The Institute of Cancer Research

PHD STUDENTSHIP PROJECT PROPOSAL : MRC iCASE SCHEME

PROJECT DETAILS	
Project Title:	Computational approaches to enable optimisation of small molecule binding by perturbing and interacting with water networks.

SUPERVISORY TEAM	
Primary Supervisor(s):	Swen Hoelder
Associate Supervisor(s):	Mirco Meniconi
Industry supervisor:	Mike Bodnarchuck
Backup Supervisor:	Bissan Al-Lazikani

DIVISIONAL AFFILIATION	
Primary Division:	Cancer Therapeutics
Primary Team:	Medicinal Chemistry team 4
Other Team (if applicable):	In silico medicinal chemistry team / Medchem 1

PROJECT PROPOSAL

BACKGROUND TO THE PROJECT

When small molecules bind to proteins they perturb a network of water molecules present in the unbound state of the protein. For example, some water molecules will be replaced, others will engage in interactions with the small molecule. This perturbance of the water network fundamentally affects the overall free binding energy of drug binding. Despite this fundamental effect, its contributions to and effect on the overall binding affinity of the small molecule remain incompletely understood thus limiting the predictive power of structure-based design approaches e.g. when designing ligands for unoccupied pockets and apo structures.

The overall aim of this PhD project is to make contributions to the understanding of how interfering with water networks affects the thermodynamics of small molecule binding. To achieve that, we will initially select a drug / target system where the water network in the apo as well as in the ligand bound form of the protein are well characterised through high-resolution crystal structures. We will use different *in silico* tools to calculate the binding free energy of key water molecules both in the apo and drug bound form to assess to which extend this change in the water network affects the overall binding energy. We will complement these predictions with isothermal calorimetry (ITC) measurements to determine thermodynamic footprint of key binders. We will then combine the small molecule SAR, the predictions of the binding free energy of key water molecules, crystal structure information and the ITC measurements to propose a quantitative model of drug binding to this particular protein that includes the perturbation of the water network.

PROJECT AIMS

- The overall aim of this PhD project is to make contributions to the understanding of how interfering with a water networks affects the thermodynamics of small molecule binding to proteins
- An intermediate goal is to focus on a model drug / target system and analysing in depth how perturbation of a water network contributes to and affects the binding of the small molecules and publish the conclusions

RESEARCH PROPOSAL

As described in the introduction, we will start the project by focusing on a model drug / target system and analysing in depth how perturbation of a water network contributes to and affects the binding of the small molecules. The first step is thus to select a drug / target system that will be the subject of the analysis. This system has to meet a number of requirements to be suitable for this project. Firstly, high resolution (1.5A and below) structures of the apo form and different drug bound forms are essential to allow characterising of the water network in the bound and unbound form. Secondly, small molecule SAR for at least one series of compounds is necessary to analyse how changes in the water network affects the overall binding affinity of small molecules. Thirdly, the protein must be amenable to ITC. In addition, to prospectively validate our findings and conclusion through the design and synthesis of new compounds, it is beneficial to work with a protein where SAR can be derived and additional crystal structures can be solved.

To select this drug/target system, we will initially compile a database of putative targets from the public domain but also include (legacy) targets from the ICR and Astra Zeneca. Working on an internal (ICR or Astra Zeneca) target has the advantage that in-house expertise and experimental protocols such as assays and crystallography conditions already exist. The first milestone within the project is thus establishing this database to enable the selection of one or more model system for our analysis. Moreover, the database may represent a valuable resource to investigate the contribution of perturbing water networks across several distinct proteins e.g. through using machine-learning tools (see below).

In the next stage of the project, we will select one drug/target system for detailed analysis. We will use different algorithms to characterise the water network in the apo and the ligand bound state thus gaining an understanding of the energetic consequences of the changes in the water-network upon drug binding. In the first instance, we will use 3D RISM, SZMAP and a Molecular Dynamics based package (e.g. Watermap) to predict how tightly the individual water molecules are bound. In addition, we will use ITC to characterise the enthalpic and entropic contributions of the binding of key compounds to the protein.

We will then incorporate these predictions and experimental data in an analysis of the small molecule SAR. We will particularly analyse if the conclusion from the computational analysis provide additional insights and a more granular view of the small molecule SAR. Moreover, we will develop a comprehensive model that rationalises the observed trends including the perturbations of the water network. We will then investigate if free energy perturbation (FEP) and / or Grand Canonical Monte Carlo simulations (GCMC) correctly recapitulate the observed changes in the water network upon drug binding.

