ISMB Symposium 2024

Programme & Abstracts

Tuesday 25th and Wednesday 26th June 2024
UCL Bloomsbury campus

ISMB
Institute of Structural and Molecular Biology

Programme & Abstracts

Tuesday 25th and Wednesday 26th June

ISMB Director
Professor Franca Fraternali, f.fraternali@ucl.ac.uk
ISMB Deputy Director
Professor Kostas Thalassinos, k.thalassinos@ucl.ac.uk
ISMB Coordinator
Cyndy Thooi, c.thooi@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/structural-and-molecular-biology

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/

School of Natural 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: https://www.bbk.ac.uk/school/natural-sciences

School of Computing and Mathematical Sciences
Birkbeck University of London, Malet Street , London WC1E 7HX
Tel: +44 (0)20 7631 6700; Fax: +44 (0)20 7631 6727
Web: https://www.bbk.ac.uk/school/computing-and-mathematical-sciences

Programme Day 1

Tuesday, 25th June

Medical Sciences and Anatomy Building, UCL

Venue for talks: J Z Young Lecture Theatre

9:00 – 9:30	Welcome Coffee and registration (Gavin de Beer LT)
9:30 – 9:45	Opening by ISMB Director, Franca Fraternali
Day 1 Session 1	Chair: Dr Kish Adoni
9:45 – 10:25	Speaker: Prof Alfonso Valencia
10:30 – 11:00	*Poster Session (Coffee/tea served in Gavin de Beer)*
11:10 – 11:50	Speaker: Prof Vittorio Limongelli
11:55 – 12:35	Speaker: Prof Christopher Tate
12:40 – 13:40	*LUNCH (Haldane Hub)*
Day 1 Session 2	Chair: Alice Pettitt
13:40 – 14:25	Speaker: Prof Syma Khalid
14:30 – 15:10	Speaker: Dr Josie Ferreira
15:15 – 15:45	*Poster Session (Coffee/tea served in Gavin de Beer)*
15:50 – 16:30	Speaker: Prof Joanne Santini
16:35 – 17:10	Speaker: Žiga Avsec
17:30 – 19:00	*Poster and Networking session (Haldane Hub)*
19:00	End of Day 1

Programme Day 2

Wednesday, 26th June

Medical Sciences and Anatomy Building, UCL

Venue for talks: J Z Young Lecture Theatre

9:30 – 9:55	Registration and coffee/tea (Gavin de Beer)
Day 2 Session 1	Chair: Dr Manoj Saxena
10:00 – 10:40	Speaker: Dr Brian Ho
10:45 – 11:25	Speaker: Dr Tracey Barrett
11:25 – 11:55	*Poster Session (Coffee/tea served)*
12:00 – 12:40	Speaker: Dr Christopher Waudby
12:45 – 14:00	*LUNCH (Haldane Hub)*
Day 2 Session 2	Chair: Dr Joseph Ng
14:00– 14:40	Speaker: Prof Paul Guichard
14:45 – 15:25	Speaker: Prof Ivan Gout
15:30 – 16:00	*Coffee/tea break (Gavin de Beer)*
16:05 – 16:45	Speaker: Prof Neil McDonald
16:45 – 17:00	Prizes and Closing by Franca Fraternali
17:30 – 19:30	*ISMB Summer Party (Bentham House, UCL)*
19:30	**END**

Abstracts
Guest speakers
(last name order A-Z)

Accurate proteome-wide missense variant effect prediction with AlphaMissense

Žiga **AVSEC**

Google DeepMind

Email address: avsec@google.com

Bio:
Žiga Avsec is a research scientist at Google DeepMind where he leads the genomics initiative within Google DeepMind's Science program. He is interested in developing AI systems that can help us better understand the code embedded in the genome sequence. Prior to that, he obtained his Ph.D. from the Technical University of Munich, supervised by Julien Gagneur.

Abstract:
The vast majority of missense variants observed in the human genome are of unknown clinical significance. We present AlphaMissense, an adaptation of AlphaFold fine-tuned on human and primate variant population frequency databases to predict missense variant pathogenicity. By combining structural context and evolutionary conservation, our model achieves state-of-the-art results across a wide range of genetic and experimental benchmarks, all without explicitly training on such data. The average pathogenicity score of genes is also predictive for their cell essentiality, capable of identifying short essential genes that existing statistical approaches are underpowered to detect. As a resource to the community, we provide a database of predictions for all possible human single amino acid substitutions and classify 89% of missense variants as either likely benign or likely pathogenic.

The development of broadly neutralising anti-influenza nanobodies in preparation for future pandemics

Dr Tracey **BARRETT**

School of Natural Sciences Birkbeck

Email address: t.barrett@bbk.ac.uk

Abstract:
The production of vaccines and immunotherapeutics to address the antigenic variability of influenza viruses is a major challenge. It has been shown, however, that viral epitopes largely conserved owing to their functional importance, can be successfully targeted by single domain antibodies (nanobodies). Given the nature of their targets, such nanobodies often exhibit cross reactivity against several viral variants and therefore have potential as universal influenza therapeutics. This talk will focus on one candidate nanobody R1aB6 and the approaches used to successfully broaden its cross reactivity as well as the implications based on the cryo-EM structure of an R1aB6-Haemagglutinin complex.

A unique cytoskeleton drives transmission of the malaria parasite

Dr Josie **FERREIRA**

Department of Structural and Molecular Biology, University College London

Email address: josie.ferreira@ucl.ac.uk

Abstract:
To ensure disease transmission, the malaria parasite undergoes multiple rounds of metamorphosis, as it entirely alters its cell morphology to promote uptake and establishment in the mosquito vector and human host. Two cytoskeletal components play essential roles in this process: microtubules and actin.

