



MSc in Artificial Intelligence for Drug Discovery

Queen Mary University of London

Unofficial brochure – put together by

Dr Arianna Fornili (a.fornili@qmul.ac.uk)

MSc in Artificial Intelligence for Drug Discovery

Queen Mary University of London

- One of the first programmes in the world to focus on the applications of AI in Drug Discovery.
- No prior knowledge of machine learning or programming is assumed, making this ideal for graduates from Chemistry or a related discipline who would like to specialise in the applications of AI in Drug Discovery.

Entry Requirements

UK: a 2:1 or above at undergraduate level in Chemistry, Pharmaceutical Chemistry, Medicinal Chemistry, Biochemistry, Pharmacy, Biomedical Sciences or a related discipline.

International: an international qualification of similar standing to the above.

Applications are now open for **Google DeepMind scholarships**

Deadline: 10 June 2024 at 9am (UK)

<https://www.qmul.ac.uk/scholarships/items/deepmind-scholarship-2024-2025.html>



Programme structure

- 1 year (Full Time)
- 7 Taught Modules (120 credits) + Research project (60 credits)
- Taught modules are designed to cover **coding & AI**, **drug discovery** and **application of AI to drug discovery**

Semester A Modules	Semester B Modules
Machine and Deep Learning (30 credits)	
Fundamentals of Medicinal Chemistry (15)	Fine-tuning lead compounds (15)
Scientific Programming for Drug Discovery (15)	Data Driven Drug Discovery (15)
Molecular Modelling for Drug Discovery (15)	Computational ligand-based drug discovery (15)

Research project: Semesters B and C

Taught Modules

Scientific Programming for Drug Discovery Machine and Deep Learning

These modules cover key concepts of scientific programming and machine learning with emphasis on the development of **practical skills**. Face-to-face teaching will be followed by **practical sessions in the computer lab**.

Indicative topics include:

- Python
- Tools for code development and sharing (e.g. Jupyter, GitHub)
- Using code to analyse chemical databases
- Unsupervised and supervised learning methods
- **Artificial neural networks (CNN, RNN, GNN)**
- Generative models
- TensorFlow/Keras

Taught Modules

Fundamentals of Medicinal Chemistry

Fine-tuning lead compounds

These modules will equip students with a comprehensive understanding of **principles of medicinal chemistry** and of **lead compound optimization** in drug discovery

Indicative topics include:

- Drug Targets
- Drug action and pharmacodynamics
- Lead identification
- **Structure-activity relationship (SAR)**
- **ADME/Tox profiling**
- **Chemical modifications**
- **Synthetic viability**
- Ligand efficiency
- Bioisosteres
- Prodrugs

Taught Modules

Computational ligand-based Drug Discovery

Data-driven Drug Discovery

These modules will cover advanced **machine and deep learning techniques applied to drug discovery.**

Indicative topics include:

- Molecular representations, descriptors and fingerprints
- **Deep learning for protein structure prediction (e.g. AlphaFold)**, binding affinity prediction and virtual screening
- Quantitative Structure Activity Relationship (QSAR)
- **ML-based prediction of ADMET**
- Chemical datasets for machine learning benchmarking
- **Generative models for de novo drug design**

