Martina Occhetta

London, United Kingdom

📱07936754128 | ✉ martina.occhetta2@gmail.com | 🐙 github.com/martina-occhetta | 💼 linkedin.com/in/martina-occhetta

Education

Exscientia & Digital Environment Research Institute, Queen Mary's 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"*
- Co-supervisors: Prof. Conrad Bessant (Professor of Bioinformatics, QMUL) and Mani Mudaliar (Target Analyst Director, ExScientia)

Wolfson College, University of Cambridge   *Cambridge, UK*

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

- MPhil Project: *"AlphaFold-based prediction of the propensity of proteins for phase separation"*
- Modules including Scientific Programming, Deep Learning, Biological Imaging and Analysis, BioDesign, Systems Biology

Imperial College London   *London, UK*

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

- Graduated with First Class Honours
- Final Year Project: *"Improving deep neural network-based classification of molecular dynamics trajectories of intrinsically disordered proteins through feature engineering"* - achieved First Class

International School of Milan   *Milan, Italy*

International Baccalaureate   *Sept 2017 - July 2019*

- Overall grade: 43/45. HL subjects: Chemistry (grade: 7/7), Biology (7), Mathematics (6), and Economics (7)

Work Experience

Exscientia   *Oxford, UK*

Bioinformatics Intern   *July 2022 - Sept 2022*

- Collaborated with Structural Bioinformatics and Active Learning teams.
- Focus: 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.
- Acquisition function to quantify Expected Improvement of models to prioritise domains to be experimentally solved.
- **Technical Skills:** Active Learning, Python with NumPy, matplotlib, Pandas, scikit-learn, bitbucket, AlphaFold2, MMseqs2, NCBI BLAST, Pfam.
- **Soft Skills:** Teamwork, Time Management, Communication, Logical Thinking, Presentation skills.

Anatomical Pathology Laboratory, Ospedale San Raffaele   *Milan, Italy*

Research Intern   *July 2021 - Sept 2021*

- Study samples from ovarian cancer patients using FISH and next-generation sequencing technologies.
- Downstream data analysis to discern mutations that could cause cancerous growth.
- Responsible for data interpretation - use of bioinformatics software and databases to predict impact of identified mutations on transcription and translation.
- **Technical Skills:** FISH, NGS, Torrent Suite, BLAST, ClustalO, HHpred, Phyre2, PyMOL, FFPred3, STRING.
- **Soft Skills:** Time management, team work, communication, idea exchange, creativity, flexibility, problem solving.

Projects

AlphaFold2-based prediction of the propensity of proteins for phase separation   *Cambridge, UK*

MPhil Project, Centre for Misfolding Diseases, Yusuf Hamied Department of Chemistry   *March 2022 - June 2022*

- Developed a computational approach to predict single protein liquid-liquid phase separation (LLPS) and co-condensation of protein pairs.
- Single protein LLPS predictor: AUC of 0.87, accurately classified 79% of set of experimentally validated proteins. Phase separation propensity profiles reveal correlation between phase separation propensity and disordered regions.
- Co-condensation predictor demonstrated even more remarkable performance (AUC: 0.89), illustrating the potential to provide insights into the context-dependent nature of PPS.
- Models promise to enhance understanding of biomolecular condensate evolution and contribute to identification of promising protein targets for modulating biomolecular condensate formation.
- Submitted 15000-word report. Attained First Class (79.00).
- **Technical Skills:** AlphaFold2; JAX; Haiku; graph representation; self-attention; protein encoding; Overleaf; LaTeX.
- **Soft Skills:** Time management, problem solving, logical thinking, presentation skills, report writing.