Within each new ecological niche, microtubules drive the single-celled parasite's successive cellular transformations. Microtubules have been studied extensively and their architecture and composition are established to be highly conserved. Using focussed ion beam milling and electron cryo-tomography, we recently studied distinct stages in the Plasmodium falciparum life cycle. This revealed that the parasite has microtubules which are evolved to undertake specific roles in each life cycle stage with structures that are strikingly different from the well-studied canonical microtubules in vertebrates.

While unique microtubules drive cellular transformations, filamentous actin ensures several parasite stages can migrate between different niches. These stages utilise a unique form of motility, termed gliding motility, which relies on a specialised actomyosin motor system. Our recent work on actively gliding parasites sheds light on this process and highlights novel roles of parasite actin in other cellular locations. Together, this work provides unexpected insights into adaptations of the parasite's cytoskeleton, highlighting areas of novelty where the parasite has diverged from the biology of the host.

Coenzyme A biology, but not as we know it.

Prof Ivan **GOUT**

Department of Structural and Molecular Biology, University College London

Email address: i.gout@ucl.ac.uk

Abstract:
Coenzyme A (CoA) is an essential cofactor in all living cells. CoA and its thioester derivatives (acetyl-CoA, malonyl-CoA, HMG-CoA etc.) participate in diverse metabolic processes, allosteric regulatory interactions and the regulation of gene expression. Dysregulation of CoA biosynthesis in animal models and inborn mutations in human genes involved in the CoA biosynthetic pathway have been associated with neurodegeneration and cardiomyopathy.
We have recently discovered a novel and unappreciated role of CoA in redox regulation. Using cell-based and animal models as well as developed reagents and methodologies, we demonstrated that CoA functions as a major cellular antioxidant under oxidative or metabolic stress. This function is mediated by covalent protein modification by CoA, which we termed CoAlation. To date, protein CoAlation is known as a reversible and widespread posttranslational modification induced by oxidizing agents and metabolic stress in prokaryotic and eukaryotic cells. We developed a robust mass spectrometry-based methodology for the identification of CoA-modified proteins in cells and tissues which allowed us to identify over 2100 CoAlated proteins. Protein CoAlation alters the molecular mass, charge, and activity of modified proteins. It can also protect oxidized cysteine residues from overoxidation and induce significant conformational changes. Based on these findings, we propose that under physiological conditions CoA functions as a key metabolic cofactor but acts as an antioxidant in cellular response to oxidative or metabolic stress. Our recent advances on molecular dissection of the CoAlation/deCoAlation cycle and studying the pattern of protein CoAlation in pathologies associated with oxidative stress will be presented.

Time-series reconstruction of the structural assembly of the centriole using U-ExM

Prof Paul **GUICHARD**

Department of Molecular and Cellular Biology, Université de Genève, Switzerland

Email address: paul.guichard@unige.ch

Abstract:
The centriole is a gigantic macromolecular assembly of cylindrical shape with a diameter of 250nm wide and approximately 450nm long. Due to these dimensions, conventional structural biology methods such as crystallography, or single particle analysis in cryo-EM cannot be used. In recent years, cryo-tomography has proven to be the method of choice to reveal the native architecture of centrioles at a resolution of approximately 20 angstroms, revealing a large number of substructures that support the function of the centriole in the centrosome or flagellum formation. However, understanding how the centriole assembles and when these substructures are recruited during this process is not fully understood. Here, I will present the latest findings from our laboratory where we used ultrastructural expansion microscopy U-ExM to address this crucial question. By correlating the spatial location of 24 centriolar proteins with structural features, we were able to perform a time-series reconstruction of protein distributions throughout human procentriole assembly, unveiling the molecular architecture of the centriole biogenesis steps.

Impact of Interbacterial antagonism on multispecies bacterial community dynamics

Dr Brian **HO**

Department of Structural and Molecular Biology, UCL and School of Natural Sciences, Birkbeck

Email address: b.ho@ucl.ac.uk

Abstract:
Multispecies bacterial communities can be found all over the natural world, playing critical roles in environmental ecology, agriculture, as well as human health and disease. One of the central factors contributing to the composition and dynamics of these communities are the bacterial cell-cell interactions found within them. In particular, the antagonistic interactions have emerged as being especially important – that is, interbacterial warfare is a defining feature of most multipsieces microbial populations. Over the last couple years, my research group has largely focused on one specific form of bacterial antagonism: the type VI secretion system (T6SS). The T6SS is found in a wide range Gram-negative bacteria, including a number of important human pathogens. It consists of a multisubunit molecular nanomachine that enables the delivery of several different toxic effector proteins directly into adjacent cells. In this talk, I will present two stories stemming from our efforts to develop in vivo and in vitro models to study how T6SS shapes multispecies bacterial community development.

Doing Microbiology with Computers- Focus on Gram-negative Bacteria

Prof Syma **KHALID**

Department of Biochemistry, University of Oxford

Email address: syma.khalid@bioch.ox.ac.uk

Abstract:
Gram-negative bacteria are protected by sophisticated, tripartite cell envelopes, which control entry and efflux of molecules, including antibiotics.
Thus, in addition to their fundamental biology, understanding how the molecular choreography within these cell envelopes underpins their functioning is also of biomedical interest. We are using atomistic and coarse-grained molecular dynamics to simulate portions of the E. coli cell envelope with models that incorporate biomolecular complexity and also begin to bridge some of the system size gaps that traditionally exist between experimental and computational studies of biological systems. In summary we are working towards using our simulations as 'computational assays' that will enable us to do Computational Microbiology.

