The Institute of Cancer Research

PHD STUDENTSHIP PROJECT PROPOSAL

FUNDER DETAILS	
Studentship funded by:	Medical Research Council - Doctoral Training Partnership (MRC DTP)
Funder specific requirements:	All MRC DTP students will attend taught courses one day a week for the first nine months of the PhD. This training will cover computational and thematic science training as well as core and transferable skills. Students will spend the remainder of the four years on their PhD project full time with monthly cohort activities.
PROJECT DETAILS	
Project Title:	Charting cancer interactomes at structural resolution using chemical crosslinking and mass spectrometry
Short Project Title:	**Charting Cancer Interactomes at Protein Interface Resolution**
SUPERVISORY TEAM	
Primary Supervisor(s):	Professor Jyoti Choudhary
Associate Supervisor(s):	Dr. Theodoros Roumeliotis
Secondary Supervisor:	Dr. Norman Davey
DIVISIONAL AFFILIATION	
Primary Division:	Cancer Biology
Primary Team:	Functional Proteomics

SHORT ABSTRACT

Large numbers of cancer mutations have been exposed through genomics; however, it remains unclear how most cancer genes function and exert influence on disease processes. Here we propose to develop a quantitative approach to map protein interactions on global scale at high throughput using cross-linking mass spectrometry (XL-MS). The method will be used to chart the native protein interaction landscapes across cancer cell lines, generating structural information on protein topology and protein-protein interaction (PPI) surfaces. A deeper understanding of how somatic mutations affect protein structure, function and rewiring of cellular networks, will provide mechanistic insights and underpins new intervention strategies.

BACKGROUND TO THE PROJECT

Protein-protein interactions (PPIs) are fundamental molecular events that regulate cellular behaviour. These biophysical associations, mediated via stretches of amino acid sequences can be stable or transient and dynamically regulated by post translational modifications and protein topology. Changes in protein sequence, post-translational state and abundance are key attributes that modulate PPIs. PPI networks encapsulate information on cellular organisation and functional relationships.

Cancer genomics has exposed a large repertoire of mutational events, however to better understand disease biology and identify therapeutic opportunities we need to assess context dependent functional consequences of mutations in a high-throughput and quantitative manner (Hoadley KA, 2018). Some driver mutations have been shown to alter protein structure and interactions leading to reorganisation of cohesive biological complexes that influence cellular behaviour. Our lab has been focused on developing mass spectrometry based methods to map native protein interactions (Pardo 2010, Bode 2016, Zhu 2019) and have applied these techniques to study disease associated genes (Purcell 2014, Lee 2020). Recently, we established approaches for global interactome analysis using co-fraction correlation analysis (Hillier 2017) and large-scale protein expression correlations (Roumeliotis 2017). The later approach was used to construct a protein relationship network in colorectal cancer cell lines, revealing how prominent cancer driver mutations can lead to collateral effects on proteins they are associated with through cohesive PPIs. Applying protein correlation analysis on patient samples, we revealed that distinct molecular remodelling events underpinned clinical trajectories that could guide the selection of candidates as biomarkers for patient stratification or therapeutic intervention (Chen et al 2020). To extend the functional and structural characterisation of cancer driver mutations in a global and high throughput fashion we plan to develop a quantitative mass spectrometry method using chemical reagents; this will enable the differential mapping of protein interactions, in the cancer proteins networks.

PROJECT AIMS

- Establish quantitative global interactome analysis.
- Chart the protein interactome landscape across cancer cell lines that encompass prominent somatic mutations.
- Develop data analysis methods to identify changes in network properties.
- Discover attributes for PPI using molecular modelling and bioinformatics.
- Validate interactome changes and their functional consequences.

RESEARCH PROPOSAL

The hypothesis of this project is that the somatic mutations remodel protein topology and binding interfaces, either directly through sequence changes or indirectly through regulatory changes, thereby altering their physical protein interactions that lead to a rewiring of functional modules in context specific cellular networks, which consequently shapes disease trajectories.

To address this, we will develop a quantitative methods using advanced mass spectrometry and bioinformatics analysis. The characterization of protein interaction on a large scale, will be applied to chart native protein interaction networks in a panel of cancer cells. These protein attributes will be used to delineate molecular relationships.

Key Aims are

1. Establish quantitative global interactome analysis

Protein immunoprecipitation in combination with mass spectrometry (IP-MS) is a powerful approach to identify endogenous protein interactions (Huttlin et al, 2017). However, IP-MS is limited to the study of interactions around predefined target proteins with available efficient antibodies, does not provide topological information and is not suitable for high throughput global characterization of protein interactions. As an alternative approach, protein crosslinking (XL) using bifunctional chemical reagents that stabilize proximal protein interfaces followed by proteolytic cleavage and MS identification which has the potential to map protein interactions in a global and high throughput fashion with some degree of structural resolution (Yu & Huang, 2018), has been successfully applied to study isolated protein complexes (Kao et al, 2011) or whole native cell lysates (Liu et al, 2015). However, technical challenges including enrichment efficiency of crosslinked peptides and the ability to confidently identify them has hindered the achievement of its full potential.

