Metal-ML: Learning how to regulate metallo-proteins

As part of ongoing bacterial and fungal warfare in nature, bacteria have evolved to express proteins to destroy antibacterial agents secreted by fungi, such as penicillin. Humans have long exploited these antibacterial agents as therapeutics in the form of antibiotics, drastically reducing mortality from infectious diseases. Overuse of antibiotics has promoted the emergence of resistance in bacteria, leading to the current antimicrobial resistance (AMR) crisis [1].

One way in which bacteria acquire resistance is through the expression of β-lactamases, an enzyme that attacks a common pattern in antibiotics. We are particularly interested in a class of β-lactamases called metallo β-lactamases (MBL). MBLs belong to a large group of proteins (~30% of all proteins) called metalloproteins that require metal ions to carry out their biological function. Aside from providing bacteria with antibiotic resistance, metalloproteins provide numerous important biological functions, such as signal transduction and energy storage.
In this project, we will develop a machine learning model for metalloproteins that will provide quantum mechanical accuracy to model enzyme inhibition and mutational effects, replacing costly quantum mechanical calculations with machine learning models. We will be leveraging data generated through QCArchive and the OpenForceField initiative and build on our own conditioned diffusion models SILVR [2] to identify important mutational sites and guide the design of new ways to inhibit MBLs.

Research Challenges:
Metallo-proteins are challenging to model due to their metal centres, as classical simulations methods, albeit accurate for most protein systems, face challenges due to the metal centres needing quantum mechanical treatment. To be able to model metal lo-protein systems at scale the costly quantum mechanical calculations for each system need to be replaced with faster and cheaper neural network approaches, such as neural network potentials. The challenge will be to design metal-ml a toll that will replace QM calculations with machine learning approaches based on large scale quantum calculations provided by project patterns. This will then pave the way to study challenges around antimicrobial resistance faster and more reliably by for example designing new inhibitors against MBLs.

Data and Methodology:
The project will make use of a collaboration with QCArchive and the OpenFroceField initiative for help with the acquisition of data for training ML models from QM data. These datasets will build on SPICE [4] which is already released and curated by QCArchive. The neural network potential will then be combined with our diffusion model approaches (SILVR) [3] to help with the generation of new ideas for substrates,inhibitors and mutational challenges in metallo- β-lactamases.

Responsible research and innovation:
The project does not have any immediate ethical concerns as it works with artificial data and will only generate in silico data. The generative model that will result may be able to produce harmful molecules (in silico). The student will be encouraged to think of ways to address potential misuse of generative models for molecule generation in the future. For example, considering combining these with automatic synthesis of novel molecules. While training models will be potentially costly these models will open up opportunities that replace even more costly quantum calculations that would be required at scale to be able to do the same work.

Expected outcomes and impact:
The outcomes include a tool—Metal-ML—that will allow the modelling of metalloproteins at quantum mechanical accuracy using machine learning to replace costly quantum computations. This will form the basis to then be able to model enzymatic reactions with metalloproteins, assess the effect of mutations on inhibitor design and help guide the design of substrates and inhibitors to these proteins. The proof of concept of this work will be carried out on MBL. The potential

impact will be broad from new inhibitors against antimicrobial resistance targets and the incorporation of the developed tools in industry pipelines.

References:

[1] J O'Neil Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. Accessed: https://bit.ly/2NbICLw
[2] Runcie and Mey Journal of Chemical Information and Modeling 2023, 63, 19, 5996-6005
[3] Friedrich et al. Nature Materials, 20, 750–761 (2021)
[4] Eastman et al. Scientific Data 10, 11 (2023)