Molecular binding studies in G Protein Coupled Receptors: from ligand binding to receptor dimerization

Prof Vittorio **LIMONGELLI**

Faculty of Biomedical Sciences, Euler Institute, University of Lugano (USI), Switzerland

Email address: vittoriolimongelli@gmail.com

Abstract:
Elucidating the structural and energetic properties of G Protein Coupled Receptors (GPCRs) is of paramount relevance to guide rational drug discovery studies. However, the complexity and the long timescale of receptor activation process make many aspects of GPCRs functioning elusive to both experimental and computational techniques. Here, I demonstrate that such limitations can be overcome by employing advance calculations like funnel-metadynamics [1,2] and coarse-grained molecular dynamics calculations [3,4], which allow disclosing the thermodynamic and kinetic properties of drug binding to GPCR, and investigating important processes such as receptor activation and dimerization. In particular, I show the case of adenosine A2A receptor in which ligand binding occurs by hopping in multiple binding modes and the transition between the vestibular and the orthosteric binding site determines the ligand (un)binding rate. The agonist and antagonist binding allosterically perturbs the receptor state that passes from a pseudo-active state - here described for the first time - to the full active and inactive state, respectively [5]. In addition, the minute timescale formation of receptor dimers can affect the access to the binding sites of the ligand (extracellular) and the effector G protein (intracellular), indicating that dimerization de facto is a fine allosteric regulatory mechanism of GPCR activity [4]. The advance in molecular binding simulation techniques [6] is expected to impact on future drug discovery studies where structural, energetic and dynamic properties of the receptor could be taken into account to achieve ligands with tailored bias signaling properties and great therapeutic potential for a broad spectrum of diseases.

References

1. Limongelli V, Bonomi M, Parrinello M. Proc. Natl. Acad. Sci. USA 110, 6358-6363 (2013)
2. Raniolo S, Limongelli V. Nat. Protoc. 15, 2837-2866 (2020)
3. Souza et al. Nat. Commun. 11: 3714 (2020)
4. Di Marino D, Conflitti P, Motta S, Limongelli V. Nat. Commun. 14, 6439 (2023)
5. D'Amore VM, Conflitti P, Marinelli L, Limongelli V. bioRxiv, 14.557711 (2023)
6. Limongelli V. Wiley Interdiscip. Rev. Comput. Mol. Sci. 10, e1455 (2020)

Kinase-driven control mechanisms for cell polarity

Prof Neil **McDONALD**

The Francis Crick Institute and School and Natural Sciences, Birkbeck

Email address: Neil.Mcdonald@crick.ac.uk

Abstract:
The PAR-titioning-defective or PAR proteins are key components of an essential cell polarity network conserved throughout the animal kingdom. This network underlies polarization in a wide variety of cell types from neurons, epithelial cells, migrating cells to the worm zygote, impacting on membrane identity and asymmetric cell division. The main components of the network include the serine/threonine kinases, PAR-1 and aPKC, the 14-3-3 protein PAR-5, which binds phosphorylated substrates, the PDZ-containing scaffold proteins PAR-3/Bazooka and PAR-6, the small Rho-family GTPase CDC-42, and Lethal Giant Larva (LGL), a multi-site substrate of aPKC. These proteins are functionally organised into two opposing groups within the cell polarity network that mutually antagonise each other. The precise mechanism of mutual antagonism is not known but likely involves (1) precise and tightly regulated aPKC and PAR-1 kinase-substrate-targeting (2) membrane recruitment and release of polarity proteins through phosphorylation, giving membranes a unique identity (3) wiring of network feedback circuits that enable robust and stable polarization, while ensuring that the network remains responsive to changing cellular states and spatial cues.

In my presentation I will discuss recent unpublished results explaining how aPKC selectively targets the crucial polarity substrates LGL and the PAR-1 kinase, revealing the origin of their mutual antagonism in polarised cells. By combining cryo-EM structures, biochemical and cell-based assays with in vivo models of polarity we develop an integrated model for both single-site and multi-site substrate phosphorylation and determine the consequences of substrate phosphorylation.

A story about poisons, phages and neonatal sepsis

Prof Joanne **SANTINI**

Research Department of Structural and Molecular Biology, University College London

Email address: j.santini@ucl.ac.uk

Abstract:
Arsenic contamination of drinking water poses a health risk to over 150 million people across more than 50 countries with those in the Bengal Delta most affected. I will discuss our research on arsenic-metabolising microbes that have played key roles in mobilizing arsenic in the environment as well as providing solutions for its removal and detection. Beyond these microbes are the ones in the gut of humans that detoxify or make arsenic more toxic. Studying microbes in the gastrointestinal tract of humans consuming arsenic-contaminated water has shifted our research from arsenic metabolism to bacteriophages. We started our phage journey by discovering the first ever megaphages (genomes >500 kb) and additional huge (>200 kb) phages in the GI tracts of humans and different animal species. We found that some contain alternative genetic codes, CRISPR-Cas systems and many more interesting biological features. More recently we have expanded our phage work to develop phage-based therapies for treating antimicrobial resistant bacterial infections that cause neonatal sepsis in low- and middle-income countries.