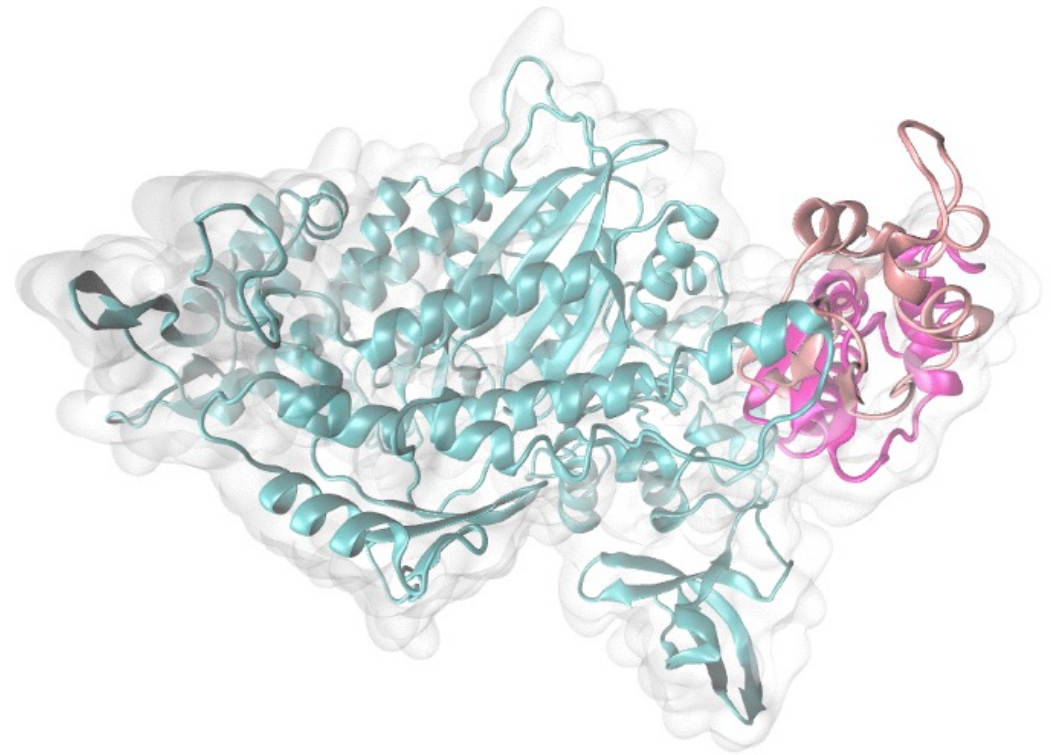
Taught Modules

Molecular Modelling for Drug Discovery

This module covers the main **molecular modelling techniques** used in drug discovery, with emphasis on structure-based approaches.

Indicative topics include:

- homology modelling
- molecular docking
- **virtual screening**
- molecular dynamics simulations
- free energy perturbation



Examples of Research Projects*

- Reducing **cardiac cytotoxicity** with computational methods
- A combined machine learning and synthesis approach to the generation of **new cannabinoid mimics with therapeutic potential**
- **Generative models for drug discovery** in skeletal and cardiac muscle disease
- **Ab initio machine learning force fields** to investigate water-protein interactions for drug discovery
- **Transfer learning** for drug discovery
- Machine learning and druggability prediction for **target prioritisation**

*Please note that these projects are not guaranteed and are subject to the availability of an appropriate supervisor.

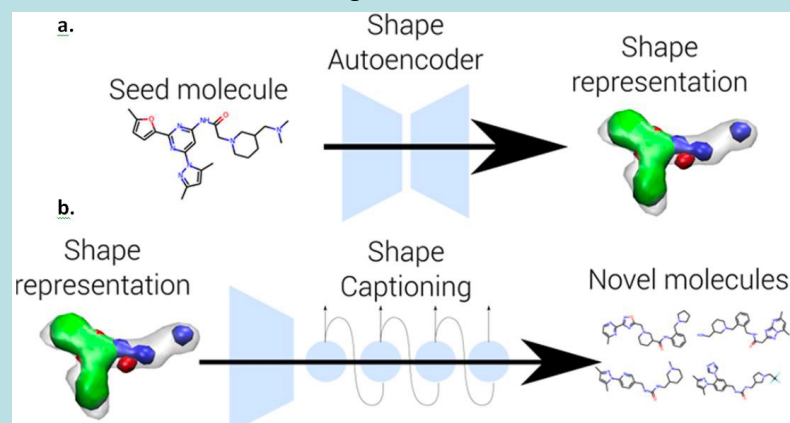
'A combined machine learning and synthesis approach to the generation of new cannabinoid mimics with therapeutic potential'

PROJECT AIM:

Adopt ML methods to generate novel libraries of original ligands that have affinity for proteins already known to be receptive to cannabinoids. Target new chemical space.

ML generation of target libraries:

- Combine Structure Based & Ligand Based Drug Design e.g. LIGDREAM, GENERATIVE, LiGANN
- Create workflow to generate library, screen for binding affinities & consider drug likeness.

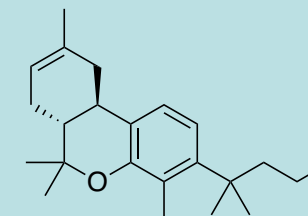


CRJ GROUP



Target based organic synthesis:

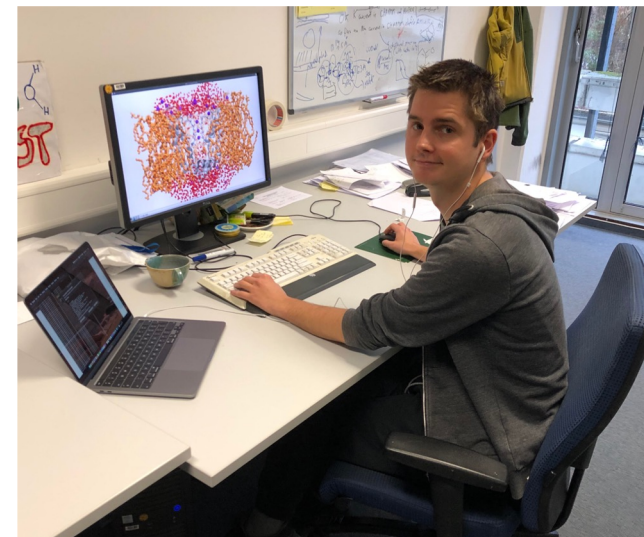
- Prepare selected targets, based on predicted binding affinity & assessed ease of synthesis.
- Standard lab techniques: e.g. reflux, extractions, inert conditions, chromatographic purification.
- Standard analytical techniques: e.g. NMR, IR, mass spectrometry.



