The Institute of Structural and Molecular Biology

ISMB 2023 Postgraduate Research Symposium

Programme & Abstracts

Thursday 15th & Friday 16th June

Institute of Structural and Molecular Biology

ISMB 2023 Postgraduate Research Symposium

Programme & Abstracts

Thursday 15th and Friday 16th June

ISMB Director
Professor Franca **FRATERNALI**, f.fraternali@ucl.ac.uk
ISMB Assistant Cyndy **THOOI,** c.thooi@ucl.ac.uk

Head of UCL Research Department of Structural Molecular Biology (HoD)
Professor Snezana **DJORDJEVIC**, s.djordjevic@ucl.ac.uk

Departmental Graduate Research Tutors (DGT):
Professor Andrew **MARTIN**, andrew.martin@ucl.ac.uk
Professor Ivan **GOUT** (Admissions), i.gout@ucl.ac.uk

UCL Research Department of Structural and Molecular Biology
University College London, Darwin Building, Gower Street, London WC1E 6BT
General Enquiries Tel: +44 (0)20 7679 2308
Email: tabitha.owen@ucl.ac.uk
Web: https://www.ucl.ac.uk/biosciences/departments/smb

Department of Chemistry
University College London**,** 20 Gordon Street**,** London WC1H 0AJ
Tel: +44 (0)20 7679 1003
Web: http://www.ucl.ac.uk/chemistry/

Department of Biological Sciences
Birkbeck University of London**,** G54, Extension Building**,** Malet Street**,** London WC1E 7HX
Tel: +44 (0)20 7631 6800; Fax: +44 (0)20 7631 6803
Web: http://www.bbk.ac.uk/biology/

Department of Computer Science and Information Systems
Birkbeck University of London, Malet Street , London WC1E 7HX
Tel: +44 (0)20 7631 6700; Fax: +44 (0)20 7631 6727
Web: http://vili.dcs.bbk.ac.uk/support/

Date: Thursday 15th June, 9:45 – 11:25
Location: Clore Management Centre, Torrington Square, WC1E 7JL (Map)

Talks

9:45 **Welcome Tea/Coffee** – In the Clore Foyer - (15 mins)

10:00 **Introduction by ISMB director**: Prof Franca **FRATERNALI** - (5 mins)

Session 1- Chair: Prof Franca **FRATERNALI**

10:05 Rinad **ALHEDAITHY**
Synthesising ncAA-containing therapeutic proteins in the chloroplast of Chlamydomonas - (10 mins)

10:15 Nathanael **PAGE**
Native SEC-IM-MS for the Assessment of Oligonucleotide-Leachable Interactions on conformation - (20 mins)

10:35 Ivanna **YIN**
Identifying functional clathrin assembly contacts in CHC17 clathrin - (10 mins)

10:45 Declan **COOK**
Synthesis and folding pathways of tarantula inhibitory cystine knot peptide toxins – potential scaffolds for novel therapeutics - (20 mins)

11:05 Ryan **DOWSELL**
Structural and molecular investigations of adenylyl cyclase 8 - (20 mins)

11:25 **Tea/Coffee Break** – in the Clore Foyer - (15 mins)

Date: Thursday 15th June, 11:40 – 3:15pm
Location: Clore Management Centre, Torrington Square, WC1E 7JL (Map)

Talks

Session 2 – Chair: Sanjib BHAKTA

11:40 Charity **HORNBY**
Using in silico methods to understand binding on an emerging nascent chain and aid in the structure-based design of co-translational therapeutics - (10 mins)

11:50 Gabriel **ING**
Challenges in Liquid-Phase Electron Microscopy: Visualizing Nano-Scale Dynamics of Molecular Biology - (20 mins)

12:10 Suniya **KHATUN**
Deciphering the molecular pathway driving cell competition using label-free mass spectrometry - (20 mins)

12:30 Rebecca **LEES**
The role of the merozoite surface protein 1 (MSP1) complex in P. falciparum egress from erythrocytes - (20 mins)

12:50 **Lunch Break** - (85 mins)

Keynote speaker

14:15 Prof Michael J **BLACKMAN**
Malaria parasite egress from the host red blood cell: a tale of PKG, proteases and puzzles - (60 mins)

15:15 End

Date: Friday 16th June, 9:45 – 11:15
Location: Clore Management Centre, Torrington Square, WC1E 7JL (Map)

Talks

9:45 **Welcome Tea and Coffee in the Clore Foyer - (15 mins)**

10:00 **Introduction from DGT:** Prof Andrew **MARTIN** - (5 mins)

Session 1 - Chair: Joseph **NG**

10:05 Yuting **LI**
How do genetic variants in complement factor H and complement C3 lead to rare kidney inflammatory diseases - (10 mins)

10:15 Reuben **MARTIN**
The development of a quantum mechanical based method for the use of high-throughput drug design - (10 mins)

10:25 Justin **BARTON**
Efficient and generalizable inference of T cell antigen specificity - (20 mins)

10:45 Rita **RAMALHETE**
The Type VI Secretion System effector toxicity towards Pseudomonas aeruginosa - (10 mins)

10:55 Sarah **VICKERS**
Investigating the Misfolding Pathways of Alpha 1 Antitrypsin using Ion Mobility Mass Spectrometry - (20 mins)

11:15 **Tea/break in the Clore Foyer** - (15 mins)

Date: Friday, 16th June, 11:30 – 2:45
Location: Clore Management Centre, Torrington Square, WC1E 7JL (Map)

Talks

Session - 2 Chair: Amandine MARECHAL

11:30 Vivian **WANG**
Interactions of peptides with oxidised lipid monolayers - (20 mins)

11:50 Debbie **WOODS**
Visualising the SWI/SNF chromatin remodeler in action - (20 mins)

12:10 Samadrita **CHATTERJEE**
Developing a Dedicated Predictor to Detect Pathogenic Variants in Protein Kinases - (10 mins)

12:20 Ines **ZOUHAIR**
Comparative in-situ characterisation of supercomplexes in human mitochondria - (10 mins)

12:30 Lauren **WOODBURN**
Multi-domain proteins fold co-translationally via stable intermediate states – (10mins)

12:40 Swathi **KUMAR**
The role of vFLIP in NF-kB activation – (20mins)

Poster Session

13:00 **Lunch/Poster session in the Clore Foyer** - (120 mins)

14:55 Prizes awards - (5 mins)

15:00 End

Date: Friday, 16th June, 12:40 -2:45pm
Location: Clore Management Centre, Torrington Square, WC1E 7JL (Map)

Poster Session

(Lunch available)

Posters (alphabetical order) - A – K:

Mahnaz **ABBASIAN**	*Impact of human genetic variation on COVID-19 disease susceptibility*
Christos **EFTHYMIOU**	*Molecular dynamics simulations clarify the effect of genetic variants in complement factor H and complement component 3 associated with atypical hemolytic uremic syndrome*
Lilian **DENZLER**	*TCR Sequence Numbering and Stability Prediction*
Christopher **DULSON**	*Drug Discovery Against African Swine Fever Virus*
Trupti **GORE**	*MHC Restriction: Towards T-Cell Function Prediction*
Julian **HERRERA BRAGA**	*Towards the next generation of variant impact predictors using deep learning and multiomics data*
Naail **KASHIF-KHAN**	*Mining Metagenomics Data for Novel Bacterial Nanocompartments*
Pokchut **KUSOLKUMBOT**	*Production in the algal chloroplast of a major capsid protein (MCP) subunit vaccine, Against the infectious spleen and kidney necrosis virus (ISKNV) of Nile tilapia*

Date: Friday, 16th June, 12:40 -2:45pm
Location: Clore Management Centre, Torrington Square, WC1E 7JL (Map)

Poster Session

(Lunch available)

Posters (alphabetical order) - L – Z:

Yvette **LEVRAY-SZABADOS**	*The unconventional trafficking mechanisms of exported transmembrane-like proteins in Plasmodium*
Weining **LIN**	*VariPred: Enhancing Pathogenicity Prediction of Missense Variants Using Protein Language Models*
Sarah **LOWEN**	*Characterisation of the alpha-1-antitrypsin pathological polymer using solution-state NMR spectroscopy*
Katherine **MORLING**	*Design, synthesis and characterisation of cofactor-mimicking HIV-1 capsid inhibitors*
Alice **PETTITT**	*Biophysical characterisation of ORF6 from SARS-CoV-2*
Gabriel **SCOGLIO**	*Unravelling the untapped potential of Aphanizomenon flos-aquae – a high-value cyanobacterium*
James **SWEET-JONES**	*abYdraw: A Graphical Tool Applying Antibody Mark-Up Language For the Standardisation of Multispecific Antibody Annotation*
Mollie **VIRGO**	*Investigation of microbial community dynamics in vivo using zebrafish infection models*

