Sam Alexander Martino

sam@1martinos.dev
1martinos
United Kingdom
1martinos.dev

About Me

Doctor of Physics from UCL specialising in developing and employing cutting-edge machine learning techniques to solve complex challenges in biophysics. Extensive experience working with large datasets as computational lead in both interdisciplinary teams and on self-driven projects. A highly-skilled programmer, research software developer, and educator with a passion for writing practical code, advancing AI technologies to solve real-world problems, and sharing my knowledge and expertise. I am looking to work on pioneering drug discovery methods by contributing to the development of scalable, practical, and performant methods and models.

Education

Doctorate in Computational Biophysics — **University College London**
↳ Applications of Machine Learning and GPGPU programming in Computational Drug Discovery — *January 2020 – September 2025*

M.Sc. Theoretical Physics — King's College London *2018-2019*
B.Sc. Theoretical Physics — University of York *2015-2018*

Selected Academic Work

Optimization of Enzymatic Reactivity via Graph Neural Network Small Molecule Screening

In Preparation

Designed and implemented an active-learning pipeline using a *GNN* to predict *QM/MM free-energy barrier calculations*, accelerating ligand selection for a *RAS (G12D)* drug discovery project. Engineered the feature pipeline linking docking/force-field results to model input, benchmarked results against *ChemProp* and multiple other *GNN* architectures, demonstrating a 10% enrichment over docking-only selection.

Kemeny Constant-Based Optimization of Network Clustering Using Graph Neural Networks

J. Phys. Chem. B 2024, 128, 34, 8103–8115 — *2024*

Developed a GNN-based method for optimising *Markov State Model* clustering of *molecular dynamics data*, benchmarked against *PyEMMA* and the standard *PCCA+* algorithm. Designed a general parallel-training framework for large-scale simultaneous model training.

CACHE Challenge #2: *Drug Discovery Challenge*

Summarised in Journal of Chemical Information and Modeling 2025 65 (13), 6884-6898 — *2023-2024, published 2025*

Computational lead within a multidisciplinary team screening the 6.75-billion-compound virtual library against NSP13. Forked and adapted an open-source GPU similarity-search codebase *(C++/CUDA)* to batch-process *~4TB* of molecular fingerprint data, alongside building a Python pipeline orchestrating consensus docking across CPU and GPU computing clusters. Parsed and analysed *QM/MMGBSA* and *FEP* results to select experimentally-tested candidates, yielding a confirmed novel binder.

Modelling the active SARS-CoV-2 helicase complex as a basis for structure-based inhibitor design

Chem. Sci., 2021,12, 13492-13505 — *2021*

Modelling and simulating RNA helicase NSP13, a protein essential for the replication process of the SARS-CoV-2 virus.
Performed kinetic and conformational analysis across multi-microsecond MD datasets spanning three engines *(NAMD, AMBER, GROMACS)* and multiple custom force fields. Later designed custom collective-variable biasing protocols using *NAMD colvars* to attempt to simulate translocation.

Selected Teaching Experience

Physics & Computing Courses at UCL — *2021 – 2024*

Lecturer, Demonstrator, Marker — *University College London*

Taught numerous undergraduate and postgraduate courses at UCL, including:

- Python programming courses, advanced mathematics, and postgraduate courses on GPGPU programming.
- Supervising undergraduate and postgraduate research projects, introducing students to HPC, ML, and biophysics research.

CCP5 Summer School: Advanced course on biomolecular simulation — *August 2022*

Lecturer & Workshop Organizer — *Durham University*

Organised, devised, and delivered a lecture and workshop introducing 30 first-year PhDs to protein simulation methods using CHARMMGUI.

Technical skills

Machine Learning	Significant experience developing and using tools for modern **ML**, including using, building, bench-marking, and fine-tuning models in both **PyTorch** and **JAX**. Highly proficient applying **Graph Neural Networks** for real problems in physics and biochemistry
Programming	Experienced writing **Python**, **bash**, **C/C++**, **CUDA**, **FORTRAN**, and **MATLAB** code in a research environment and at large scales both locally and on **HPC** computing clusters.
Software Development	Proficient at **developing shared software libraries** and using **Git** and **GitHub** for **VCS**.
Domain Knowledge	Direct experience working with **biological** and **chemical data**. Proficient setting up, running, and processing the results of **molecular dynamics** calculations **across several engines and force fields**. Including using **enhanced sampling methods** and experience working with **free-energy calculations**