The Institute of Cancer Research

PHD STUDENTSHIP PROJECT PROPOSAL:

PROJECT DETAILS

<table>
<thead>
<tr>
<th>Project Title:</th>
<th>In silico discovery of the degrons and the degron-mediated recruitment modules of the human Ubiquitin E3 Ligase system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Project Title:</td>
<td>In silico discovery of the degrons and the degron-mediated recruitment modules of the human Ubiquitin E3 Ligase system</td>
</tr>
</tbody>
</table>

SUPERVISORY TEAM

<table>
<thead>
<tr>
<th>Primary Supervisor(s):</th>
<th>Norman Davey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associate Supervisor(s):</td>
<td>Ylva Ivarsson (University of Uppsala)</td>
</tr>
<tr>
<td>Backup Supervisor:</td>
<td>To be decided</td>
</tr>
<tr>
<td>Lead contact person for the project:</td>
<td>Norman Davey</td>
</tr>
</tbody>
</table>

DIVISIONAL AFFILIATION

<table>
<thead>
<tr>
<th>Primary Division:</th>
<th>Cancer Biology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Team:</td>
<td>Short Linear Motifs</td>
</tr>
</tbody>
</table>

PROJECT PROPOSAL

BACKGROUND TO THE PROJECT

The addition of ubiquitin to a protein can affect multiple aspects of its function and thereby fine-tune its activity. For example, modification of a target protein with a ubiquitin chain, catalysed by a ubiquitin ligase, can lead to its degradation, alter its localisation or regulate its interaction with other proteins. Ubiquitin-mediated processes are frequently deregulated in human diseases ranging from cancer to neurodegeneration. A common theme arising from the few enzymes that have been studied in-depth is that they recognise short linear interaction motifs (SLIMs - also known as consensus motifs or degrons in the case of ubiquitin ligases) within their targets, a recognition principle likely to be widely applicable to E3 ligases and DUBs.
PROJECT AIMS

- Develop and apply in silico approaches for substrate recognition module discovery
- Apply the substrate recognition module discovery tool to human Ubiquitin E3 ligases to build a set of high confidence predictions for degron-binding proteins and test these predictions using Proteomic Peptide Phage Display (ProP-PD)
- Develop and apply in silico machine learning approaches to integrate diverse omics datasets related to E3s experimental data from within the project, publicly available data and in silico predictions
- Apply these in silico approaches to define a high-confidence conditional motif-mediated interactome of the human Ubiquitin system
● Create a community resource for the high-confidence conditional motif-mediated interactome of the human Ubiquitin system

RESEARCH PROPOSAL

The proposed PhD studentship is part of the UBIMOTIF MSCA H2020 Innovative Training Network titled “Short linear interaction motifs as specificity determinants in the ubiquitin system – discovery, mechanisms and therapeutic opportunities”. The project includes 15 students from eleven groups across nine countries (Figure 2).

![Partners and Beneficiaries of UBIMOTIF](image)

*Figure 1. Participants in UBIMOTIF, the MSCA H2020 Innovative Training Network “Short linear interaction motifs as specificity determinants in the ubiquitin system – discovery, mechanisms and therapeutic opportunities” project.*

There have been two barriers to the discovery, functional analysis and exploitation of SLiMs recognised by ubiquitin ligases and DUBs: (i) lack of efficient methods for systematically identifying SLiMs interacting with globular domains in ubiquitin ligases and DUBs, and (ii) limited interactions between researchers focusing on SLiMs and those exploring ubiquitin ligases and DUBs. The UBIMOTIF will overcome these barriers by (i) exploiting recent advances in technology spearheaded by consortium members that enables the high-throughput discovery of SLiMs binding to globular domains and integrating this with novel substrate discovery methods and (ii) uniting SLiM and ubiquitin researchers in a highly interdisciplinary consortium with a common goal.

The overarching scientific objective of UBIMOTIF is to identify, characterise and exploit SLiMs that mediate interactions with ubiquitin ligases and DUBs (Davey et al. 2012; Davey and Morgan 2016). This goal will be addressed through two focused but interrelated scientific objectives:

- To identify SLiMs, the specificity determinants of SLiMs and interactors for ubiquitin ligases and DUBs.
- To understand how fundamental biological processes are regulated by ubiquitin ligases and DUBs through SLiM-mediated interactions.

The ICR studentship on the grant will be central to both objectives. The student has three key roles in the project. Firstly, using big data, evolutionary and structural analysis they will define targets for screening from the roughly 700 E3s of the ubiquitin system. Secondly, they will integrate the data produced by the project from the various sources
The PhD student will develop and apply approaches for E3 substrate recognition module discovery (Di Fiore et al. 2015).

Output 1: Motif-binding binding pocket prediction - Develop and apply in silico approaches for motif-mediated substrate recognition module discovery.

The PhD student will develop and apply approaches for E3 substrate recognition module discovery.