Abstracts
Last name order – A-Z

Impact of human genetic variation on COVID-19 disease susceptibility (Poster)

Mahnaz **ABBASIAN**, Vaishali WAMAN & Christine ORENGO

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: mahnaz.abbasian.19@ucl.ac.uk

Abstract:
Severe viral acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first reported in China in December 2019, rapidly spread globally and was responsible for the COVID-19 pandemic
In addition to differences in socio-economic and demographic factors, clinically approved host factors such as age, gender, and body-mass index in individuals with no chronic underlying health condition are correlated with a wide range of COVID-19 morbidity (1); although the influence of host genome variability is less studied.
The human protein variants may play a role in COVID-19 morbidity by enhancing the binding affinity to SARS-CoV-2, either in a receptor that facilitates viral entry into the host cell and causes infection or in a mediator of innate immune-associated pathways (2).
By employing data on the human genetic variation and viral genus from several web-resources, bioRxiv and the literature analysis (of well-studied complexes and other less studied complexes), this study aimed to understand the impact of changes in binding affinity following disease-independent amino acid substitutions in three human cell-receptor (ACE2, KREMEN1, and AXL) in individuals across different ethnicities with the SARS-CoV-2 spike protein.
Therefore, human genetic databases such as gnomAD and All of us were used to collect populations and subpopulations' genetic variations and their allele frequencies. It is crucial to comprehend how variations in human proteins affect a person's vulnerability to COVID-19.

References:
1) Hu, B., Guo, H., Zhou, P. et al. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 19, 141–154 (2021). doi.org/10.1038/s41579-020-00459-7
2) Niemi MEK, Daly MJ, Ganna A. The human genetic epidemiology of COVID-19. Nat Rev Genet. 2022 Sep;23(9):533-546. doi: 10.1038/s41576-022-00478-5. Epub 2022 May 2. PMID: 35501396; PMCID: PMC9060414.

Synthesising ncAA-containing therapeutic proteins in the chloroplast of Chlamydomonas (7 minute talk)

Rinad **ALHEDAITHY**, Harry JACKSON & Saul PURTON

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: rinad.alhedaithy.22@ucl.ac.uk

Abstract:
Genetic code expansion aims to reassign codons to incorporate non-canonical amino acids (ncAA) at specific sites in a polypeptide chain to expand the functionality of the target protein. In nature, the full potential of protein functionality has not been unlocked due to limited set of 20 proteogenic amino acids, although some archaeal species naturally contain an extended number of amino acids, which can be exploited. The incorporation of a ncAA at a specific site in a target protein relies on a translation system that has an unused codon that can be reassigned to the ncAA, an orthogonal pair of aminoacyl tRNA synthetase (aaRS) and its cognate tRNA for the ncAA, and the presence of the ncAA itself within the system.

Chlamydomonas reinhardtii is a 'generally recognised as safe' (GRAS) certified microalga, which allows for its use as a platform for the manufacturing of therapeutic proteins. The chloroplast is ideal for genetic engineering due to its small chloroplast genome that benefits from a lack of gene-silencing mechanisms, no overlapping genes, and DNA integration via homologous recombination. Importantly, the chloroplast genome does not use the UGA stop codon, allowing for its reassignment. Using synthetic biology techniques, elements of the ncAA-incorporating machineries can be engineered into the chloroplast to synthesise novel therapeutic proteins in C. *reinhardtii*.

Efficient and generalizable inference of T cell antigen specificity (15 minute talk)

Justin **BARTON**[1], Trupti GORE[1], Michele MISHTO [2 3], Adrian SHEPHERD[1]

[1]Institute of Structural and Molecular Biology, Birkbeck, University of London, London, UK
[2] Centre for Inflammation Biology and Cancer Immunology (CIBCI) & Peter Gorer Department of Immunobiology, King's College London (KCL), London, UK
[3] Francis Crick Institute,, London, UK

Email address: jbarto02@student.bbk.ac.uk

Abstract:
Predicting T cell antigen specificity poses several challenges due to the diversity of T cell receptors (TCRs) and the complexity of antigen presentation and recognition. The paucity of publicly available assay data has further frustrated efforts to computationally model specificity. To date, models have only been able to demonstrate significant predictive accuracy on a small set of epitopes for which substantial assay data is available. Attribution analyses have raised questions about whether models are learning underlying biological dynamics or exploiting confounding correlations in the data. Recent advances in transfer learning offer a potential solution to ameliorate some of these issues and pave the way for models that reflect the biology of T cell antigen recognition. In this work we set out to improve predictive accuracy on out of distribution epitopes.

Keynote Speaker: ***Malaria parasite egress from the host red blood cell: a tale of PKG, proteases and puzzles*** (60 minute talk)

Professor Michael J **BLACKMAN** PhD FMedSci

Malaria Biochemistry Laboratory, Francis Crick Institute, London NW1 1AT, UK and Faculty of Infectious Diseases, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK

Email: mike.blackman@crick.ac.uk

Abstract:
Malaria is a major cause of mortality, morbidity and poverty across much of the globe. Clinical malaria results from replication of the causative agent, a protozoan parasite, within circulating host red blood cells. Successive cycles of intraerythrocytic replication within a parasitophorous vacuole (PV) culminates in active egress of merozoites with associated red blood cell destruction and invasion of fresh red cells. Merozoite egress is a tightly controlled, explosive event that takes place over the course of just a few minutes. Egress is controlled by a parasite enzyme cascade in which activation of a single parasite cGMP-dependent protein kinase called PKG triggers the discharge of a parasite serine protease called SUB1 into the lumen of the PV. There, SUB1 rapidly activates several merozoite surface and PV-located proteins, including a cysteine protease called SERA6 which catalyses the final step of red cell membrane rupture. Maturation of SUB1 is itself mediated by an aspartic protease called plasmepsin X. In this talk I will summarise recent work in this area and describe new insights into the molecular mechanisms underlying egress as well as some of the continuing mysteries surrounding the process. Finally I will illustrate how attempts to target enzymes of the egress pathway are providing new potential approaches to antimalarial drug discovery.

Developing a Dedicated Predictor to Detect Pathogenic Variants in Protein Kinases (7 minute talk)

Samadrita **CHATTERJEE**[1], Vaishali P. WAMAN[1], Paul ASHFORD[1], Andrew C.R. MARTIN[1], Christine ORENGO[1]

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: samadrita.chatterjee.21@ucl.ac.uk

Abstract:
Many computational tools have been developed to predict the deleteriousness of genetic variations leading to residue mutations in protein sequences. These exploit both structure and sequence based features and are regularly evaluated by the CAGI independent assessment. We conducted a review of several of the best performing methods. A study in 2020 [1] suggested that performance of predictors trained for specific protein families are much more reliable than generic predictors. The Orengo Group collaborate with the Swanton Group in the Crick Institute to analyse mutations associated with lung cancer. Studies have shown that mutations in kinases play a significant role in the development and prognosis of cancer. My PhD aims to develop a specific predictor for kinases. To date, I have tested two popular and widely used generic predictors, namely CADD and MutPred2, to assess their performance. I have applied them to a set of kinase cancer mutations from TRACERx to analyze the variants. In the future two in-house predictors, SAAPpred and VariPred, will be adapted for kinase based cancer mutations. VariPred exploits protein sequence embeddings derived from protein language models and has been shown to outperform other state-of-the-art pathogenicity predictor whilst depending only on a single sequence feature.