We estimate that these first stages of the project will be completed towards the end of the first / beginning of the second year and that we will publish the results from this first stage of the project. The project can then be continued alternative directions. The direction that we will take will depend to some extend on the preference and skill set of the student.

One exciting option is to continue the project is to build on the database of ligand/protein systems compiled during the first stage of the project and to explore machine and deep learning algorithms to derive rules and insights on how water networks can be perturbed across different proteins.

The second option is to further deepen the analysis of the model system in the first stage of the project by using quantum mechanics based algorithms called fragment molecular orbital (FMO). The FMO method is a general ab initio method that can be applied to studying large molecular systems.

The final option to continue the project is to prospectively validate the results of the first stage of the project by designing compounds that perturb the water network in a distinct way. For example by first translating our conclusion how the water network can be perturbed into a series of pharmacophore queries. In a second steps, these queries will then be used to select compounds from a virtual library constructed around the structure of the small molecule ligand. This is the option with the strongest emphasis on experimental methods and design and synthesis. There is also the potential to conduct this arm of the project in collaboration with experimental scientists at the ICR.

Supervision of the project

At The ICR, the project will be jointly supervised by Professor Swen Hoelder and Dr Mirco Meniconi. The student will be situated both in the *in silico* chemistry team and the Medicinal Chemistry Team 4 at the CRUK Cancer Therapeutics Unit and benefit from the significant experience in drug design and medicinal chemistry in both teams. In addition, in the cause of the project we will have to opportunity to obtain input from other experts at The ICR, for example Dr. Rob van Montfort (crystallography) and Prof. Bissan Al-Lazikani (bioinformatics and data science). In addition, Dr Mike Bodnarchuk from Astra Zeneca will be part of the supervisory team and contribute his extensive expertise in water modelling to the project. Moreover, the student will spend time within the Astra-Zeneca computational chemistry group at their research site.

Benefit of conducting this project in the context of a collaboration with Astra Zeneca

Astra Zeneca is one of the leading pharmaceutical companies with a strong emphasis on small molecule discovery. Astra Zeneca also has strong track record in structure-based design and *in silico* methods. This expertise will benefit the project in a number of ways for example by contributing insights and learning about perturbation of water networks observed within Astra Zeneca projects as well as giving the opportunity to include Astra Zeneca data and drug target systems in our analysis. In addition, given the strong emphasis on small molecule drug discovery, the outcome of this project will also be relevant to in silico chemist at Astra Zeneca.

PLEASE INDICATE WHETHER THIS PROJECT ALIGNS WITH ANY OF THE MRC STRATEGIC SKILLS PRIORITY AREAS?

☒ Quantitative skills ☒ Interdisciplinary skills ☐ Whole organism physiology

See the link below for further details: https://mrc.ukri.org/documents/pdf/mrc-strategic-skill-priorities/

LITERATURE REFERENCES

CANDIDATE PROFILE

Note: the ICR's standard minimum entry requirement is a relevant undergraduate Honours degree (First or 2:1)	
Pre-requisite qualifications of applicants: e.g. BSc or equivalent in specific subject area(s)	Candidates must have a first class or upper second class honours BSc Honours/MSc in chemistry or computational science or related disciplines.
Intended learning outcomes (including those arising from the industry collaboration):	• Computational chemistry, particularly detailed insights into methods to characterise the role of water molecules in small molecule binding • Structure based design • Biophysical methods There is also the opportunity to apply and gain insights into some of the following topics • Docking and pharmacophore modelling • Molecular dynamics • Quantum mechanics particularly FMO • Organic synthesis • Machine learning
Potential publications arising from project:	Given the relevance of this topic for small molecule drug discovery, there will be the potential to publish in high impact medicinal chemistry and drug design journals.
Estimated amount and distribution of time spent with industrial partner:	To be agreed with successful candidate
ADVERTISING DETAILS	
Project suitable for a student with a background in:	☐ Biological Sciences ☐ Physics or Engineering ☒ Chemistry ☐ Maths, Statistics or Epidemiology ☒ Computer Science ☐ Other (provide details)
Keywords:	**1.** Drug design
	2. Molecular Modelling
	3. Computational chemistry
	4. Medicinal chemistry
	5. Cancer Therapy
	6.