To establish a robust and high coverage quantitative method all aspects of the workflow will be considered; crosslinkers will be evaluated, different enrichment approaches, including liquid chromatography and click chemistry, followed by optimisation of sensitivity and specificity of mass spectrometry at both data capture and analysis will be explored. In addition, to characterise protein interaction changes and dynamics in cancer cells, we will establish high precision quantitation of the interaction pairs using peptide quantitation (Burke et al, 2015). Together these developments will generate a workflow providing quantitative interactome at high coverage with low FDR that can be deployed at high throughput.

2. Apply quantitative global interactome in cancer cell lines

The developed quantitative methods will be applied in a panel of cancer cell lines representing diverse human cancer subtypes for which whole genome and exome sequencing have detailed the somatic mutation status.. Moreover, molecular data including RNA sequencing and Proteomics (generated by our group) are available. Genetic screens, cellular and drug response data have also been systematically collected providing deep characterisation.

The aim is to better understand the effect of mutations on protein interactions landscape at baseline conditions, and with specific perturbations such as irradiation included to assess dynamic features. Quantitative PPI measurements will enable comparative analysis of the interactomes between conditions and cell lines.

3. Interaction Network assembly and analysis.

The data in AIM2 will be used to generate *de novo* interactome networks. These will be annotated using public experimental data to discriminate known and novel interactions. Data analysis methods will be developed to characterize the global interactome and detect stable or state dependent PPIs. The acquired data will be integrated with published data (Roumeliotis 2017) to validate correlations of novel protein interactions and specific findings will be validated in AIM5.

4. Discover structural attributes for PPI using molecular modelling and bioinformatics.

Mass spectrometry is a powerful technology capable of yielding structural insights on native cellular protein topology on a global scale and interaction interfaces at amino acid resolution. Data generated, will be compared with model protein structures from structural databases (Tunyasuvunakool 2021).

5. Validate PPIs and their functional consequences.

The method will be transferable with minor modifications to other cellular contexts such as CRISPR-cas9 gene deletions or cellular systems such as organoids (and if possible Patient Derived Xenografts (PDX)) providing additional validation. Selected findings will be followed up using alternative approaches for interaction mapping such as co-

immunoprecipitation or binary reporter assays. Further follow-up could include *in vivo* imaging and functional assays depending on the exact interaction pairs.

Overall, with this project we aim to provide a differential interaction map across a panel of cancer cells at structural resolution.

LITERATURE REFERENCES

Bode D, Yu L, Tate P, Pardo M, Choudhary JS*. (2016) Characterization of Two Distinct Nucleosome Remodeling and Deacetylase (NuRD) Complex Assemblies in Embryonic Stem Cells. Mol Cell Proteomics. 15(3):878-91.

Chen YJ, Roumeliotis TI, Chang YH, Chen CT, Han CL, Lin MH, Chen HW, Chang GC, Chang YL, Wu CT, Lin MW, Hsieh MS, Wang YT, Chen YR, Jonassen I, Ghavidel FZ, Lin ZS, Lin KT, Chen CW, Sheu PY, Hung CT, Huang KC, Yang HC, Lin PY, Yen TC, Lin YW, Wang JH, Raghav L, Lin CY, Chen YS, Wu PS, Lai CT, Weng SH, Su KY, Chang WH, Tsai PY, Robles AI, Rodriguez H, Hsiao YJ, Chang WH, Sung TY, Chen JS, Yu SL, Choudhary JS, Chen HY, Yang PC, Chen YJ.(2020) Proteogenomics of Non-smoking Lung Cancer in East Asia Delineates Molecular Signatures of Pathogenesis and Progression. Cell. 182(1):226-244.

Hillier C, Pardo M, Yu L, Bushell E, Sanderson T, Metcalf T, Herd C, Anar B, Rayner JC, Billker O, Choudhary JS*. (2019) Landscape of the Plasmodium InteractomeReveals Both Conserved and Species-Specific Functionality. Cell Rep. 28(6):1635-1647

Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, Shen R, Taylor AM, Cherniack AD, Thorsson V, Akbani R, Bowlby R, Wong CK, Wiznerowicz M, Sanchez-Vega F, Robertson AG, Schneider BG, Lawrence MS, Noushmehr H, Malta TM; Cancer Genome Atlas Network, Stuart JM, Benz CC, Laird PW. (2018) Cell of Origin Patterns Dominate the Molecular Classification of Tumors from 33 Types of Cancer. Cell. 2018;173(2):291-304.

Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Cannon, J. R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M. P., Parzen, H., Szpyt, J., Tam, S., Zarraga, G., Pontano-Vaites, L., Swarup, S., White, A. E., Schweppe, D. K., Rad, R., Erickson, B. K., Obar, R. A., Guruharsha, K. G., Li, K., Artavanis-Tsakonas, S., Gygi, S. P. & Harper, J. W. (2017) Architecture of the human interactome defines protein communities and disease networks. Nature, 545(7655), 505-509.

Kao, A., Chiu, C. L., Vellucci, D., Yang, Y., Patel, V. R., Guan, S., Randall, A., Baldi, P., Rychnovsky, S. D. & Huang, L. (2011) Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. Mol Cell Proteomics, 10(1), M110 002212.

Krystkowiak I, Davey NE. (2017) SLiMSearch:a framework for proteome-wide discovery and annotation of functional modules in intrinsically disorder regions. Nucleic Acids Res. 3;45(W1):W464-W469.

Lee YR, Khan K, Armfield-Uhas K, Srikanth S, Thompson NA, Pardo M, Yu L, Norris JW, Peng Y, Gripp KW, Aleck KA, Li C, Spence E, Choi TI, Kwon SJ, Park HM, Yu D, Do Heo W, Mooney MR, Baig SM, Wentzensen IM, Telegrafi A, McWalter K, Moreland T, Roadhouse C, Ramsey K, Lyons MJ, Skinner C, Alexov E, Katsanis N, Stevenson RE, Choudhary JS, Adams DJ, Kim CH, Davis EE, Schwartz CE. (2020) Mutations in FAM50A suggest that Armfield XLID syndrome is a spliceosomopathy. Nat Commun. 23;11(1):3698

Liu, F., Rijkers, D. T., Post, H. & Heck, A. J. (2015) Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. Nat Methods, 12(12), 1179-84.

Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary JS*. (2010) Anexpanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell. 2;6(4):382-95.

Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, O'Dushlaine C, Chambert K, Bergen SE, Kähler A, Duncan L, Stahl E, Genovese G, Fernández E, Collins MO, Komiyama NH, Choudhary JS, Magnusson PK, Banks E, Shakir K, Garimella K, Fennell T, DePristo M, Grant SG, Haggarty SJ, Gabriel S, ScolnickEM, Lander ES, Hultman CM, Sullivan PF, McCarroll SA, Sklar P. (2014) A polygenic burden of rare disruptive mutations in schizophrenia. Nature 13;506(7487):185-90.

Roumeliotis, T. I., Williams, S. P., Goncalves, E., Alsinet, C., Del Castillo Velasco-Herrera, M., Aben, N., Ghavidel, F. Z., Michaut, M., Schubert, M., Price, S., Wright, J. C., Yu, L., Yang, M., Dienstmann, R., Guinney, J., Beltrao, P., Brazma, A., Pardo, M., Stegle, O., Adams, D. J., Wessels, L., Saez-Rodriguez, J., McDermott, U. & Choudhary, J. S. (2017) Genomic Determinants of Protein Abundance Variation in Colorectal Cancer Cells. Cell Rep, 20(9), 2201-2214.

Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Žídek A, Bridgland A, Cowie A, Meyer C, Laydon A, Velankar S, Kleywegt GJ, Bateman A, Evans R, Pritzel A, Figurnov M, Ronneberger O, Bates R, Kohl SAA, Potapenko A, Ballard AJ, Romera-Paredes B, Nikolov S, Jain R, Clancy E, Reiman D, Petersen S, Senior AW, Kavukcuoglu K, Birney E, Kohli P, Jumper J, Hassabis D. (2021) Highly accurate protein structure prediction for the human proteome. Nature 596(7873):590-596

Yu, C. & Huang, L. (2018) Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. Anal Chem, 90(1), 144-165.

Zhu F, Collins MO, Harmse J, Choudhary JS, Grant SGN, Komiyama NH. (2020) Cell-type-specific visualisation and biochemical isolation of endogenous synaptic proteins in mice. Eur J Neurosci. 51(3): 793–805.

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:	BSc or equivalent in Biological sciences or Chemistry
Intended learning outcomes:	• Quantitative large scale protein analysis • Preparative liquid chromatography • Advanced tandem mass spectrometry • Quantitative statistical data analysis and visualization • Analysis of genomic mutation data in cancer • Protein structure modelling and bioinformatics
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.** Protein interactions
	2. Mass spectrometry proteomics
	3. Cancer cell profiling