Reference:

1. Zaucha J, Heinzinger M, Tarnovskaya S, Rost B, Frishman D. Family-specific analysis of variant pathogenicity prediction tools. *NAR GenomBioinform*. 2020;2(2):lqaa014. Published 2020 Feb 28. doi:10.1093/nargab/lqaa014

Synthesis and folding pathways of tarantula inhibitory cystine knot peptide toxins – potential scaffolds for novel therapeutics (15 minute talk)

Declan **COOK**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: declan.cook.19@ucl.ac.uk

Abstract:
Disulfide-rich peptides (DRPs), such as venom toxins, are molecules of therapeutic interest. Their disulfide linkages provide a conformational constraint that ensures the bioactive conformation predominates. For the synthesis of these peptides, it is important to elucidate and replicate their disulfide connectivities. Current methods for disulfide installation rely on either laborious orthogonal protecting group chemistry, or the use of redox buffers that need subsequent removal. This work aims to explore the formation of disulfide bonds in wildtype Protoxin-II, a triply bridged tarantula toxin, when left to fold in water. Chemical methods to direct and control disulfide formation in water are investigated, including the use of novel pseudoprolines as removable turn-inducers. Tandem mass spectrometry (MS^n) is used to try and determine the disulfide connectivities of both wildtype Protoxin-II, and misfolded intermediates. Preliminary studies have shown that DRPs are refractory to collisional and electron-based fragmentation techniques in MS^n, prompting investigations into chemical methods that improve the fragmentation efficiency of peptides that contain disulfide linkages. Conversion of the disulfides into extended thioacetals is explored as a means of increasing collision-based fragmentation. To conclude, this work aims to (i) improve synthetic strategies to produce DRPs; (ii) explore the folding pathways of Protoxin-II; and (iii) develop MS workflows that can determine the disulfide connectivities of DRPs.

TCR Sequence Numbering and Stability Prediction (Poster)

Lilian **DENZLER**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: lilian.denzler.17@ucl.ac.uk

Abstract:
The software package TCRnum was created to reliably number T-Cell Receptor Sequences.

A robust sequence numbering method for T-cell receptors is important for all work with T-cell receptor sequence data. Universal numbering schemes and correct residue numbering is vital for sequence analysis and comparison. We present a numbering software for T-cell receptor sequences that will reliably implement a set of popular numbering schemes. The software also enables T-cell receptor sequences to be labelled using numbering schemes commonly used for antibodies, which will facilitate studying the differences and commonalities of T-cell receptors and antibodies. It will also enable antibody-based tools to be adapted to T-cell receptor sequences.

Furthermore, NLP methods are used to analyse TCR sequences and predict the stability of a construct based on sequence alone. The likelihood of a TCR having a melting temperature within a critical range is predicted. Methods for flagging stability-decreasing sequence patches are explored.

Molecular dynamics simulations clarify the effect of genetic variants in complement factor H and complement component 3 associated with atypical hemolytic uremic syndrome (Poster)

Christos **EFTHYMIOU**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: christos.dereschuk.20@ucl.ac.uk

Abstract:
Abstract - Molecular dynamics (MD) simulations have proven to be a valuable tool for understanding the molecular mechanisms underlying disease. In this study, genetic variants in the complement proteins C3d and CFH (SCR 19/20) associated with atypical hemolytic uremic syndrome (aHUS) were studied with MD simulations. We evaluated the ability of MD simulations to detect changes in the binding interaction and behaviour of the complement proteins C3d and CFH via controls and compared the results to simulations of the aHUS-associated mutations D1119N (CFH SCR 19), I1169T (CFH SCR 20), and G1224D (C3d). We employed a combination of structural, energetic, and dynamic information to obtain a comprehensive view of the impact of the mutations on protein function and classify variants as benign or malignant. Our results showed that MD simulations can be used to accurately predict the functional consequences of mutations, matching known classifications. The simulations surpassed the predictive capabilities of SIFT and Polyphen-2, which regularly predicted conflicting classifications compared to each other and experimental results. This study demonstrates the potential of MD simulations for the classification of aHUS-associated mutations as either deleterious or benign. These findings have important implications for the development of more accurate and reliable predictive tools for disease diagnosis and treatment, and highlight the potential of MD simulations for guiding experimental validation of the functional effects of mutations.

Structural and molecular investigations of adenylyl cyclase 8 (15 minute talk)

Ryan **DOWSELL**

Email address: ryan.dowsell.19@ucl.ac.uk

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Abstract:
Adenylyl cyclases (AC) catalyse the biosynthesis of cAMP, a ubiquitous second messenger essential for most biological processes. The synaptically enriched AC8 is activated by calmodulin but the structural mechanisms underlying this interaction are unclear. We have co-purified AC8 with calmodulin for cryo-EM and crosslinking mass spectrometry to address this. Furthermore, a key site of regulation may exist at the 12 transmembrane helices, which comprise ~40% of the polypeptide but have no known function besides subcellular anchoring. To investigate this, we have developed a FRET-based plate reader assay using purified AC8 and isolated catalytic domains to screen for compounds that bind at the transmembrane domain to regulate catalytic activity. The transmembrane domain, along with the N and C terminal domains may also be important sites for protein interactions. We have fused AC8 to the biotin ligase, TurboID, for proximity labelling in primary hippocampal neurons to identify novel protein interactors. Collectively, our findings will elucidate how cAMP is spatiotemporally regulated at the synapse.

Drug Discovery Against African Swine Fever Virus (Poster)

Christopher **DULSON**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: christopher.dulson.20@ucl.ac.uk

Abstract:
African swine fever virus (ASFV) poses a significant threat to global food security, and currently there are no approved vaccines or antiviral therapies. ASFV requires a viral DNA dependent RNA polymerase for viral transcription which we have successfully expressed and purified recombinantly. The core ASFV viral RNA polymerase is active in non-specific *in vitro* transcription assays which have been developed for each product of the reaction. One assay is based on an absorbance measurement relative to the pyrophosphate product and the other assay is based on a fluorescence measurement relative to the RNA product. The intention is to screen derivatives of known RNA polymerase inhibitors, compound libraries and nucleotide analogues. Another avenue for discovery of novel therapeutics is though disrupting the protein-protein interactions that lead to the assembly of protein complexes that are crucial to ASFV transcription. Nano-luciferase complementation assays have been developed to quantify the dimerization of the viral RNA polymerase assembly platform subunits and the early transcription factors. Peptide and compound libraries will be screened for the disruption of the dimer. Additionally, the recombinantly expressed and purified early transcription factors were also discovered to have ATPase activity which will be targeted for drug discovery with nucleotide analogues.

MHC Restriction: Towards T-Cell Function Prediction (Poster)

Trupti **GORE**[1,2], Justin BARTON[1] & Adrian SHEPHERD[1]

1 Birkbeck College, 2 University College London

Email address: trupti.gore.20@ucl.ac.uk

Abstract:
T cell function prediction is a very challenging problem because it involves predicting how the six flexible loops of a TCR bind to an epitope. A further complication arises from the fact that a TCR does not simply bind to an epitope on its own but to a complex formed between the epitope and an MHC molecule (the latter's function is to present epitopes on the surface of host cells for T cell inspection). Hence, accurate MHC prediction, i.e., predicting which MHC molecule(s) a given TCR can bind to, is likely an essential precursor for TCR function prediction.

This task differs from TCR-epitope binding prediction in two important aspects. Firstly, different parts of the TCR's CDR loops play a dominant role in MHC binding compared to epitope binding, with the contributions of different CDRs varying between TCR-pMHC complexes.[1] Secondly, a single individual's CD8+ T cell repertoire can engage with a small number of different MHC molecules, which is fixed, whereas the potential number of epitopes that the same repertoire engages with is vast and changes over time. The goal here is to develop a deep learning method that addresses the *MHC restriction prediction task* and contributes to the broader goal of T cell function annotation.

References:

Burrows, S. R. *et al.* Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10608–10613 (2010).

Towards the next generation of variant impact predictors using deep learning and multiomics data (Poster)

Julian **HERRERA BRAGA**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: julian.braga.22@ucl.ac.uk
Supervisor: Professor Franca Fraternali (f.fraternali@ucl.ac.uk)

Abstract:
The classification of missense variants regarding their consequences to pathogenicity is a complex task. The presence of variants does not necessarily indicate pathogenicity. Typically, sequence conservation and variant incidence are features used in variant impact predictors, but these are often inconclusive. Features derived from other levels of biological complexity such as transcription, protein structure, dynamics, protein abundance and interactions could be informative, but they have not been systematically integrated and applied in variant impact prediction. The aim of this research is to investigate the application of machine learning to a multiomic database covering these new features to achieve predictive capabilities of missense variant impact. I am currently building an ontology-based multiomics graph database in Neo4J using BioCypher framework with the aim to integrate common variations data from gnomAD and pathogenic variations from ClinVar, and investigate the contribution of transcriptomics, proteomics and interactomics data in discerning variant impact. I have investigated, as a case study, protein structural data, by extracting features which distinguish physiological protein dimeric complexes from artefacts of X-ray crystallographic experiments. Using a simple logistic regression model, I found that hydrophobicity of residues alone contains good signal (AUC = 0.745) to identify physiological interfaces. Building on this, I am applying classical and deep machine learning techniques onto the graph multiomics database, with the goal of classifying damaging variations.