The molecular basis for GPCR pharmacology illuminated through structural biology

Prof Christopher G. **TATE**

MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH

Email address: cgt@mrc-lmb.cam.ac.uk

Abstract:
G protein-coupled receptors are the cornerstone of intercellular communication throughout the human body and impact a multitude of physiological processes in both healthy and disease states. There is therefore great interest in developing new and better drugs to GPCRs, particularly through the development of pathway-specific activators of the receptors (biased agonists). There is also great interest in higher oligomeric states of GPCRs and how these may impact on the pharmacology of the receptors. We have been addressing such questions over the last 15 years through the structure determination (X-ray crystallography and cryo-EM) of specific states of receptors bound to ligands of different efficacy (inverse agonists, partial agonists, full agonists, biased agonists) and to different intracellular modulators (G proteins, G protein mimetics, arrestin).

I will discuss two aspects of allosterism in the talk. Firstly, I will discuss the effect of different intracellular modulators (G protein versus arrestin) on the orthosteric binding site of the β1-adrenoceptor and how this impacts agonist affinity and suggests ways forward in developing biased agonists (1,2). Secondly, I will discuss how dimerisation of the yeast pheromone receptor Ste2 can impact G protein affinity (3,4).

1. Warne, T, Edwards, PC Doré, AS, Leslie, AGW, Tate, CG (2019) Molecular basis for high-affinity agonist binding in GPCRs. Science 364, 775-778
2. Lee, Y., Warne, T., Nehmé, R., Pandey, S., Dwivedi-Agnihotri, H., Edwards, P.C., García-Nafría, J., Leslie, A.G.W., Shukla, A.K. & Tate, C.G. (2020) Molecular determinants of β-arrestin coupling to formoterol-bound β1-adrenoceptor. Nature 583, 862-866
3. Velazhahan, V., Ma, N., Pándy-Szekeres, G., Kooistra, A.J., Lee, Y., Gloriam, D.E., Vaidehi, N. & Tate, C.G. (2021) Dimeric structure of the Class D GPCR Ste2 coupled to two G proteins. Nature 589, 148-153
4. Velazhahan, V., Ma, N., Vaidehi, N. & Tate, C.G. (2022) Activation mechanism of the class D fungal GPCR dimer Ste2. Nature 603, 743-748

The convergence of large scale biomedical data, digital twins and Ai applications

Prof Alfonso **VALENCIA**

Life Sciences, Barcelona Supercomputing Centre, Spain

Email address: alfonso.valencia@bsc.es

Abstract:
In this presentation, I will explore the transformative potential of three parallel development in data access, simulations of biological systems (Digital Twins) and AI/Large Language Models.
Data. For the first time Europe is converging towards a single space for the use of the medical data for research (European Health Data Space). We participate in two large track projects: European Genomic Data Infrastructure - GDI and Cancer Image Europe- EUCAIM, dedicated to the implementation of the technology to make the federated access and analysis of the data possible, with their direct transposition to the implementation in the data infrastructure of the Spanish Genomic Medicine Initiative (IMPaCT-Data).

Digital Twins. Following the European Virtual Human Twins manifesto promoted by EDITH (European Digital Innovation Hubs Network) and followed by the EU declaration for the promotion of development of virtual twins leveraging new methodologies and the available High-Performance Computing capacities, in particular those related with virtual (pre-)clinical trials. we have developed a first generation of simulations at the cellular and intracellular levels (PhysiBoSS software), to facilitate the generation of hypotheses regarding the mechanistic consequences of perturbations within cellular systems.

AI and synthetic data. Based on the development of national and European initiatives (for the development of Foundational Models with emphasis on national languages, in our case Spanish, Catalan, Basque and Galician, as well as the application to biomedical text (i.e., the Carmen I corpus of annotated medical records), and the extensions to other biological objects.
Throughout this presentation, I will highlight recent advancements in these three areas, and the possibilities that their convergence opens for biomedical research.

Accelerating fragment-based drug discovery using NMR spectroscopy

Dr Christopher **WAUDBY**

UCL School of Pharmacy

Email address: c.waudby@ucl.ac.uk

Abstract:
Fragment-based drug discovery (FBDD) is a powerful approach for generating small-molecule lead compounds from weak-binding fragments. However, FBDD faces a unique challenge in detecting and characterising weak interactions (high μM to low mM binding affinities). As a consequence of this, biophysical assay methods are generally preferred, but the low affinities associated with typical fragment hits can still push detection methods to their limits, leading to potential false positives and false negatives. Ligand-observed NMR spectroscopy serves as a crucial component of the screening toolkit. Several experiments have been developed, including saturation transfer difference (STD), waterLOGSY, and nuclear spin relaxation (through CPMG measurements). The latter can also be applied to fluorine-containing ligands using 19F spectroscopy. However, the translation of 'hits' from these experiments into affinity measurements, even only in relative terms, presents a challenge and inhibits the ranking of different compounds.

In this talk, we introduce a new approach that employs 19F ligand-observed NMR to detect binding-induced relaxation dispersion. This method has enabled us to successfully measure bound ligand populations of less than 0.5% and dissociation rates as rapid as 58,000 s-1. We anticipate that this approach will accelerate the early stages of FBDD by facilitating the characterization, screening, and ranking of fragment libraries.

Lastly, we will provide an update on the establishment of an in-house library of fluorinated fragments available for screening within UCL and the ISMB.