JWH-133
(classic synthetic cannabinoid)

Reducing Cardiac Cytotoxicity with Computational Methods

Supervisor: Dr. Wojciech Kopec

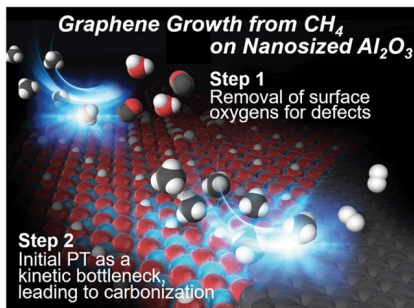


- Molecular Dynamics / Computational Electrophysiology Simulations
- Free Energy Calculations
- Machine Learning / Artificial Intelligence for Small Molecule Design

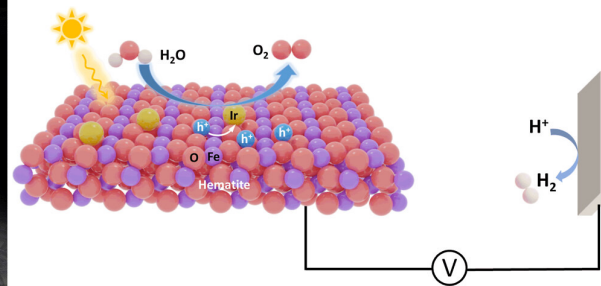
DOI: 10.3389/fphar.2019.01572 ; 10.1016/j.jmb.2021.167002 ; 10.1021/acs.jctc.2c00752