Using in silico methods to understand binding on an emerging nascent chain and aid in the structure-based design of co-translational therapeutics. (7 minute talk)

Charity **HORNBY**, Julian STREIT, Tomek WLODARSKI, Lisa CABRITA & John CHRISTODOULOU

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: charity.hornby.16@ucl.ac.uk

Abstract:
During translation the nascent chain can go from an unfolded polypeptide to its folded native state and along this folding path form intermediate states in a mechanism known as co-translational folding[1]. This translation window could be a target for therapeutics to bind to a protein which would be particularly significant for proteins that are hard to drug after release from the ribosome. The structures that form along the co-translational pathway could reveal pockets/epitopes that may have been buried in the native state and no longer available for binding. However, structural understanding of these states on the ribosome is difficult to study experimentally. Therefore, if there is to be structure-based design of therapeutics for an emerging nascent chain the use of *in silico* methods could be a vital tool to supplement experimental work by generating high resolution ensembles of the nascent chain which can then be used to screen virtual libraries and identify early hits which can be tested *in vitro*. The aim of this work is to put together a computational strategy where from structural ensembles of the ribosome-nascent chain complex that characterise unfolded and intermediate states they can be used to identify early hits and be used for structure-based optimisation of experimental hits. This strategy can then complement the ongoing experimental effort to study the mechanism of co-translational folding and test nascent chain binders.

References:

1. Chan, S.H.S., Włodarski, T., Streit, J.O. *et al.* The ribosome stabilizes partially folded intermediates of a nascent multi-domain protein. *Nat. Chem.* **14**, 1165–1173 (2022).

Challenges in Liquid-Phase Electron Microscopy: Visualizing Nano-Scale Dynamics of Molecular Biology (15 minute talk)

Gabriel **ING**[1], Andrew STEWART[1], Lorena RUIZ-PEREZ[2] & Guiseppe BATTAGLIA[2]

1. University College London, Department of Chemistry, Christopher Ingold Building, 20 Gordon Street, London,WC1H 0AJ
2. Institute for Bioengineering of Catalonia, IBEC, C. Baldiri Reixac, 10-12 · 08028 Barcelona

Email address: gabriel.ing.19@ucl.ac.uk

Abstract:
In the last two decades, electron microscopy of liquid samples has been developed into an effective and impressive method for visualising nano-scale dynamics in real time. While the developments in cryo-EM have been revolutionary in molecular biology, liquid-phase electron microscopy (LPEM) has primarily focused on inorganic systems due to difficulties and limitations in the technique. These challenges include creating an effective liquid cell for the system and limiting radiation damage. Recent advanced in LPEM have expanded the scope for LPEM for studies of biomolecules.

In my talk, I will discuss my attempts to develop the use of the technique by imaging ferritin and apoferritin in the liquid-phase and the challenges this yielded. Apoferritin is an iron transport protein, which can be filled with iron ions to produce ferritin. While iron-containing ferritin is easily seen, the protein component has proved difficult to visualise by LPEM due to low contrast.

Finally, I will show results from an applied study on amyloid-β aggregation, a dynamic system highly associated with Alzheimer's disease. In this system, the peptide amyloid-β aggregates into larger structures called oligomers and then fibrils, but the pathway for this aggregation process is unclear. Attempts to image this process has shown formation of liquid-droplets, this result has since been published by another method. These preliminary results demonstrate LPEM can provide useful information on biological samples.

Mining Metagenomics Data for Novel Bacterial Nanocompartments (Poster)

Naail **KASHIF-KHAN**[1,2], Stefanie Frank[1] & Renos Savva[2]

1. Department of Biochemical Engineering - UCL
2. Department of Biological Sciences - Birkbeck, University of London

Email address: naail.kashif-khan.20@ucl.ac.uk

Abstract:
Encapsulins are icosahedral protein compartments found in prokaryotes. They encapsulate specific cargo proteins and are implicated in several physiological functions, including stress response, iron storage, and secondary product metabolism. These virus-like nanocompartments have many promising uses in biotechnology, synthetic biology, and the development of vaccines and gene therapies. To date, there has not been a comprehensive survey of metagenomics databases to uncover encapsulin sequences. In this work, a dataset of novel encapsulin sequences from the MGnify Protein Database is presented. Bioinformatics and deep learning methods show that these encapsulins may have new, previously unseen biological function, and encapsulate cargo proteins that have not been associated with encapsulin genomic loci in the literature. Large-scale structure prediction of these encapsulins shows that they may also exhibit novel structural features and new fusion domains. This dataset may be a valuable resource for future encapsulin engineering experiments, and raises questions about the evolution of virus-like capsids.

Deciphering the molecular pathway driving cell competition using label-free mass spectrometry. (15 minute talk)

Suniya **KHATUN**

Institute of Structural and Molecular Biology, Division of Bioscience, University College London, London, UK

Email address: suniya.khatun@ucl.ac.uk

Abstract:
Cell competition is phenomenon that results in the elimination of sub-optimal cells from tissue. This process is crucial for development and maintaining tissue homeostasis. It is believed that the fate of cells in competition is determined by their local cellular neighbourhood as sub-optimal cells are only eliminated in the presence of fit cells. To gain a deeper understanding of this elimination process, we employed a mass spectrometry-based proteomics approach to investigate the molecular mechanisms involved in cell competition. Specifically, we focused on studying the molecular pathways that determine the fate of $MDCK^{WT}$ (Madin-Darby Canine Kidney) cells and MDCK scribble depleted cells ($scribble^{kd}$) during competitive interactions in a 50:50 co-culture. Our analysis revealed significant alterations in proteins associated with cellular stress response, Wnt, NF-kB, Hippo and MAPK signalling pathways. Additionally, we observed dysregulation of cell adhesion, cell-cell communication, perturbation of cytoskeletal organisation, cell migration, as well as disruption of mitochondrial function and metabolism in $scribble^{kd}$ when in competitive condition. Our findings suggest that cells engaged in competition undergo complex molecular and cellular changes, employing diverse mechanisms for growth and survival in order to gain a competitive advantage.

Production in the algal chloroplast of a major capsid protein (MCP) subunit vaccine, Against the infectious spleen and kidney necrosis virus (ISKNV) of Nile tilapia (Poster)

Pokchut **KUSOLKUMBOT**[1*], Patai CHAROONNART[2,3], Harry JACKSON[1], Sarocha JITRAKORN[2,3], Vanvimon SAKSMERPROME[2,3] & Saul PURTON[1]

[1]Algal Research Group, Institute of Structural and Molecular Biology Department, University College London, Gower Street, London WC1E 6BT, UK
[2]Center of Excellence for Shrimp Molecular Biology and Biotechnology, Mahidol University, Bangkok, 10400, Thailand
[3]National Center for Genetic Engineering and Biotechnology (BIOTEC) Thailand Science Park, Pathumthani, 12120, Thailand
Keywords: Chlamydomonas; Chloroplast; major capsid protein (MCP); infectious spleen and kidney necrosis virus (ISKNV)

Email address: pokchut.kusolkumbot.20@ucl.ac.uk

Abstract:
The chloroplast of the green microalga *Chlamydomonas reinhardtii* offers various advantages for the sustainable production of recombinant proteins. The alga is Generally Recognized as Safe (GRAS), and the chloroplast compartment represents an enclosed system that allows transgene expression to a high level and correct protein folding. The Infectious Spleen and Kidney Necrosis Virus (ISKNV) causes a severe pandemic in fish. ISKNV infection results in epidermal lesions with significant petechial haemorrhages and abdominal edema. The algal strain selected for this experiment is *C. reinhardtii* CC-4033 which is engineered to contain a bacterial gene that allows selective growth of the strain in media containing phosphite. The strain also loses chlorophyll when cultures are transferred from the light to the dark. The two features will reduce scale-up costs and palatability. A novel method to resolve downstream processing costs is to use the whole of the edible algae, allowing vaccine encapsulation in the dried algae and formulation into the aquaculture feed. Moreover, a codon reassignment strategy results in the biocontainment of the transgenes in the chloroplast and prevents horizontal gene transfer. This approach is improving the genetic features of the *C. reinhardtii* host, which will be beneficial for oral vaccine production in the future.

