ICR The Institute of Cancer Research

PhD Project Proposal

Funder details

Studentship funded by:	ICR

Project details

Project title:	Novel mass spectrometry-based approaches for peptide interactomics

Supervisory team

Primary Supervisor:	Norman Davey
Associate Supervisor(s):	Jyoti Choudhary
Secondary Supervisor:	Jon Pines

Divisional affiliation

Primary Division:	Cancer Biology
Primary Team:	Short Linear Motifs
Site:	Chelsea

Project background

Many protein–protein interactions are mediated by short linear motifs (SLiMs; 3–10 amino acid stretches; typically found in intrinsically disordered regions; low-to-mid micromolar affinities) (1). The human proteome is expected to contain tens of thousands of SLiMs. However, to date only ~5000 SLiM instances have been discovered (2). The field requires scalable and accurate experimental methods to tackle SLiM discovery. Several approaches have been developed for large scale screening of SLiM-binding domains against libraries of peptides to define the SLiM-mediated interactomes. However, until recently the opposing screen, finding a SLiM-binding domain partner of a motif using a SLiM-containing peptide as bait at a medium/high throughput scale has been difficult. However, the recent development of the MS-based peptide interactomics methods PRISMA and nHU has opened the field for large scale screening of peptides (3,4,5). The major aim of this PhD proposal is to develop the experimental and computational framework required to uncover the elusive SLiM-mediated interactome of the human proteome using PRISMA and nHU. The resulting framework will be a blueprint for large scale studies to characterise the function of IDRs on a proteome-wide scale. In this project, we will focus on the key regulators of the cell cycle - up to half of the residues in the cell cycle proteins are intrinsically disordered and the vast majority of these regions have not been functionally characterised (1). The proposal will apply the complementary PRISMA and nHU approaches to sets of putative SLiM-containing peptides from key cell cycle proteins. The main output of the project will be the definition of an optimal experimental design for data collection, a framework for the analysis of mass spectrometry-based peptide interactomics data and prioritisation of biologically relevant binders.

Project aims

- Develop framework for the analysis of mass spectrometry-based peptide interactomics data and define the optimal experimental design to collect high quality peptide interactomics data.
- Develop methods for the data integration of SLiM-mediated PPI data from complementary PPI methods and datasets to define confidence in given interactions.
- Predict the conditionality of PPIs within the SLiM-mediated interactions based on the integration of complementary data.
- Apply the approach to human cell cycle proteins to discover novel motif mediated interaction in the human proteome
- Produce a resource for the dissemination of SLiM data produced from mass spectrometry-based peptide interactomics.

Research proposal

The project will apply a cutting-edge MS-based approaches to discover and characterise functional SLIM elements in IDRs. The goal of the proposal is to produce **(i)** a framework to characterise an amino acid resolution map of the IDR-mediated interactome of the cell. The project will integrate the data produced by peptide-based screens conducted with available motif specificity and interactomics information to produce high confidence reference maps for the SLIMs.

The proposal has three work packages:

Project 1: Develop framework for the analysis of mass spectrometry-based peptide interactomics data.

The PhD student will develop a framework for large-scale screens using the MS-based peptide interactomics PRISMA and nHU approaches (4,5).

- develop methods to process MS data produced by peptide interactomics screens.
- **design and screen a set of peptides with known binding partners to allow comprehensive benchmarking of the framework.**
- define the statistical framework or machine learning based approaches to extract the signal from biologically relevant binders from background noise by integrating information across interactomics screens for distinct baits.
- define the optimal and most cost-effective experimental design to collect high quality peptide interactomics data by applying leave one out approaches to controls and replicates.

Project 2: Integration, annotation, and prioritisation of SLiM-mediated PPI data

The PhD student will develop and apply approaches to integrate data collected by the peptide interactomics screens with complementary datasets to define confidence in given interactions (7,8).

Subproject 2.1: *The development of tools to integrate a priori peptide-binding knowledge into peptide MS data analysis*

- leverage the MoMAP database (http://slim.icr.ac.uk/momap/), the largest publicly available repository of motif instances and motif specificity determinants hosted at the ICR, to annotate known binders in MS peptide screens and define potential direct interactions at the amino acid resolution.
- the tool will scan bait peptide-binding prey for known motif binding domains, collect the relevant binding peptides and specificity determinants for these preys from MoMaP, and analyses the bait peptide for similarity to binding peptides and the presence of binding determinants indicative of direct motif-mediated interactions.
- software will be made available as an integrated tool in the MoMaP resource and as open-source software.