AE-based classification of c-Myc MD trajectories *London, UK*

BSc Final Year Project, Imperial College London *March 2022 - June 2022*

- Use molecular dynamics simulations to determine conformational changes that occur in mutant $^{91-160}$c-Myc and lead to functional activation.
- Use local compaction-based dimensionality reduction and feature selection methods as starting point for autoencoder-based machine learning methods of trajectory analysis.
- Autoencoder, trained on the most relevant local compaction data, is capable of separating conformational ensembles into similar sub-classes.
- Workflow has wide applicability towards systematic characterisation of the conformational states of IDP ensembles.
- Code available on GitHub. Submitted 6000-word report. Attained First Class (72.70).
- **Technical Skills:** Python with Keras, TensorFlow, NumPy, Pandas, matplotlib, seaborn, scikit-learn; AMBER; CPPTRAJ; Overleaf; LaTeX.
- **Soft Skills:** Time Management, teamwork, presentation skills, report writing.

The SARS-CoV-2 M^{pro} as a target for peptidomimetic and small molecule inhibitors *London, UK*

BSc Tutored Dissertation, Imperial College London *Feb 2021 - June 2021*

- Researched efforts undertaken to identify repurposed drugs and newly synthesised leads that target the main protease (M^{pro}) of SARS-CoV-2.
- Following literature review, critically analysed which drugs are most likely to be effective with reference to mechanisms of inhibition.
- Submitted 4000-word report. Attained First Class (83.75).
- **Technical Skills:** PDB, PyMOL, BioRender, ClustalO WebLogo, ChemDraw, Microsoft Office.
- **Soft Skills:** Research skills, problem solving, time management, logical thinking, communication, report writing, presentation skills.

Skills

Programming	Python (NumPy, Pandas, PyTorch, PyTorch Lightning, TensorFlow, Scikit-learn, BioPython, etc.), R, MatLab, Julia.
Software	AlphaFold2, Phyre2, PyMOL, Autodock4, Coot, CcpNmr, ChemDraw, AUGUSTUS, NCBI BLAST, MMseqs2, JPred, ClustalO.
Databases	EMBL, GenBank, Ensembl, InterPro, Pfam, Swiss-Prot, DisProt, PDB, SCOP, CATH, PROSITE, PANTHER.
Miscellaneous	Ubuntu Linux, Virtual Machines, LaTeX(Overleaf), Markdown, Microsoft Office, Git.

Achievements

2022	**Beloff-Chain Prize**, Imperial College London	*London, UK*
2022	**Convener's Prize (Bioinformatics)**, Imperial College London	*London, UK*
2022	**Dean's List**, Imperial College London	*London, UK*
2021	**Biochemistry Second Year Prize**, Imperial College London	*London, UK*
2021	**Dean's List**, Imperial College London	*London, UK*
2021	**Python Data Analysis Course**, LinkedIn Learning	*London, UK*
2020	**Dean's List**, Imperial College London	*London, UK*
2019	**Steve Marczylo Prize for Sciences**, International School of Milan	*Milan, Italy*
2018	**Basic Life Support Certification**, Croce Rossa Italiana	*Milan, Italy*
2017	**HSK Chinese Proficiency Test**, Level 1, (Score: 191/200)	*Milan, Italy*

Interests

Mentee	Selected to be a mentee at LeadTheFuture, leading non-profit organisation that offers a mentoring program for STEM students, with an acceptance rate below 20%.
Volunteering	Volunteer at Fondazione Francesca Rava. Participation in yearly *In Farmacia per i Bambini* event, contributing to collection of pharmaceutical products for donation. In 2017, spent two weeks at NPH República Dominicana, assisting orphaned children.
Tutoring	Worked with Velasca Tutors (2019 to 2023), mentoring GCSE and IB students in Chemistry, Biology, Economics, and Mathematics.
Travelling	Visiting over 20 countries has exposed me to different cultures and reinforced my organisational, communication, adaptability, and problem-solving skills.
Reading	I enjoy reading fiction (especially crime novels) and non-fiction books (mainly on science and the social sciences).

Languages

English	Bilingual proficiency
Italian	Native proficiency
Spanish	Intermediate proficiency
Chinese	Elementary proficiency

References available upon request.