Posters (last name alphabetical order):

Aboagye-Mensah, Dorothy	*The effect of divalent salt on the structure of negatively supercoiled DNA*
Adoni, Kish	*A Novel Crosslinker Improves Monolink and Crosslink Identification for Alphafold2 based Integrative Protein Structure Determination Pipeline*
Bangera, Mamata	*Structural and dynamic visualization of the interaction between the Microtubule-Associated Protein 7 (MAP7) and Microtubles*
Bezzera da Cruz, Carlos Henrique	*Design of antimicrobial peptides self-assembled into virus-like capsids*
Charlton, Georgina	*Using Crosslinking Mass Spectrometry to Understand the Structure of Proteins Implicated in Huntington's Disease*
Chaterjee, Samadrita	*Developing a kinase optimized meta-predictor to detect the implication of pathogenic mutations in diseases*
Cowell, Alana	*Exploring isoform dependent Arp2/3 structural flexibility through cryo-EM*
Dulson, Christopher	*Targeting the RNA polymerase of African Swine Fever Virus for Drug Discovery*
Faniyi, Amy	*Islet amyloid polypeptide disruption of the pancreatic beta-cell plasma membrane: A new mechanism of amyloid induced toxicity?*
Filippov, Ivan	*Evolutionary Origins of Amyloidogenicity - Reconstruction and Analysis of Ancestral IAPP Sequences*
Gracia I Carmona, Oriol	*Addressing biases in large language models for variant impact prediction in macro proteins*

Guo, Dongjun	*ImmunoMatch: Illuminating the design of antibody heavy and light chain pairs using deep learning approaches and structure analysis*
Herrera Braga, Julian	*Towards the next generation of variant impact predictors using deep learning and multiomics data*
Kyriakou, Stella	*The role of oxidation in primary nucleation and further prion aggregation*
Liu, Chunan	*Antigen-specific Antibody Design – Preliminary Progress*
Liu, Tianyang	*Structural comparison of Arp2/3 iso-complexes*
Lucic, Anka	*The structure-function relationship of virus lasers*
Martin, Reuben	*Quantum Mechanics for High-Throughput Drug Discovery*
Ng, Joseph	*BHive+: An integrated computational toolkit for B-cell systems immunology*
Pettitt, Alice	*An integrative characterisation of proline cis and trans conformers in a disordered peptide*
Lera Ramirez, Manuel	*ShareYourCloning: towards a FAIR standard for sequence provenance*
Silvey, Kirsten	*Characterisation of monomeric isoforms of amyloid-β using solvation-based Nuclear Magnetic Resonance techniques*
Stiens, Jennifer	*Making 'antisense' of the PhoPR two-component system in the Mycobacterium tuberculosis complex*
Trewby, Will	*Nanoscale mapping of biosynthesis and molecular insertion in the Escherichia coli outer membrane*

Troman, Luca	*Using purified Xenopus tubulin as a model to decipher the tubulin code*
Waman, Vaishali	*The Encyclopaedia of protein Domains (TED) from AlphaFold2-predicted structures: expansion of CATH domain space and insights into structural diversity*
Virgo, Mollie	*Different modes of interbacterial competition elicit distinct host responses and health outcomes*
Yang, Tangweina	*On the importance of being amidated: Analysis of the role of the conserved C-terminal amide of amylin in amyloid formation and cytotoxicity*
Zouhair, Ines	*Comparative in-situ characterisation of respiratory supercomplexes in human mitochondria*

Poster Abstracts
(Last name order A-Z)

The effect of divalent salt on the structure of negatively supercoiled DNA

Dorothy **ABOAGYE-MENSAH**, Aleksander Klimczyk, and Graeme A. King

Department of Structural and Molecular Biology, Division of Biosciences, Gower Street, London, WC1E 6BT, United Kingdom

Email address: d.aboagye-mensah@ucl.ac.uk

Abstract:
Negative supercoiling (underwinding) of DNA frequently occurs in vivo and plays a vital role in regulating gene expression by modulating genome accessibility. It is known that negatively supercoiled DNA can adopt a range of topologies; however, the influence of environmental factors, such as ionic strength, on the structure of underwound DNA is still poorly understood. Here, we develop a novel analytical strategy to quantify the mechanical properties of underwound DNA, based on the measured force-distance curves. We then apply this approach to explore the effect of divalent cations on the structure of underwound DNA. We demonstrate that moderate concentrations of divalent cations can substantially increase the bending stiffness of underwound DNA, consistent with the formation of a left-handed structure known as Z-DNA. In the presence of $\geqslant$ 5 mM divalent salt, we furthermore observe multiple large ruptures in the force-distance curves of underwound DNA. This suggests that divalent cations can also substantially increase the stability of writhed structures (plectonemes). Together, our findings reveal that the structural composition and stability of underwound DNA is highly sensitive to the concentration of divalent cations, which has implications for the biological interactions of supercoiled DNA.

A Novel Crosslinker Improves Monolink and Crosslink Identification for Alphafold2 based Integrative Protein Structure Determination Pipeline

Kish **ADONI**

Institute of Structural and Molecular Biology
Department of Structural and Molecular Biology, University College London

Email address: k.adoni@ucl.ac.uk

Abstract:
Introduction
AlphaFold2 has revolutionised the role of computational techniques for protein structure determination. However, such algorithms struggle to predict protein structures that contain intrinsically disordered regions or occupy multiple conformations. We have developed a pipeline that combines AlphaFold2 and experimental crosslink/monolink data to target these dynamic structures. Added to this, we demonstrate the application of this pipeline in deconvoluting the structural rearrangement of Insulin like growth factor 2 mRNA binding protein 1 (IMP1) upon its coordination with RNA. Further, we introduce the novel crosslinker, TOLE165, for improved recovery of structurally relevant crosslinks and monolinks on the human 26S proteasome complex, and Glutamine Binding Protein (QBP). Thus, TOLE165 represents a novel addition to the XL-MS toolbox for integrative protein structure determination algorithms.