As part of the UBI-MOTIF project, the student will spend 3 months (months 12-15) in the Ivarsson lab (University of Uppsala, Sweden) to design and analyse the ProP-PD experiments used to test the substrate recognition module discovery predictions.

Output 2: Ubiquitin system data integration - Develop and apply in silico big data approaches to define the high-confidence conditional motif-mediated interactome of the Ubiquitin system (Davey et al. 2017; Hertz et al. 2016).

The PhD student will develop and apply approaches to integrate curated PPI data and data collected by the interactomic and proteomic screens performed by the UBIMOTIF consortia including data from AP-MS; proteomic phage display (ProP-PD); BioID, protein stability (MPS) profiling, and motif-binding pocket specificity determinant modelling. They will integrate the data produced by the project with publicly available interactomics data and discriminatory information for motif functionality using machine learning methods to produce a map of the motif-mediated interactions of the Ubiquitin system (Davey et al. 2017; Krystkowiak and Davey 2017; Hertz et al. 2016). The student will also assist in the benchmarking and optimisation of the interaction screens produced during the UBIMOTIF project.

The development of tools for the integration of SLiM-mediated PPI data of E3 Ubiquitin ligases will include:

- integrating diverse omics datasets related to E3s and DUB experimental data from within the project, publicly available data and in silico predictions
- developing machine learning approaches to train classifiers based on the experimentally validated motif instances
- applying these approaches to the interaction data to build an amino acid resolution map of the conditional motif-mediated enzyme-substrate interactions of the Ubiquitin system

As part of the UBIMOTIF project, the student will spend 3 months (months 24-27) in the Khmelinskii lab (Institut für Molekulare Biologie, Mainz, Germany) to integrate multiplexed protein stability (MPS) profiling data into the data analysis pipeline.

Contingency plan: If the experimental data from within the project is delayed the in silico data analysis tools will be built on the publicly available data and applied to novel data as it is produced. Pilot data for the project has already been produced.

Output 3: Dissemination - Create a web-based community resource of motif-related data for the Ubiquitin system from the UBIMOTIF project (Kumar et al. 2019).
LITERATURE REFERENCES


CANDIDATE PROFILE

SPECIAL REQUIREMENTS

• Applicants can be of any nationality.
• If the applicant’s first language is not English:
  They must be able to demonstrate a proficiency in English to the equivalent of an IELTS score of 7.0, with a minimum of 6 in any one component;
  -or- within the last two years in a majority English speaking country, have either education experience in English for a minimum of 1 year; or work experience in English for a minimum of 18 months and be able to satisfy Home Office visa criteria where necessary.

• Applicants must be eligible to enroll in a PhD programme at the ICR.
• Being a PhD student in an ITN includes a considerable amount of travelling including participation in meetings and workshops and complete two 3 month secondments outside the host laboratory (at the University of Uppsala, Sweden and the IMB Mainz, Germany). Therefore, a requirement is the willingness to travel and stay abroad.

Candidates will be required to meet the Marie Skłodowska-Curie Early Stage Researcher eligibility criteria:

• Mobility Rule: researchers must not have resided or carried out their main activity (work, studies, etc.) in the United Kingdom for more than 12 months in the 3 years immediately before the recruitment date. Compulsory national service, short stays such as holidays, and time spent as part of a procedure for obtaining refugee status are not taken into account. The earliest possible recruitment date is 1 March 2020.
• Early Stage Researcher (ESR) criteria: ESRs must, at the date of recruitment by the host organisation, be in the first four years (full-time equivalent research experience) of their research careers and have not been awarded a doctoral degree. Full-Time Equivalent Research Experience is measured from the date when the
researcher obtained the degree entitling him/her to embark on a doctorate (either in the country in which the degree was obtained or in the country in which the researcher is recruited, even if a doctorate was never started or envisaged).

- **You will only be considered for the position if you fulfill (and can prove so) the above eligibility criteria.**

The project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 860517”

Note: the ICR’s standard minimum entry requirement is a relevant undergraduate Honours degree (First or 2:1)

<table>
<thead>
<tr>
<th>Pre-requisite qualifications of applicants:</th>
<th>Candidates must have a first class or upper second class honours BA or BSc Honours/MSc or equivalent in computational biology or computer sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. BSc or equivalent in specific subject area(s)</td>
<td>- Structural and evolutionary analysis of proteins</td>
</tr>
<tr>
<td></td>
<td>- Integrative analysis of diverse proteomics datasets</td>
</tr>
<tr>
<td></td>
<td>- Machine learning for data integration and filtering</td>
</tr>
<tr>
<td></td>
<td>- Biology of short linear motifs and the Ubiquitin system</td>
</tr>
<tr>
<td></td>
<td>- Best practices for biological software development</td>
</tr>
</tbody>
</table>

**ADVERTISING DETAILS**

**Project suitable for a student with a background in:**

- [ ] Biological Sciences
- [ ] Physics or Engineering
- [ ] Chemistry
- [ ] Maths, Statistics or Epidemiology
- [x] Computer Science
- [ ] Other (provide details)