Catalytic activation of small molecules

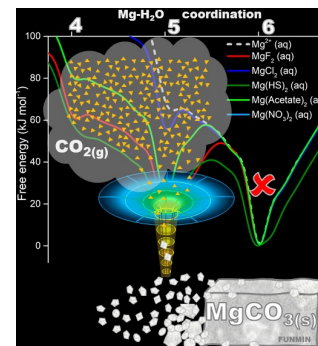


Chemical Sciences, 2022

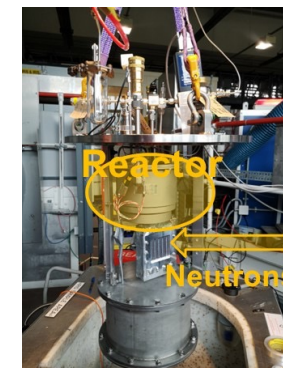


JACS, 2022

CO₂ mineralization

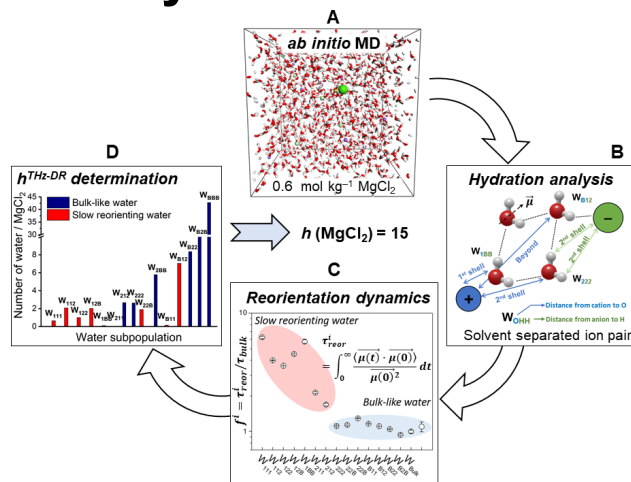


CrystEngComm, 2021



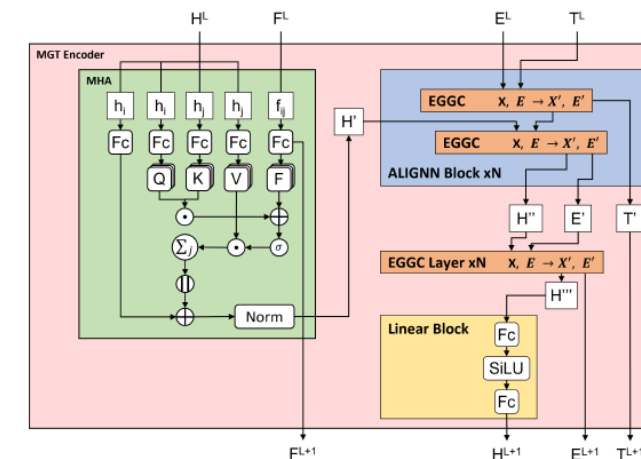
Rev. Sci. Instrum., 2023

Thermodynamic models for electrolyte solutions



J. Chem. Phys., 2022, 156,024502.

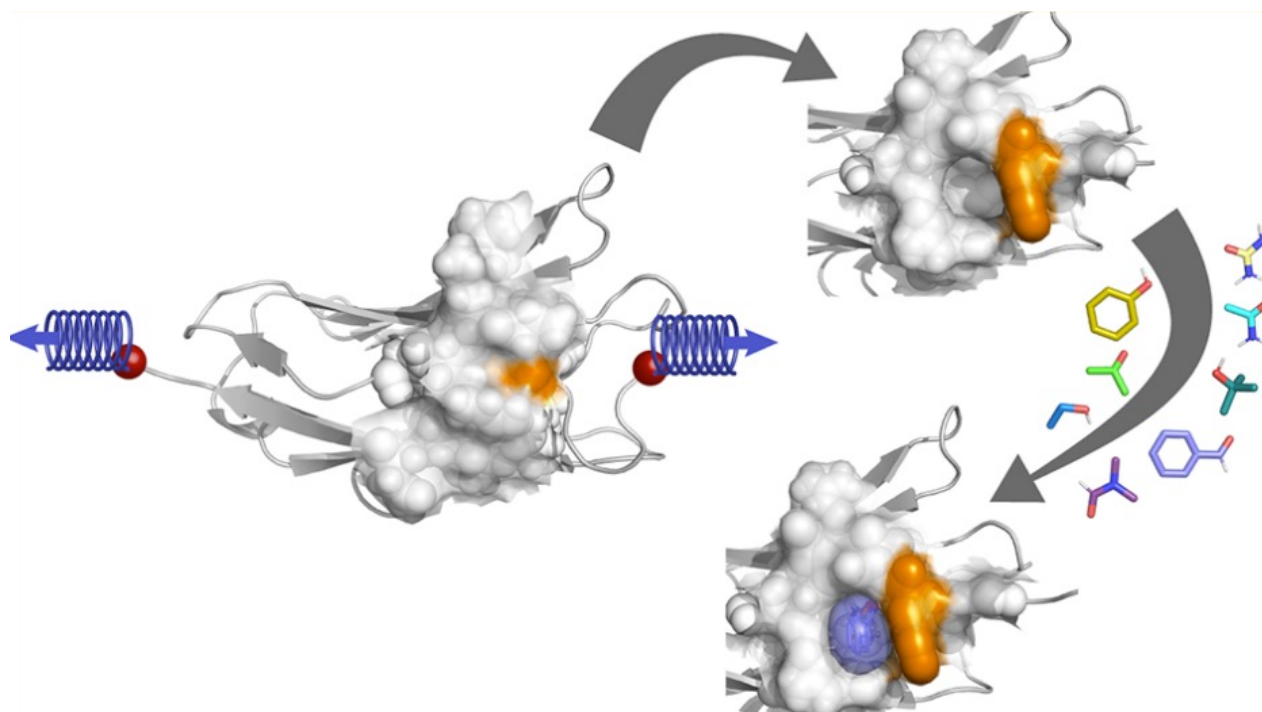
Ab initio machine learning force fields





Fornili's research group

- Machine learning for the **classification of drug binding sites** on targets
- **Generative models** for drug discovery in skeletal and cardiac muscle disease



PLoS Comput. Biol. 19: e1011099
Biophys. J. 122:54–62
J. Chem. Inf. Model. 60:6438–6446
PLoS Comput. Biol. 13: e1005826
J. Chem. Theory Comput. 15:1

Techniques: molecular dynamics, homology modelling, molecular docking, virtual screening, machine learning, R/Python coding

Career paths

Technical skills at the end of the programme:

- Coding in Python + TensorFlow, Keras, DeepChem, RDKit
- Ability to work in Linux environments
- Advanced use of drug discovery suites (such as Schrodinger Maestro)
- Ability to use High Performance Computing facilities
- Use of state of the art molecular modelling software (such as DeepMind's AlphaFold)
- Ability to design, implement and evaluate Machine Learning/Deep Learning pipelines to analyse drug discovery data and develop predictive tools
- Ability to critically evaluate the impact of medicinal chemistry on drug discovery and development.
- Deep understanding of the principles and applications of medicinal chemistry

Possible career paths:

- Computational Drug Discovery Scientist
- AI consultant for Pharmaceutical Companies
- Data Sciences & AI Graduate Programmes
- Computational Chemist
- Data Scientist

The computational skills developed can be applied to a wide range of sectors and industries.

MSc in Artificial Intelligence for Drug Discovery

Queen Mary University of London

More Info:

<https://www.qmul.ac.uk/postgraduate/taught/coursefinder/courses/artificial-intelligence-for-drug-discovery-msc/>

E-mail: a.fornili@qmul.ac.uk