The role of the merozoite surface protein 1 (MSP1) complex in P. falciparum egress from erythrocytes (15 minute talk)

Rebecca A **LEES**[1,2], Claudine BISSON[3,4], Trishant UMREKER[5], Christine R COLLINS[2], Abigail J PERRIN[2,6], Michael J BLACKMAN[2,7] and Helen R SAIBIL[1]

1 Department of Biological Sciences, ISMB, Birkbeck, University of London, United Kingdom.
2 Malaria Biochemistry Laboratory, The Francis Crick Institute, London, United Kingdom
3 Centre for Ultrastructural Imaging Department, Kings College London, United Kingdom
4 Evotec, Milton Park, Oxfordshire, United Kingdom
5 Department of Life Sciences, Imperial College London, United Kingdom
6 Department of Biology, University of York, United Kingdom
7 Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, United Kingdom

Email address: rebecca.lees.19@ucl.ac.uk

Abstract:

In the intraerythrocytic cycle of the malarial parasite *P.falciparum*, merozoites invade red blood cells (RBC) in which they asexually replicate. Daughter merozoites then burst out of the RBC in a process known as egress and go on to invade new erythrocytes. Merozoite Surface Protein 1 (MSP1) is a GPI-anchored protein that decorates the merozoite surface membrane. Just before egress, MSP1 undergoes proteolytic maturation by a parasite serine protease called SUB1; the processing products remain associated at the merozoite surface where they form a complex with other malarial proteins. Both MSP1 processing and complex formation are thought to be essential for parasite egress from the RBC. However, the composition, structure and function of this MSP1 complex, and the precise role of SUB1 processing, are poorly understood. MSP1 has been suggested to be involved in degradation of the RBC cytoskeletal component b-spectrin, allowing RBC membrane rupture at egress. Here we show that conditional depletion of MSP1 results in defective egress, but has no effect on b-spectrin cleavage or rupture of the RBC and parasitophorous vacuole membrane. Electron microscopic examination revealed that defective egress of the MSP1-null mutant is likely due to a segmentation defect. Parasites expressing a mutant MSP1 refractory to SUB1 processing at a key site also display abnormal egress, suggesting cleavage at this site is important for MSP1 function. For further insight into MSP1 function, we have purified the native protein complex in both SUB1 processed and unprocessed forms for analysis by single particle cryo-EM. Initial models of the structure in comparison to the published structure indicate the presence of partner proteins. Whilst there are minimal structural differences between processed and unprocessed MSP1, differences between the composition of the complex were seen. Identifying the positions of binding of these partner proteins will help determine the role of the MSP1 complex.

The unconventional trafficking mechanisms of exported transmembrane-like proteins in Plasmodium (Poster)

Yvette **LEVRAY-SZABADOS**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: y.levray@ucl.ac.uk

Abstract:

Symptoms of malaria occur during the blood stage of infection when the malaria parasite resides inside human red blood cells. Hundreds of proteins, synthesized by the parasite, are exported into the host cell where they modify its properties. Some exported proteins become embedded in membranes and are referred to as membrane proteins. Despite their importance in disease pathology, how these membrane proteins are transported into the red blood cell is poorly understood. Some exported membrane proteins are thought to become membrane-embedded in the parasite endoplasmic reticulum and are subsequently extracted from the parasite plasma membrane before being transported into the red blood cell where they are then re-inserted into the appropriate membrane.

Contrary to this, we find that a subset of membrane proteins such as Pf332 and SBP1 enter into the lumen of the endoplasmic reticulum and do not insert into the endoplasmic reticulum membrane. This behaviour is very different from the behaviour of membrane proteins that have been studied in model eukaryotic systems and suggests, that the trafficking of exported membrane proteins is a mechanistically distinct process that may represent a unique drug target.

How do genetic variants in complement factor H and complement C3 lead to rare kidney inflammatory diseases (7 minute talk)

Yuting **LI**[1], Joshua CARMICHAEL[2], Xin GAO[1] & Stephen J. PERKINS[1,*]

[1] Department of Structural and Molecular Biology, Division of Biosciences, University College London, UK
[2] Department of Renal Medicine, Royal Free Hospital, University College London, UK
* Email: s.perkins@ucl.ac.uk

Email address: ucbtibe@ucl.ac.uk

Abstract:
Genetic variants in the complement proteins of innate immunity are associated with the development of inflammatory diseases in the kidney (termed aHUS and C3G), in particular those variants in complement factor H (CFH) and complement C3. The same also applies to age-related macular degeneration (AMD) in the eye. The mechanistic link of these variants with these diseases remains unclear. We maintain a successful interactive web database for genetic variants in CFH. In this project, we will use our most recently upgraded interactive web database from 2021 to investigate the complement variants associated with aHUS and C3G, and also AMD in order to plan and perform rational mutagenesis experiments on the interaction between CFH and C3 fragments. The mutants will be used to test the effect on the binding affinity between C3 and CFH and their stabilities, and compare these with structural predictions. This work will establish the distinct structural locations of the complement genetic variants in CFH and C3 that influence the development of either aHUS, C3G, or AMD, and will determine experimentally the effect of the variants on their molecular interactions that lead to these three different inflammatory diseases. By performing these experimental studies, we aim to clarify the molecular mechanism leading to these kidney and eye diseases, and in turn inform the targeted development of improved bioinformatics prediction methods and clinical therapeutics.

VariPred: Enhancing Pathogenicity Prediction of Missense Variants Using Protein Language Models (Poster)

Weining **LIN**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: w.lin.16@ucl.ac.uk

Abstract:
Computational approaches for predicting the pathogenicity of genetic variants have advanced in recent years. These methods enable researchers to determine the possible clinical impact of rare and novel variants. Historically these prediction methods used hand-crafted features based on structural, evolutionary, or physiochemical properties of the variant. In this study we propose a novel framework that leverages the power of pre-trained protein language models to predict variant pathogenicity. We show that our approach VariPred (Variant impact Predictor) outperforms current state-of-the-art methods by using an end-to-end model that only requires the protein sequence as input. By exploiting one of the best performing protein language models (ESM-1b), we established a robust classifier, VariPred, requiring no pre-calculation of structural features or multiple sequence alignments. We compared the performance of VariPred with other representative models including 3Cnet, Polyphen-2, FATHMM and 'ESM variant'. VariPred outperformed all these methods on the ClinVar dataset achieving an MCC of 0.727 vs. an MCC of 0.687 for the next closest predictor.

Characterisation of the alpha-1-antitrypsin pathological polymer using solution-state NMR spectroscopy (Poster)

Sarah **LOWEN**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: sarah.lowen.20@ucl.ac.uk

Abstract:
Alpha-1-antitrypsin (α1AT) is a member of the serpin superfamily best known to regulate the proteolytic activity of neutrophil elastase (NE). Mutations in the SERPINA1 gene can lead to α1AT deficiency (α1ATD) that may present as chronic obstructive pulmonary disease (COPD) and liver cirrhosis. A1AT polymers have been observed by electron microscopy as 'beads on a string' that is hypothesised to arise from a domain swap mechanism. Yet, the size and flexibility of these polymers are difficult to structurally characterise by certain structural techniques and the polymerisation pathway is still contested. I have used solution-state NMR spectroscopy to investigate the conformational dynamics and structure of α1AT polymerisation using recombinant and patient-derived samples. By completing a residue-specific assignment of a self-inserted α1AT molecule that mimics the structure of α1AT polymers we now have a sequence-wide probe for α1AT polymers at atomic detail. I have also purified α1AT polymers from ZZ liver explant and acquired a 1H fingerprint by solution state NMR. The methyl region chemical shifts for the liver-derived α1AT were identified using the residue-specific assignments from the recombinant self-inserted probe. By studying ex-vivo material I can characterise the true conformation of polymers derived directly from clinical patients and work towards elucidating the interactions between α1AT polymers that underlie α1ATD.

The Development of a Quantum Mechanical Based Method for the Use of High-Throughput Drug Design. (7 minute talk)

Reuben L. **MARTIN**[1,2], Andrea TOWNSEND-NICHOLSON[1] & Alexander HEIFETZ[2]

1. Research Department of Structural & Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, United Kingdom
2. Evotec (UK) Ltd., 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire OX14 4RZ, United Kingdom

This work is funded by the MRC DTP, in partnership with Evotec

Email address: reuben.martin.18@ucl.ac.uk

Abstract:
High-Throughput Structure-Based Drug Design (HT-SBDD) is a set of rapidly evolving tools for faster and more cost-effective methods of lead drug discovery. These sets of methods offer a computational replacement to traditional high-throughput screening drug discovery techniques. Structural data of a target protein is utilised in conjunction with large databases of drug candidates, where the most likely drug candidates are output and tested for clinical potential.

