeps with automated testing.

PHORAGER (GitHub)

- AI-driven protein engineering platform integrating ESM3, Boltz-2, MCMC with simulated annealing, and HADDOCK3 docking. Modular OOP design with CI/CD pipelines and full API documentation.

Plasmid.AI (GitHub)

- Largest open-source toolkit for plasmid foundation models. Mamba2-based generative model trained on 137K+ plasmid sequences with BPE tokenization.

COBRA-FSEOF (GitHub)

- Python (COBRApy) implementation of the FSEOF (Flux Scanning based on Enforced Objective Flux) algorithm for bacterial strain design via constraint-based metabolic modeling, based on the work of Choi et al. (2010).

AWARDS & HONORS

- **iGEM 2025**: **Gold Medal**, Best Therapeutics Project. Project: PHORAGER.
- **iGEM 2024**: **Gold Medal**, Best Computational Model Award, Best Entrepreneurship Award + nominations for Best Foundational Advance Project and Best Presentation. Project: Plasmid.AI: an intelligent platform to write new life
- **Summer Undergraduate Data Science (SUDS) Research Scholarship**, University of Toronto (CAD $7,200, 2024)—fully funded research position at Lunenfeld-Tanenbaum Research Institute; declined due to concurrent commitment at Sanofi.
- **iGEM 2023**: **Gold Medal**, Best Computational Model Award + nominations for Best Climate Crisis Project, Best Wiki, Best Integrated Human Practices, Best Entrepreneurship, Best Presentation. Project: A comprehensive solution to harness landfill gas emissions
- **iGEM 2022**: **Gold Medal** + nomination for Best Conservation Project. Project: On-site early diagnostic tool for Oak Wilt disease
- **University of Toronto Entrance Scholarship** (2021)

ADDITIONAL EXPERIENCE

Robotics Student & Competitor 2009 – 2021
Schools: RobotiX, Nética *Mexico City, Mexico*

- Participated in over 35 competitive events at regional, national, and international levels, including FIRST Lego League, RoboCup, RoboRave, and World Robot Olympiad—represented Mexico at the international stage in Doha, Qatar.
- Gained hands-on experience in electrical and mechanical engineering, Arduino programming, and Raspberry Pi projects. Developed teamwork and creative problem-solving under pressure.

INTERESTS

Playing electric guitar (solo and ensemble performances), Scuba Diving (PADI Open Water certified).