Charles Harris

07792 541377 | cch57@cam.ac.uk | GitHub | Website | LinkedIn | British Citizen

EDUCATION

University of Cambridge   Oct. 2021 – Present
PhD in AI for Drug Discovery and Structural Biology   Cambridge, UK
- **Research**: Focus on molecular design, protein dynamics and discovering new biology with self-supervised learning
- **Supervisors**: Prof Sir Tom Blundell, Prof Pietro Liò, Prof Andres Floto
- **Funding**: Cambridge Centre for AI in Medicine (CCAIM) Studentship, sponsored by AstraZeneca and GSK

Imperial College London   Oct. 2020 – Sept. 2021
*MSc in Bioinformatics and Theoretical Systems Biology - **Distinction - 76.5%***   London, UK
- **Supervisors**: Prof Michael Bronstein (Imperial, Twitter), Prof Bruno Correia (EPFL), Prof Michael Sternberg

Imperial College London   Oct. 2017 – Sept. 2020
BSc in Biochemistry - 2:1   London, UK

SELECTED PROJECTS

MSc Project 3 - Distinction (76.9%) - Paper   Supervisors: Prof Michael Bronstein, Prof Bruno Correia
Graph generative models for structure-based drug design   Jun. 2021 – Sept. 2021
- Developed a molecular graph generation procedure which improves previous work by conditioning multiple stages of the generative procedure with an attention mechanism between the molecule and protein pocket (using dMaSIF)

MSc Project 2 - Distinction (81.5%) - Website   Supervisors: Prof Michael Bronstein, Prof Bruno Correia
Self-supervised learning of interpretable protein dynamics   Apr. 2021 – Jun. 2021
- Created a novel message-passing layer for protein graphs that learns interpretable protein dynamics

MSc Project 1 - Distinction (74.5%) - Paper   Supervisor: Prof Michael Sternberg
Prediction of the effect of missense variants involved in ligand binding   Jan. 2021 – April. 2021

Graphein - Contributor - Repository - *Paper Under Review*
Graphein aims to democratise access to biological data for geometric deep learning models

PREPRINTS

Graphein - a Python Library for Geometric Deep Learning and Network Analysis on Protein Structures and Interaction Networks. A. R. Jamasb, R. Viñas Torné, E. J. Ma, **C. Harris**, K. Huang, D. Hall, P. Lió, T. L. Blundell.

TECHNICAL SKILLS

Deep Learning: PyTorch, PyTorch Geometric, NetworkX, Tensorflow/Keras , some experience with JAX/Jraph
Computing: Linux, SQL, GPU machines, computing clusters
Statistics: Proficient in R, Probability, Likelihood statistics, Bayesian statistics, Linear regression, Classification
Structural Biology: Deep understanding of protein structure. Highly proficient in PyMol. Highly experienced in structure prediction (AlphaFold2, Phyre2) Experience with building atomic models in Coot and NMR analysis in CCPNmr
Chemoinformatics/drug design: RDKit, Conventional docking (AutoDock, SWISSDOCK)
Biochemistry and molecular biology: Strong background in and understanding of common lab data collect techniques
Bioinformatics: Genome assembly and annotation, RNA-seq data analysis, Sequence alignment, Protein function and structure prediction, NGS data analysis in R and Artemis
Systems biology/MatLab: ODEs, Master equations, Gillispie simulations, Dynamical systems

COMMUNICATION

Chair and Founder - 1st Cambridge AI in Drug Discovery Conference <u>Link</u>
Organised for late February 2022. Hearing from top academics and business leaders Feb. 2022

Guest - iGEM Synthetic Biology Podcast <u>Link</u>
Discussed AlphaFold2, my research and the impact of computation and AI on biology in general Aug. 2021

Chair and Founder - 1st Imperial AI in Drug Discovery Conference <u>Handbook</u>
Sold over 1,400 tickets Feb. 2021

- Moderated two panel discussions (first one with 4 CEOs/Founders of AI in Drug Discovery companies and second with Prof Sir Tom Blundell, Prof Michael Bronstein and Dr Andreas Bender)

SUPERVISING

Part III Computer Science Tripos Project
Self-supervised learning of ligand-binding Oct. 2021 – Present

MPhil Computer Science Project
AI for reconstruction of ligand-induced conformational rearrangement from Cyro-EM Jan. 2020 – Present

Visiting Masters Project
AI for enzyme kinetics predictions using multi-modal representations Jan. 2022 – Present

ADVISORY POSITIONS

Advisor - *Cambridge University International Genetically Engineered Machine (iGEM) Team* Dec. 2021 - Present
Advisor - *Imperial College International Directed Evolution Competition (iDEC) Team* Jun. 2021 - Oct. 2021
Mentor/Organiser - *Catalyse Competition - <u>SynBioUK</u>* Dec. 2020 - Present

TEACHING

Representation Learning on Graphs/Networks - *Part III/MPhil CS, University of Cambridge* Jan. 2022 - Mar. 2022
Bioinformatics - *Part II Computer Science Tripos, University of Cambridge* Nov. 2021 - Dec. 2021

AWARDS

CCAIM PhD Studentship - *University of Cambridge* Oct. 2021
Associateship - *Royal College of Science* Jul. 2020
Gold - *UK Chemistry Olympiad* Jun. 2017
David Lean Scholar - *Leighton Park School* Sep. 2015

VOLUNTEERING

Events and Conference Officer
Cambridge University Artificial Intelligence Society Oct. 2021 – Present

- Chair and organiser for the 'Cambridge AI in Drug Discovery Conference' planned for 2022

Chair and Founder <u>Twitter</u> - <u>Instagram</u>
Imperial College Computational Biology Society Oct. 2019 – Aug. 2021

- Organised regular webinar schedule with top companies and academics in computational biology (including DeepMind (Dr Petar Velikovic), BenevolentAI, Nvidia, LabGenius)
- Organised six workshops to teach biologist computational and coding skills
- Quickly grew the Society to over 340 members and developed a strong social media presence

- Showed strong leadership, organisational and team building skills by creating an effective team (20+ volunteers)

Trustee

Antony McGreevey Charitable Trust   Oct. 2021 – Present

- Focusing on awarding grants to charities that tackle the problems of Dyslexia for those from disadvantaged backgrounds

Ambassador

Helen Arkell Dyslexia Charity   Oct. 2019 – Present

REFERENCES - all are past and present supervisors

Prof Sir Tom Blundell FRS - University of Cambridge   tom@cryst.bioc.cam.ac.uk

Professor Emeritus of Biochemistry, Former President - BBSRC, Biochemical Society

Prof Peitro Lio - University of Cambridge   pl219@cam.ac.uk

Professor of Computational Biology. Member - Artificial Intelligence Group and CCAIM

Prof Michael Bronstein - Imperial College London, Twitter   m.bronstein@imperial.ac.uk

Chair of Machine Learning at Imperial, Head of Graph Learning Research at Twitter

Prof Bruno Correia - EPFL   bruno.correia@epfl.ch

Head - Laboratory of Protein Design and Immunoengineering

OTHER SKILLS

Languages: Mandarin (A at GCSE Level)
Interests: Glider pilot, Hockey, Running, Drone photography, Science communication