HT-SBDD is fundamentally dependent on the large-scale processing of many complex calculations, recent developments in High-Performance Computing (HPC) have enabled these calculations to be completed in manageable timeframes. This has led to the development of more refined and detailed computational techniques to better predict protein-ligand interaction energies. It is important to acknowledge that atomic-level biological interactions are not solely governed by Newtonian physics, and that quantum mechanics is responsible for a significant proportion of interaction energies. The Fragment-Molecular Orbital (FMO) method applies quantum mechanical calculations to biologically relevant targets to determine the nature and relative strengths of these non-intuitive interactions. Throughout this talk I will introduce the application of HPCs in drug design and explore the application of the FMO method, using the adenosine A1 GPCR as an example of my research.

Design, synthesis and characterisation of cofactor-mimicking HIV-1 capsid inhibitors (Poster)

Kate **MORLING**[1], Morten LARSEN[2], Justin WARNE[1], Ben GRAHAM[1], Sally OXENFORD[1], Emma TOUIZER[2], Lauren HARRISON[2], Lucy THORNE[2], Nikos PINOTSIS[3], Greg TOWERS[2], David SELWOOD[1]

1. Wolfson Institute for Biomedical Research, UCL, Gower St, London, WC1E 6BT.
2. Division of Infection and Immunity, UCL, Gower St, London, WC1E 6BT.
3. Birkbeck College, Malet St, London, WC1E 7HX.

Email address: katherine.morling.20@ucl.ac.uk

Abstract:
HIV combination therapy has been hugely successful at suppressing viral replication and extending life expectancy. However, challenges such as resistance and side effects persist, necessitating the development of drugs with new mechanisms. The conical HIV-1 capsid is formed of 250 hexamers and 12 pentamers of capsid protein. It's crucial functions during the HIV lifecycle, including protecting the viral genome and regulating reverse transcription, make it an attractive drug target. Although early capsid inhibitor PF74 inhibits DNA synthesis and disrupts cofactor interactions, it has limited potency and unfavourable metabolic properties. Recently, ultrapotent capsid inhibitor Lenacapavir was approved as a first-in-class drug, but resistance rapidly arises from point mutations.

We used rational design to derive a series of capsid inhibitors targeting a conserved host cofactor binding site. These inhibit HIV-1 infection with nanomolar potency. We solved structures of inhibitor-capsid complexes and confirmed their binding to a hydrophobic pocket between capsid monomers which interacts sequentially with cofactors Sec24C, Nup153, and CPSF6. Additionally, we discovered that these inhibitors trigger innate immune sensing during infection and influence capsid stability in *in vitro* assembly and thermal shift assays. Furthermore, we investigated the potency of our inhibitors against Lenacapavir resistant mutants. These results are guiding design of improved inhibitors with enhanced potency and reduced sensitivity to resistance mutations.

Native SEC-IM-MS for the Assessment of Oligonucleotide-Leachable Interactions on conformation (15 minute talk)

Nathanael **PAGE**[1]
Supervisor: Professor Konstantinos THALASSINOS[2,3]*

[1] Drug Development Solutions Ltd trading as Resolian, Newmarket Rd, Fordham, Ely CB7 5WW. Nathanael.Page.17@ucl.ac.uk
[2] Institute of Structural and Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, United Kingdom
[3] Institute of Structural and Molecular Biology, Birkbeck College, University of London, London, WC1E 7HX, United Kingdom

* **Contact email:** k.thalassinos@ucl.ac.uk
Email address: nathanael.page.17@ucl.ac.uk

Abstract:
Oligonucleotide therapeutics (ONTs) have become invaluable tools in the ongoing battle with contagious diseases and genetic malfunction. However, as with any therapeutic compound, there are a range of opportunities for ONT related impurities to be introduced into the drug product. These impurities can often be detected and removed as part of production through purification processes, however, when these products are stored or prepared for use there is an additional opportunity for impurities to be formed through interactions with components leached from container closure system.

The components which can be leached from container closure systems are known as Leachables and can vary greatly in physio-chemical properties. As a result, these components have the potential to interact with oligonucleotides in a variety of ways, some of which impact the function of the product and thus the effectiveness of the therapy. While there are analytical approaches capable of measuring Impurities in relatively short oligonucleotides, these approaches are denaturing and are traditionally challenging to use for the characterisation of conformation impurities in ONTs who's function is reliant on its conformation. This talk will discuss novel native SEC-IM-MS approaches with the potential to simultaneously detect and elucidate the impact of Leachables on the safety and function of ONTs.

Biophysical characterisation of ORF6 from SARS-CoV-2 (Poster)

Alice J. **PETTITT**[1,2], Lydia NEWTON[1], Vaibhav Kumar SHUKLA[1], Stephen MCCARTHY[3], Alethea B. TABOR[3], Gabriella T. HELLER[1], Christian D. LORENZ[2], D. Flemming HANSEN[1]

1. Department of Structural and Molecular Biology, Division of Biosciences, UCL Darwin Building, Gower Street, London, WC1E 6BT
2. Department of Physics, Faculty of Natural, Mathematical and Engineering Sciences, King's College London, Strand, London, WC2R 2LS
3. Department of Chemistry, Faculty of Mathematical and Physical Sciences, UCL

Email address: alice.pettitt.16@ucl.ac.uk

Abstract:
The 61-residue accessory protein ORF6 from SARS-CoV-2 is highly flexible and binds to host proteins, which is believed to suppress the antiviral immune response. Despite the importance of ORF6 as a potent interferon antagonist, there is currently a lack of structural data available for this protein. To characterise ORF6 and the ORF6 C-terminal region ($ORF6_{CTR}$) in the unbound state, we are using a combination of biophysical techniques including nuclear magnetic resonance spectroscopy (NMR), small-angle X-ray scattering (SAXS) and molecular dynamics (MD) simulations with the enhanced sampling technique metadynamics. Thus far, NMR chemical shifts have been used as restraints to improve force field accuracy (metadynamic metainference). Here, we present molecular scale detail on the conformational sampling of the $ORF6_{CTR}$ and provide a comparison of our MD simulations to experimental results. We also use our approach to validate the MD simulations by expressing and purifying isotopically labelled ORF6 and the $ORF6_{CTR}$. Our work highlights the importance of integrating MD simulations with experimental data, such as NMR chemical shifts, which thus far has allowed us to show that the $ORF6_{CTR}$ ensemble samples a mainly disordered state. We anticipate that by characterising the $ORF6_{CTR}$ in the unbound and binding-competent state, insights into the mechanisms of ORF6 can be gained.

The Type VI Secretion System effector toxicity towards Pseudomonas aeruginosa (7 minute talk)

Rita **RAMALHETE**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: r.ramalhete@mail.bbk.ac.uk

Abstract:
The type VI secretion system (T6SS) is a contractile nanomachine used by Gram-negative bacteria to deliver toxic effector proteins into adjacent cells. The T6SS of *Vibrio cholerae* has been shown to successfully kill several Gram-negative species. However, *Pseudomonas aeruginosa* is a notable exception in that it is largely unaffected by *V. cholerae* T6SS attacks. Interestingly, *P. aeruginosa* is not inherently resistant to all T6SS attacks, given it is sensitive to attacks from the T6SS of other species, including *Acinetobacter baylyi*. We endogenously expressed three *V. cholerae* effectors (VgrG3, VasX and TseL) in *P. aeruginosa* to see if *P. aeruginosa* is intrinsically immune to the unique set of effectors delivered by *V. cholerae* T6SS. We found that expression of the effectors in the periplasm but not the cytosol was toxic to *P. aeruginosa*, suggesting that *P. aeruginosa* immunity to *V. cholerae* T6SS may be to the inability of the *V. cholerae* effectors to reach their periplasmic targets. To investigate whether *V. cholerae* effectors are only being delivered to the *P. aeruginosa* cytosol or if the effectors are not able to pass the outer membrane altogether, we developed a fluorescent reporter assay to quantify non-lethal delivery of T6SS substrates to the cytosolic compartment of target cells. In addition to detecting cytosolic delivery of T6SS secretion substrates, this reporter can be used to quantitatively measure what proportion of recipient cells receive T6SS attacks in multispecies bacterial communities.