Methods
Apo-QBP and Ca2+ bound QBP (1:1000 QBP:CaCl2), as a known cofactor of QBP that drives conformational change, were incubated with BS3-d0/d4 crosslinker to quantitatively probe changes in crosslinks and monolinks upon QBP-Ca2+ coordination. Data generated from an Orbitrap Eclipse Tribrid MS was then incorporated into our AlphaFold2-XLMS pipeline, with benchmarking of predicted open and closed QBP conformations vs standalone AlphaFold2. The same approach was used to reveal IMP1 structural dynamics upon RNA binding.
Following this, our novel TOLE165 crosslinker, developed by our collaborators (Maulide lab) and demonstrated to be more stable against retro-Michael addition directed degradation (using NMR reaction monitoring), was benchmarked, using the above workflow, against DSSO for the structural investigation of the h20S Proteasome complex, and QBP monomer.

Preliminary Data
Our AlphaFold2-XLMS pipeline generated models for C3 protein, open and closed conformation (simulated crosslinks and monolinks), luciferase, open and closed conformation (simulated

crosslinks and monolinks) and QBP, open-conformation (experimental crosslinks and monolinks) that outperformed the top pTM scoring AlphaFold2 prediction in terms of RMSD from the native crystal structure. Interestingly, it was specifically the monolinks that boosted the predictive capacity of our pipeline for the QBP open conformation, highlighting their importance for such workflows. These data place our pipeline as a useful tool for the prediction of such proteins that occupy differential conformations upon ligand binding. To demonstrate it's potential for protein structure elucidation, experimental XLMS data on apo-IMP1 and RNA-bound IMP1 were incorporated into our pipeline to reveal novel IMP1 domain rearrangements upon coordinating RNA. IMP1 has been implicated in decreased metastatic potential of mammary tumours and as such, these findings could shed light on drug-protein interactions of IMP1. Despite the success of our methodology, we recognised its reliance on the recovery of experimental crosslinks and monolinks. Thus, we identified the novel crosslinker: TOLE165, as a potentially useful crosslinker for boosted structurally informative crosslinks and monolinks, on account of its increased stability against degradation when compared with DSSO (as validated by NMR reaction monitoring). Whilst DSSO generated 80 % more crosslinks than TOLE165, 50 % more monolinks were recovered from TOLE165 against the human 20S proteasome complex. To supplement these findings, TOLE165 identified double as many intramolecular crosslinks and more than double as many intramolecular monolinks, compared to DSSO, when applied to monomeric QBP open and closed conformations, suggesting its potential in revealing useful structural information for our integrative AlphaFold2-crosslinking proteins structure determination algorithm.

Novel Aspect

The use of experimental monolinks/crosslinks for AlphaFold2 based predictions. TOLE165 novel crosslinker with improved stability vs DSSO for this pipeline.

The effect of divalent salt on the structure of negatively supercoiled DNA

Agnes Adler[1], Mamata **BANGERA**[2], J. Wouter Beugelink[1], Salima Bahri[1], Hugo van Ingen[1], Marc Baldus[1], Carolyn A. Moores[2]

[1]Utrecht University, Utrecht, Netherlands
[2]Institute of Structural and Molecular Biology, School of Natural Sciences, Birkbeck College, London, United Kingdom

Email address: m.bangera@bbk.ac.uk

Abstract:
Microtubules are key components of the cytoskeleton, and along with motor and non-motor microtubule-associated proteins (MAPs) form the architectural foundation of eukaryotic cells. Microtubule associated protein 7 (MAP7) is a unique microtubule binding protein that is involved in recruitment and activation of the plus-end directed motor kinesin-1. MAP7 is also implicated in neuronal axon branching and nuclear positioning in myotubes. Full length MAP7 is mainly structurally disordered, with the exception of its helical microtubule binding (MTBD) and kinesin binding domains (KBD). Here we probe the interaction of MAP7 MTBD with the microtubule lattice using a combination of electron microscopy and nuclear magnetic resonance (NMR) spectroscopy. Previous NMR studies and AlphaFold2 predictions showed that the 110-residue MAP7 MTBD forms an elongated 130Å helix and our cryo-EM 3D reconstruction identified its binding site along the microtubule protofilament, independently of the mode of microtubule stabilization. However, the register of the helix along the length of the microtubules is obscured in cryo-EM studies due to helical averaging of the microtubule over a tubulin dimer (~80Å). Hence, NMR was also used to identify rigid and flexible regions along the entire length of the microtubule bound MAP7 MTBD helix. In combination with AlphaFold2 prediction of the MTBD-tubulin complex, a full model of its microtubule interaction was calculated. Further we demonstrate interactions between the tubulin C-terminal tails and MAP7 MTBD that might play an important role in the formation of the MAP7 MTBD-microtubule complex. Overall, our study provides a comprehensive picture of both the well-defined and flexible interactions that contribute to binding of MAP7 MTBD to microtubules. In doing so, we outline a general framework for investigating dynamic, extended interactions between microtubules and their binding partners.

