Martina Occhetta

PhD student · ML for Drug Discovery

London, England, United Kingdom

(+44) 7936-754-128 | m.occhetta@qmul.ac.uk | martina-occhetta.github.io | linkedin.com/in/martina-occhetta

Summary

PhD student in AI for Drug Discovery with academic and industry experience applying machine learning to biology and healthcare. Internships at IBM Research and Exscientia focused on biomedical foundation models, single-cell omics, and protein structure prediction. Skilled in Python, deep learning, and scalable workflows, with publications in top venues. Driven to translate AI research into real biological and healthcare impact.

Work Experience

IBM Research

Hartree Centre, Daresbury, UK

Research Intern — *May 2025 – Aug 2025*

- Collaborated with the AI for Healthcare and Life Sciences team on biomedical foundation models (BMFM).
- Implemented and optimized BMFM finetuning workflows for domain-specific biomedical tasks, improving model adaptability.
- Extended Transcriptominer, an automated pipeline for large-scale gene expression data retrieval and analysis, to support single-cell data.
- **Technical Skills:** Kubernetes, biological foundation models, machine learning for healthcare.
- **Soft Skills:** Cross-functional collaboration, scientific communication, problem solving, adaptability in fast-paced environments.

Yusuf Hamied Department of Chemistry, University of Cambridge

Cambridge, UK

Student Researcher, supervised by Prof. Michele Vendruscolo — *May 2023 - Aug 2023*

- Developed a computational approach to predict single protein liquid-liquid phase separation (LLPS) and co-condensation of protein pairs.
- Research findings published in *PNAS* (2024).
- **Technical Skills:** AlphaFold2; JAX; Haiku; graph representation learning; self-attention; protein encoding; LaTeX.
- **Soft Skills:** Time management, problem solving, logical thinking, presentation skills, report writing.

Exscientia

Oxford, UK

Bioinformatics Intern — *July 2022 - Sept 2022*

- Collaborated with Structural Bioinformatics and Active Learning teams.
- Focused on using active learning to explore how to improve AlphaFold2 performance, with final aim of ensuring full domain coverage by highly confident (pLDDT score > 70) AlphaFold2 models.
- **Technical Skills:** Active Learning, Python, bitbucket, AlphaFold2, MMseqs2, NCBI BLAST, Pfam.
- **Soft Skills:** Teamwork, Time Management, Communication, Logical Thinking, Presentation skills.

Education

Recursion & Queen Mary University of London

London, UK

PhD in AI for Drug Discovery — *Sept 2023 - Sept 2027*

- Project: *"Target identification from multi-omics data using systems biology and machine-learning approaches"*
- Supervisors: Prof. Conrad Bessant (Professor of Bioinformatics, QMUL), Dr. Mani Mudaliar (Director, Quantitative Biomarkers, Recursion)
- Research interests: GNNs, network biology, multi-omic data integration, biological foundation models, perturbation response prediction
- Teaching assistant: AI and Data Science in Biology (2023 & 2024), Coding for Bioscientists (2024), Research Methods and Communication (2024)

publications and preprints:

- Wenteler, A.*, **Occhetta, M.***, et al. 'PertEval-scFM: Benchmarking Single-Cell Foundation Models for Perturbation Effect Prediction'. bioRxiv, 3 October 2024. https://doi.org/10.1101/2024.10.02.616248. Published at *ICML 2025*. Presented at *NeurIPS 2024 Workshop AIDrugX*.
- **Occhetta, M.**, et al. 'MORGaN: self-supervised multi-relational graph learning for druggable gene discovery'. bioRxiv, 16 September 2025. https://doi.org/10.1101/2025.09.10.675402.

University of Cambridge

Cambridge, UK

MPhil in Computational Biology, Department of Applied Mathematics and Theoretical Physics — *Oct 2022 - Sept 2023*

- Graduated with First Class Honours
- Modules including Scientific Programming, Deep Learning, Biological Imaging and Analysis, BioDesign, Systems Biology

publications and preprints:

- Zhang, S.*, Lim, C.*, **Occhetta, M.***, and Vendruscolo, M. (*equal contribution). 'AlphaFold2-Based Prediction of the Co-Condensation Propensity of Proteins'. Proceedings of the National Academy of Sciences 121, no. 34 (20 August 2024). https://doi.org/10.1073/pnas.2315005121.

Imperial College London

London, UK

BSc in Biochemistry, Department of Life Sciences — *Sept 2019 - July 2022*

- Graduated with First Class Honours; placed on the Dean's List each year for outstanding academic achievement.
- Final Year Project: *"Improving deep neural network-based classification of molecular dynamics trajectories of intrinsically disordered proteins through feature engineering"* - achieved First Class