Unravelling the untapped potential of Aphanizomenon flos-aquae – a high-value cyanobacterium (Poster)

Gabriel **SCOGLIO &** Saul PURTON

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: gabriel.scoglio.17@ucl.ac.uk

Abstract:
My project focuses on *Aphanizomenon flos-aquae* (AFA): a filamentous nitrogen-fixing edible cyanobacterium known for its high nutritional value. Specifically, I am working with a wild AFA strain that blooms naturally in Klamath Lake, Oregon, since this strain is already distributed worldwide as a nutritional supplement. I am further investigating the commercial and biotechnological potential of two AFA strains from culture collections: CCAP AFA and NIES AFA. Overall, my project is divided into three main aims: 1) whole genome sequencing (WGS) and characterisation; 2) cultivation optimisation with the goal of commercial exploitation; and 3) AFA genetic engineering for the generation of bespoke strains with improved traits for biotechnology.

The WGS has been obtained for both the Klamath and CCAP strains. From the obtained genetic data, proof-of-concept experiments to transform CCAP AFA have been attempted. Specifically, a method to deliver antibiotic resistance genes into CCAP AFA's genome via either natural transformation or electroporation is currently being developed and honed. In terms of cultivation, focus has been placed on CCAP AFA: its growth has been optimised to achieve the fastest doubling time to date (<30h) and has been demonstrated to grow well in 10L photobioreactors, specifically a low-cost bubble column PBR made from cheap polythene tubing, reaching final biomass values of 2–3 g/L. Next steps include optimisation of growth of the Klamath and NIES AFA strains in the laboratory, CCAP AFA growth evaluation in other photobioreactor systems, from open ponds to flat panels (particularly in relation to its production of C-phycocyanin, a high-value pigment), and the generation of a CCAP AFA transformant line.

The role of vFLIP in NF-kB activation (Poster)

Swathi **KUMAR**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: swathilsk@outlook.com

Abstract:
Kaposi Sarcoma-associated Herpes Virus (KSHV) causes Kaposi Sarcoma (KS), Primary Effusion Lymphoma (PEL) and Multicentric Castleman Disease in immunocompromised individuals. Owing to the rarity of KS and PEL, there is no effective treatment for late-stage disease, leaving a need for the development of novel therapeutics.

Constitutive activation of the canonical NF-kB pathway is a major factor in KSHV pathogenesis where it is essential for tumour survival. KSHV produces a virally encoded oncoprotein called vFLIP which activates the NF-kB pathway by persistently upregulating the IKK kinase complex. vFLIP binds to the modulatory subunit of the IKK kinase complex known as IKKγ. Although the physical interaction between vFLIP and IKKγ has been well-characterised, it has remained unclear how vFLIP activates the two IKK kinase complex subunits IKKα and IKKβ.

Using a combination of cell-based assays, biophysical techniques, and structural biology, we demonstrate that vFLIP alone is sufficient to activate the IKK kinase complex by inducing the trans-autophosphorylation of IKKβ. We have further established that vFLIP-induced IKKβ autophosphorylation is dependent on high molecular weight, multimeric vFLIP-IKKγ assemblies.

This research has shown a novel mechanism by which the KSHV protein vFLIP activates a signalling pathway that is not only activated by other viruses, but is overactivated in the majority of inflammatory diseases and cancers.

abYdraw: A Graphical Tool Applying Antibody Mark-Up Language For the Standardisation of Multispecific Antibody Annotation (Poster)

James **SWEET-JONES**

Supervisors: Andrew MARTIN (UCL) and Franca FRATERNALI (KCL)

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: james.sweet-jones.20@ucl.ac.uk

Abstract:
Abstract: Multispecific antibodies (MsAbs) are engineered molecules that may bind to more than one type of antigen with applications in diagnostics assays and as therapeutic agents. Currently three MsAbs are FDA approved but they occupy a significant proportion of therapeutic antibodies in clinical trials. As the methods of generating MsAbs have diversified, the different formats which MsAbs may be presented in have also grown, yet no format to annotate them has become standardised in drug applications.

Using an annotation language from previous work in our group, Antibody Markup Language (AbML) we have developed abYdraw. abYdraw is a graphical interface tool allowing users to input AbML strings to draw a schematic of the MsAb format represented by the string. Alternatively, users may draw MsAb formats in order to obtain an AbML string which describes the drawn MsAb. Both AbML string and schematics may be exported for use in publication or drug approval applications. With abYdraw we aim to apply AbML as a new standard of MsAb annotation for its intuitive use and accessibility as a Windows, MacOS or Linux app.

Investigating the Misfolding Pathways of Alpha 1 Antitrypsin using Ion Mobility Mass Spectrometry (15 minute talk)

Sarah **VICKERS**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: sarah.vickers.20@ucl.ac.uk

Abstract:
Alpha-1 antitrypsin (AAT) is a serine protease inhibitor essential to the control of proteolytic pathways. AAT has a complex mechanism for inhibition involving a conformational change, leaving it vulnerable to misfolding. Consequently, patients suffer from AAT deficiency, caused by mutant AAT, leading to two complications: The mutants can misfold and polymerise causing cirrhosis. AAT thus fails to reach the lungs and inhibit elastase, leading to emphysema. Here, a combination of Ion Mobility Mass Spectrometry (IMMS) and preparative techniques were used to investigate the misfolding of AAT. We see both conformational and stability differences in the misfolding pathway relevant to AAT deficiency.

An in-depth characterisation of AAT and disease-associated mutants using a combination of native ion mobility mass spectrometry and traditional proteomics experiments was performed. The stabilising effect of a small molecule inhibitor was observed in the changes in mobility and the molecule-bound protein's response to the application of supplemental energy. The activation of cleaved AAT with supplemental energy induced the release of the 'c-terminus' peptide which was isolated and ECD was performed to characterise it.

Further to these studies of monomeric AAT, we have developed IMMS methods for isolating and studying ex vivo aggregates of antitrypsin. In our preliminary investigations of liver polymer samples using IMMS, high-order polymers and their multiple conformations were observed. Further experiments will investigate the polymerisation pathway by in-depth analysis of conformations of heat-induced, denaturant induced and liver polymers.

This work shows how the use of a variety of mass spectrometry methods can be used in combination to characterise misfolding proteins and their aggregates. The use of native ion mobility mass spectrometry and top-down methods give structural insights while proteomics experiments can provide complimentary cellular information. Future work will involve the application of these methods to Hep 1.5 cells and native mass spectrometry analysis of the cellular environment. Furthermore, intermediates on the polymerisation pathway will be investigated by ion mobility isolation and activation.

Investigation of microbial community dynamics in vivo using zebrafish infection models (Poster)

Mollie **VIRGO**[1,2], Serge MOSTOWY[2], Brian HO[1]

1. Department of Biological Sciences, Birkbeck, London, UK.
2. Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, UK.

Email address: mvirgo01@student.bbk.ac.uk

Abstract:
Multispecies microbial communities are found ubiquitously throughout nature and have crucial roles in human health, agriculture, and industry. Bacterial cell-cell interactions are some of the key factors dictating the dynamics, structure, and diversity of these communities and consequently determine how they behave as a system. One widely conserved system that mediates a large portion of such interactions is the type VI secretion system (T6SS). The T6SS is a nanomachine used by Gram-negative bacteria, including *Vibrio cholerae*, to deliver toxic effector proteins directly into target cells.

Here, we develop a multi-species zebrafish hindbrain ventricle (HBV) infection model to explore how *V. cholerae* T6SS activity impacts bacterial populations and influences host responses to infection. We found that when *V. cholerae* alone colonises the HBV there is no T6SS-depenent host response. However, following co-infection with *E. coli* we observed T6SS-depenent bacterial antagonism, activation of host inflammatory cytokines and accelerated host death. Moreover, T6SS-medatited antagonism towards *E. coli* appeared to enhance *V. cholerae's* ability to colonise the HBV. These results indicate that although T6SS does not directly affect the host, T6SS-mediated interbacterial antagonism induces host signals and enhances pathogen virulence.