Design of antimicrobial peptides self-assembled into virus-like capsids

Carlos Henrique **BEZERRA DA CRUZ,** Franca Fraternali

Institute of Structural and Molecular Biology, University College London

Email address: carlos.cruz@ucl.ac.uk

Abstract:
Over the recent past decades, the constant threat of antimicrobial resistance has encouraged the development of new drugs and strategies to combat it. However, the discovery of novel approaches for targeting many bacterial strains remains challenging. In this context, the capability of antimicrobial peptides (AMPs) to disrupt bacterial activity has been extensively explored as a potential weapon against drug-resistant bacteria. Our work proposes the design of antimicrobial peptides self-assembled into virus-like capsids for delivering drugs and antimicrobial products by using state of the art rational protein design combined with multi-scale molecular dynamics simulations. The virus's ability to infect cells has been used as inspiration to design sophisticate delivery systems. Matssura et al.[1,2] have explored the potential of short beta-annulus peptides from the genus Sobemoviruses to self-assemble into regular virus-like capsid structures, and to use these to encapsulate anionic dyes, DNA and bifunctional proteins. Despite the mechanism of assembling and capsid structure details remains unknown, there is robust evidence that three copies of beta-annulus peptides interact with each other to form a trigonal beta-annulus structure that is at the basis of the generated capsule nano-capsule. Recently, the versatility of such peptides to form virus-like nano-capsules for antimicrobial delivering systems was explored by us in two publications.[3,4] By the use of in-silico modelling and coarse-grained molecular dynamics simulations, we reveal the ability of a modified beta-annulus peptide from the Tomato Bushy Stunt virus carrying an antimicrobial sequence to form a stable nano-capsule and to perturb a model bacterial membrane. In our current work, sophisticate learning-based protein design methods were applied to design novel trigonal beta-annulus structures and sequences, using the topology of the beta-annulus from the Sebasnia virus as a seed of our design strategy. With high prediction accuracy, these designed peptides can form the framework needed to nano-capsule assembling. Our hypotheses have been explored by extensive full-atom and coarse-grained molecular dynamics simulations. We highlight challenges involving current coarse grained force fields in exploring these complex self-assembled nanoparticles.

Reference [1]K. Matsuura, K. Watanabe, T. Matsuzaki, K. Sakurai, N. Kimizuka, Angew. Chem. Int. Ed., 49, 9662-9665 (2010) [2]K. Matsuura, Y. Mizuguchi, N. Kimizuka, Biopolymers., 106, 470-475 (2016) [3]I. Marzuoli, C. Cruz, C. Lorenz, F. Fraternali, Nanoscale, 13, 10342-10355 (2021) [4]C. Cruz, I. Marzuoli, F. Fraternali, Faraday Discuss., 232, 448-462 (2021)

Using Crosslinking Mass Spectrometry to Understand the Structure of Proteins Implicated in Huntington's Disease

Georgina **CHARLTON**, Tristan Cragnolini, Rob Goold, Sarah Tabrizi, Konstantinos Thalassinos

[§]Institute of Structural and Molecular Biology, University College London, Gower Street, London, United Kingdom

Email address: georgina.charlton@ucl.ac.uk

Abstract:
Introduction:
Huntington's disease (HD) is an inherited neurodegenerative illness which causes the breakdown of nerve cells. HD is caused by CAG repeat expansion in the huntingtin (HTT) gene. Expression of the protein FAN1 is closely linked to a delay in age-of-onset and slower disease progression. FAN1 is known to interact with the mismatch repair protein MLH1. MLH1 is essential for somatic expansion of the HTT CAG repeat, a process strongly linked to disease progression. FAN1 binding prevents MLH1 from interacting with the DNA repair protein complex MSH3-MSH2, leading to the stabilisation of the CAG repeat. Large regions of FAN1 and MLH1 are unstructured and cannot be probed using established structural biology methods, making mass spectrometry an ideal method for their study.

Methods:
To better understand the FAN1 interactome we performed pull down experiments, where cells were formaldehyde crosslinked to fix interactions before cell lysis. Then a FAN1 pull down was performed to enrich for FAN1 and its interactors. The formaldehyde crosslinks were broken before trypsin digestion. To get more information on the structure of FAN1 and its interactors, particularly the unstructured regions we performed on bead crosslinking with BS3 before an on-bead digest was performed. In both cases samples were run on the Orbitrap Eclipse (Thermo Scientific) and data processed by MaxQuant and pLink.

Preliminary data:
Previous studies have shown that FAN1 binds with MLH1 via the SPYF motif and MIM box in the N-terminal region of FAN1. Here, mutating the SPYF motif and MIM box, showed a decrease in the binding of FAN1 to MLH1 compared to the wildtype. Using on-bead XLMS, we showed sites of interest on FAN1 and MLH1 which could directly interact. AlphaFold predictions were used to better understand the structure of the FAN1/MLH1 complex and zooming in on the binding site

gives a better idea of the interaction taking place between the SPYF on FAN1 and E669 and Y625 on MLH1.

We were able to show that MLH1 directly interacts with FAN1, and sites important for this interaction. Further studies in this research will study the binding of MLH1 and MSH3, which promotes CAG repeat expansion leading to the progression of Huntington's disease. A better understanding of how MLH1 binds to FAN1 and to MSH3 will give insights into how Huntington's disease progresses and allow the creation of an MLH1 mutant which binds to FAN1 but not MSH3 in the hopes of slowing down Huntington's disease progression.

