

Murto Hilali

Toronto, ON | 647-569-4616 | mhilali@uwaterloo.ca | [LinkedIn](#) | [GitHub](#) | [Medium](#)

SUMMARY OF QUALIFICATIONS

- 6+ years of experience in biological data analysis & collaborative workflow development (Python, R, WDL, Bash, Git).
- Expertise with HPC clusters and containerization methods from OICR & LatchBio (Docker, SLURM, Cromwell).
- Proficiency with large -omics databases and tools for literature review (RCSB PDB, UniProtKB, NCBI, ClinVar, IntAct).
- Skilled in 3D-protein modeling & structural biology from interactomics study at SickKids (PyMol, ChimeraX, AlphaFold).
- Strong understanding of core machine learning skills, including exploratory analysis, visualization, & feature engineering.
- Influential communication (interpersonal, written, presentation) skills.

EXPERIENCE

Computational Chemistry Research Student | QDX Technologies | Singapore, SG 10/2024–12/2024

- Benchmarked performance of six protein-ligand docking models across three datasets (Davis, KIBA, PDBBind) using 30+ metrics, including speed, physical plausibility (PoseBusters), and experimental binding correlation.
- Optimized retraining of DiffDock-Pocket and FABind on 11,500+ protein-ligand complexes, integrating early stopping mechanisms in PyTorch to save hours in training while ensuring reproducibility.
- Validated ~25 protein-ligand poses using PoseBusters and PyMol, uncovering steric clashes, misalignment, and deep learning models' limitations in producing physically plausible outputs.
- Evaluated 12+ protein-ligand databases (10K+ entries each, several GB) for binding affinity data, identifying leakage-free datasets to improve the reliability of retrained docking models.

Bioinformatics Researcher | The Hospital for Sick Children | Toronto, ON 01/2023–12/2023

- Implemented a machine learning algorithm (XGBoost) to classify and predict structural impacts of de novo missense variants on protein interactions with a micro-average Area Under Curve of 0.91.
- Leveraged biocomputational best practices and tools (exploratory data analysis, Jupyter notebooks, GPU parallelization, hyperparameter tuning, SHAP values) to find 2 potential protein mechanisms involved in Autism Spectrum Disorder.
- Pioneered a novel approach to predict disease mechanisms of de novo mutations in Cerebral Palsy by mastering AlphaFold2-Multimer and structural biology fundamentals, generating interaction data on 25+ proteins and leading to upwards of 500+ new research paths per data point.
- Contributed to a paper published in Nature Genetics, collaborating with 10+ researchers and 3 organizational stakeholders, including Holland Bloorview Kids Rehabilitation Hospital.
- Presented project insights to 45+ researchers and executives at the Scherer Lab in The Center for Applied Genomics.
- Achieved authorship credits on a scientific abstract presented at the International Cerebral Palsy Genomics Consortium.

Bioinformatics Workflow Developer | OICR | Toronto, ON 01/2022–04/2022

- Accelerated workflow execution time by 50% via parallelization techniques.
- Developed several workflows from scratch in a high-performance computing (HPC) cluster environment.
- Collaborated via GitHub, Bitbucket, Jira, and Confluence to enhance four existing pipelines in OICR's codebase.
- Constructed regression tests for all pipelines via Jenkins that accounted for speed and stochastic elements in data.

SKILLS

- *Scientific*: Protein-Protein Interactions, Immunology, Structural Bioinformatics, WGS/WES Data, Biostatistics
- *Development*: Distributed Model Training, Large Dataset Management, Parallelized Hyperparameter Tuning, Workflow Development, Regression Testing, Agile Development, Agile Research, Machine Learning, Deep Learning, Statistics
- *Software*: Linux/Unix environments, Python, R, Jira, Confluence, BitBucket, WDL, Bash, Git, SLURM, NumPy, SciPy, BioPython, Seaborn, Scikit-Learn, PyTorch, TensorFlow, XGBoost, AlphaFold

EDUCATION

Honour's Bachelor of Science (BSc) in Biotechnology/Economics | University of Waterloo, Waterloo, ON

Relevant Courses: Mathematical Modeling of Molecular Systems, Methods in Bioinformatics, Biostatistics & Experimental Design, Advanced Cell Biology

Expected Graduation: 04/2026