Subproject 2.2: *The development of pipeline tools for the integration and prioritisation of SLiM-mediated PPI data will include:*

- integrating motif specificity determinant data from the software developed in **Subproject 2.1** into the MS peptide interactomics analysis pipeline developed in **Project 1.**
- develop methods for the analysis and integration of MS peptide interactomics data with complementary interactomics datasets including full-length interactomics, structural studies and phage display screens (6).
- Build simple pipeline to perform brute force AlphaFold-multimer docking of bait peptides with prey proteins in screens without clear binding partners from a priori information to define novel peptide domain interfaces.
- integrating the available PTM, interactomic, proteomic and genomic data with the available motif information to predict the spatiotemporal and cell state conditionality of cell cycle motif-mediated interactions that allow integrative decision-making in cell regulation.
- creating a condition-specific motif-mediated interactome of proteins to produce testable predictions regarding the gain and loss of interactions in response to cell state perturbations.

Project 3: Peptide interactomics of human cell cycle proteins

The PhD student will apply the developed framework to a set of peptides from the cell cycle to define the binding partners of highly conserved regions in the IDRs of cell cycle proteins.

- Design and screen a set of **cell cycle** peptides.
- Process data and define conditionality of peptides to define PTM-based switching of the interactions using the framework developed in Project 2.
- Create a web-based resource to disseminate peptide interactomics data produced by the project.

Literature references

(1) “Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation.” Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE*. Chem Rev. 2014*. Jul 9;114(13):6733-78.

(2) “A million peptide motifs for the molecular biologist.” Tompa P, Davey NE, Gibson TJ, Babu MM*. Mol Cell. 2014 Jul 17;55(2):161-9.*

(3) “The next wave of interactomics: Mapping the SLiM-based interactions of the intrinsically disordered proteome”. Davey NE, Simonetti L, Ivarsson Y. *Curr Opin Struct Biol. 2023 Jun;80:102593.*

(4) “Protein Interaction Screen on a Peptide Matrix (PrISMa)” Perez-Hernandez D, Jones M, Dittmar G*Methods Mol Biol . 2023;2690:269-280*

(5) “Native holdup (nHU) to measure binding affinities from cell extracts Zambo B, Morlet B, Negroni L, Trave G, Gogl G. *Sci Adv . 2022 Dec 21;8(51):eade3828.*

(6) “Proteome-scale mapping of binding sites in the unstructured regions of the human proteome.” Benz C, Ali M, Krystkowiak I, Simonetti L, Sayadi A, Mihalic F, Kliche J, Andersson E, Jemth P, Davey NE, Ivarsson Y. *Mol Syst Biol. 2022 Jan;18(1):e10584.*

(7) “A Conserved Motif Provides Binding Specificity to the PP2A-B56 Phosphatase”. Hertz EPT, Kruse T, Davey NE, López-Méndez B, Sigurðsson JO, Montoya G, Olsen JV, Nilsson J. *Mol Cell. 2016 Aug 18;63(4):686-695.*

(8) “Systematic Discovery of Short Linear Motifs Decodes Calcineurin Phosphatase Signaling.” Wigington CP, Roy J, Damle NP, Yadav VK, Blikstad C, Resch E, Wong CJ, Mackay DR, Wang JT, Krystkowiak I, Bradburn DA, Tsekitsidou E, Hong SH, Kaderali MA, Xu SL, Stearns T, Gingras AC, Ullman KS, Ivarsson Y, Davey NE, Cyert MS. *Mol Cell. 2020 Jul 16;79(2):342-358.e12.*

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:

B.Sc in biological science or computational biology
Experience with big data analysis / MS data analysis

Intended learning outcomes:

- MS data analysis
- Large scale data analysis
- Motif biology
- Peptide interactomics

Advertising details

Project suitable for a student with a background in:

- ☒ Biological Sciences
- ☐ Physics or Engineering
- ☐ Chemistry
- ☐ Maths, Statistics or Epidemiology
- ☒ Computer Science