Interactions of peptides with oxidised lipid monolayers (15 minute talk)

Vivian **WANG**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: vwangx01@student.bbk.ac.uk

Abstract:
Pulmonary surfactant is vital in regulating the biophysical activity of the lungs by minimising the surface tension of the air-water interface in the alveoli. Surfactant proteins such as Sp-B contribute to the biophysical function of surfactant, by facilitating the breathing mechanics where a lowered surface tension results in larger gaseous transfer during inhalation and the prevention of alveolar collapse during expiration. Due to the extreme hydrophobicity of Sp-B, an experimental structure of the full-length peptide has yet to be determined. Hence, different purification methods were explored to optimise the expression of His-6 Sp-B, which was observed to be more stable at lower pH. Circular dichroism spectroscopy data suggested that the addition of trifluoroethanol to His6 Sp-B resulted in the formation of predominantly alpha helical structures. This research also investigates the effect of oxidative damage on the interactions of a peptide mimic of Sp-B, Sp-$B_{1\text{-}25}$, with lipid monolayers using molecular dynamics simulations. Oxidatively damaged lipids resulted in significant loss of surfactant function, re-spreading and partial uncoiling of the peptide's alpha helix. In a related set of experiments, neutron scattering data showed that the peptide protegrin embeds further into the unoxidized lipid monolayers than the oxidatively damaged monolayers.

Multi-domain proteins fold co-translationally via stable intermediate states (7 minute talk)

Lauren **WOODBURN**[1], Sammy CHAN[1], Lisa CABRITA[1] & John CHRISTODOULOU[1]

[1] Institute of Structural & Molecular Biology, Department of Structural & Molecular Biology, University College London, London, UK and Department of Biological Sciences, Birkbeck College, London, UK

Email address: lauren.woodburn.16@ucl.ac.uk

Abstract:
Proteins are synthesized by the ribosome and the majority must fold into their biologically-active three-dimensional structure. Most of the proteome comprises multidomain proteins, but in contrast to single domains which can efficiently fold from their denatured states, the refolding of multidomain proteins in isolation is prone to misfolding and aggregation, processes that are associated with a variety of human diseases. During biosynthesis, progressive emergence of the nascent protein from the ribosome exit tunnel is thought to enable co-translational, independent folding of each domain, although a structural understanding of this process remains elusive. In this work, we study a model tandem repeat protein, FLN, whose folding of its fifth domain has been well-characterised both on and off the ribosome, to examine the role of its neighbouring domains in modulating its co-translational folding. Using ^{1}H-^{15}N, ^{13}C-methyl, and ^{19}F NMR spectroscopy and protein engineering, we examined the length-dependent folding of FLN4-FLN5 by varying the FLN6 linker. We find that FLN4 is fully folded before FLN5 begins folding. In both the absence and presence of FLN4, FLN5 folds via two stable partially folded intermediates that are not found in isolation of the ribosome (1); however, the addition of the preceding domain provides additional thermodynamic stability (up to ~2 kcal/mol) to these intermediates across a wider folding transition. We then systematically assessed the factors that could result in such changes, such as interactions with the highly charged ribosome surface of both the unfolded and folded domains, and proximity of neighbouring domains to each other. We thus demonstrate that domains of tandem repeat proteins fold sequentially during translation and can modulate the thermodynamics, rather than the pathway, of folding of subsequent domains.

References:

1. Chan SHS, Włodarski T, Streit JO, Cassaignau AME, Woodburn LF, Ahn M, et al. The ribosome stabilizes partially folded intermediates of a nascent multi-domain protein. Nature Chemistry. 2022;14(10):1165-73.
2. Cassaignau AME, Włodarski T, Chan SHS, Woodburn LF, Bukvin IV, Streit JO, et al. Interactions between nascent proteins and the ribosome surface inhibit co-translational folding. Nature Chemistry. 2021.
3. Burridge C, Waudby CA, Włodarski T, Cassaignau AME, Cabrita LD, Christodoulou J. Nascent chain dynamics and ribosome interactions within folded ribosome–nascent chain complexes observed by NMR spectroscopy. Chemical Science. 2021;12(39):13120-6.

Visualising the SWI/SNF chromatin remodeler in action (15 minute talk)

Debbie **WOODS**, Manoj SAXENA, Tim CRAGGS & Philip ROBINSON

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: debbie.woods.19@ucl.ac.uk

Abstract:
Chromatin remodelers are large multi-subunit complexes able to restructure the nucleosome, increasing the accessibility of DNA for processes such as transcription, repair, or replication. SWI/SNF is an evolutionarily conserved ATP-dependent chromatin remodeler that uses the energy of ATP hydrolysis to move, remodel and evict nucleosomes. Mutations in various subunits of human SWI/SNF, BAF, have been found in 20% of human tumours, highlighting the complex as one of the most commonly affected targets in cancer. Insights into the remodelling mechanism of SWI/SNF is important to understand its role in cancer. The project aims were to produce a “stalled” reaction intermediate of nucleosome bound BAF for biochemical, biophysical, and structural analysis to shed light on the remodeling mechanism. We have successfully generated a novel internal DNA-histone site-specific thiol linkage within the nucleosome and demonstrated BAF remodeling is blocked by this cross-link using a restriction enzyme accessibility assay. To follow the reaction at higher spatial and temporal resolution, a single molecule FRET system has been established. In the context of this project, these methods can be used to understand the effect of the cross-link on remodeling and to inform cryoEM structural studies to visualise transient intermediate states of BAF remodelling. The development of a “stalled intermediate” has broader applications to other fields of nucleosome study including transcription regulation and DNA replication, recombination, and repair.

Identifying functional clathrin assembly contacts in CHC17 clathrin. (7 minute talk)

Ivanna **YIN**

ISMB, Division of Biosciences, Darwin building, University College London, Gower Street, London, WC1E 6BT

Email address: ivanna.yin.18@ucl.ac.uk

Abstract:
Clathrins are vesicle coat proteins that facilitate intracellular membrane traffic in eukaryotic cells. In humans, there are two types of clathrin with differential cellular distributions and functions, which are generated respectively by clathrin heavy chains (CHC), CHC17 and CHC22. This PhD project will focus on the conventional clathrin, CHC17 clathrin, which is ubiquitously expressed in human tissues, including within the brain, where it plays an important role in neurotransmission in addition to vital housekeeping functions. Clathrin mediates these processes by self-assembling into a basket-like structure to deform membranes and transport synaptic and other protein cargo between different cellular compartments. Efficient self-assembly is essential for clathrin function in membrane traffic, and a recent cryo-EM structure of CHC17 clathrin has identified key inter-molecular contacts that drive this assembly process. Several patients with intellectual disability (ID) possess mutations in CHC17 that are close to these predicted assembly contacts. This study will test the effects of these mutations on clathrin assembly properties, including the rate and efficiency of assembly and the size and shape of clathrin baskets, using recombinant approaches and biochemical analysis. Eventually, this study will establish the functional residues in CHC17 that control clathrin assembly and contribute to the understanding of neurological diseases.

Comparative in-situ characterisation of supercomplexes in human mitochondria (7 minute talk)

Ines **ZOUHAIR** [1,2], Cristina UGALDE[3], Amandine MARÉCHAL[1] & Karen DAVIES[2]

[1] Department of Structural and Molecular Biology, University College London, London WC1E 6BT, UK
[2] eBIC, Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK
[3] Instituto de Investigacion Hospital 12 de Octubre, Madrid 28041, Spain

Email address: ines.zouhair.19@ucl.ac.uk

Abstract:
Mitochondria are the main energy source in eukaryotic cells and are at the heart of many fundamental cellular mechanisms. The protein complexes of the mitochondrial respiratory chain (MRC) haven been described to form supercomplexes (SC). To this day, little is known in regards to SC biogenesis or biological relevance. Such key elements to cellular functioning have a pressing need to be further explored and characterised, in particular as they have been linked to several pathologies.

Complex IV is the terminal oxidase of the human MRC. One of its subunits, COX7A, can exist as 1 of 3 isoforms. Previous studies on mutant human cell lines selectively expressing only one COX7A isoform showed differences in the MRC organisation and cell bioenergetic phenotypes, suggesting a direct link between COX7A isoform, SC population and metabolism.

This project aims to have an *in-situ* comparative study of the MRC in those COX7A mutant cell lines. We will use cryo-Electron Tomography to observe the SCs formed *in situ* and therefore gain an understanding of the MRC organisation depending on COX7A isoform. These observations will be complemented with a range of other techniques, such as Mass spectrometry and Soft X-Ray Tomography, to further characterise the structural and bioenergetic impact of each isoform.