The structural role of CaMKII in long-term potentiation (Poster)

Samadrita **CHATTERJEE**[1], Paul Ashford[1] , Vaishali P. Waman[1], Andrew C.R. Martin[1], Christine Orengo[1]

[1]Institute of Structural and Molecular Biology, University College London, London, UK

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

Abstract:
Mutations in kinases play a significant role in the development and prognosis of cancer. A study has shown that family-based predictors outperform generic predictors. Our study is aimed at developing a kinase optimized meta-predictor to determine the pathogenicity of mutations in kinases of human proteome. VariPred, an in-house predictor 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. We developed a meta-predictor by combining other state-of-the-art predictors namely Alphamissense and Polyphen-2 with VariPred and testing on a dataset from Humsavar. Logistic Regression, Multi Layer Perceptron, Random Forest, CatBoost, XGBoost are the 5 classifiers that have been used to build the model and are tested on the dataset. Varipred was retrained with kinase variants and the meta-predictor was built. It is found that retrained Varipred (MCC = 0.596) performs better than original Varipred (0.530) on the test dataset. The meta-predictor was built by combining Alphamissense, Polyphen-2 and the retrained Varipred , trained and tested using 10 fold cross-validation and after hyperparameter tuning of the Multi Layer Perceptron classifier, it gives an MCC of 0.611. We plan to combine retrained VariPred with other state-of-the-art predictors namely, 3CNet and EVE for building a kinase optimized meta-predictor and evaluate peerformance. The best performing meta- predictor shall be applied to detect the mutations involved in renal cancer by using TRACERx data from Dr. Charles Swanton's group.

Exploring isoform dependent Arp2/3 structural flexibility through cryo-EM

Alana **COWELL**, Ramteen Shayan, Miroslav Mladenov, Michael Way, Carolyn Moores

Institute of Structural and Molecular Biology
School of Natural Sciences, Birkbeck University of London

Email address: a.cowell@bbk.ac.uk

Abstract:
Actin branch nucleator Arp2/3 is a seven-subunit complex which coordinates multiple cellular processes. In mammals, three of the subunits of Arp2/3 have two different isoforms which can become part of the full complex, creating 8 possible variants. Each of these variants have been shown to have different actin nucleating properties in vitro and in cellulo (Abella et al., 2016, Galloni et al., 2021). Previous data revealed that complexes containing subunit isoforms ARPC1B/C5L are the most efficient at polymerising actin filaments, whereas ARPC1A/C5 containing complexes are least efficient (Abella et al., 2016). Here we explore the differences in structure and intrinsic flexibility between these Arp2/3 variants to understand how isoform composition impacts the conformational changes occurring during branch formation. Using single particle cryo-EM we were able to obtain 3D reconstructions of the Arp2/3 complex with C1B/C5L and C1A/C5 subunits at 3.1Å and 2.7Å respectively. Analysis of sample heterogeneity and 3D variability reveals key differences in the different Arp2/3 complexes that modify overall flexibility of the complex, which is hypothesised to be important in driving the differences observed in actin nucleation efficiency.

Targeting the RNA polymerase of African Swine Fever Virus for Drug Discovery

Christopher **DULSON**, Gwenny Cackett, Simona Pilotto, Michal Sykora, Finn Werner

Institute of Structural and Molecular Biology, University College London

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

Abstract:
African Swine Fever Virus (ASFV) causes a lethal haemorrhagic fever in domestic pigs and has been reported in 80 countries since 2005. There are no approved antiviral drugs or effective vaccines, therefore infections continue to threaten global food security. ASFV produces its own 8-subunit viral RNA polymerase (RNAP) that is required for viral transcription within the cytoplasm of the infected cell. The essential role provided by the RNAP makes it a key target for drug discovery. To facilitate the discovery of specific inhibitors from within a low biosafety containment laboratory, the RNAP was produced recombinantly via a baculovirus expression system. The purified recombinant RNAP was transcriptionally active in vitro and the optimal performance over ranges of temperature, pH, ionic strength, divalent cation concentration and with different DNA templates, were determined by quantification of α-32P-UTP incorporation into the RNA transcription product. The optimised conditions were then used to develop a high-throughput assay for the detection of the RNA product by interaction with RiboGreen dye which induces fluorescence. A range of compounds that inhibit polymerases and ATPases were screened for inhibition of ASFV RNA polymerase, which led to the discovery of novel inhibitors for the ASFV RNA polymerase.

Islet amyloid polypeptide disruption of the pancreatic beta-cell plasma membrane: A new mechanism of amyloid induced toxicity?

Amy **FANIYI**, Daniel Raleigh, Alexander Zhyvoloup

Institute of Structural and Molecular Biology, University College London

Email address: amy.faniyi.19@ucl.ac.uk

Abstract:

Human Amylin plays a central role in regulating glucose homeostasis, but within type-2 diabetes Amylin is known to aggregate, forming islet amyloid polypeptide (hIAPP). The process of islet amyloid formation is toxic to the insulin producing pancreatic beta-cells, leading to a reduced beta-cell mass and insulin resistance within those with type-2 diabetes. Despite this, the mechanisms of toxicity are not fully understood. Membrane disruption has been proposed as a possible reason for the polypeptide's toxicity. However, this hypothesis is historically based upon work with model vesicles, and it is not clear how those studies translate to the cellular context, especially as the model vesicles employed have very non-physiological lipid composition. They have a high content of anionic lipids, are symmetric and usually lack cholesterol. It is well known that hIAPP interacts effectively with anionic vesicles and induces their leakage, but the outer leaflet of the plasma cell membrane contains a very low percentage of anionic lipids, as they are localized in the inner leaflet. Using INS-1 832/13 cells, a standard pancreatic beta-cell line, we have examined the effect of human IAPP and a non-amyloidogenic variant of IAPP (rodent IAPP, rIAPP) on the beta cell plasma membrane. hIAPP induces the translocation of phosphatidylserine (PS), an essential anionic phospholipid, to the outer leaflet of the beta cell plasma membrane while rIAPP does not. This is of a particular interest because hIAPP is known to interact significantly more efficiently with anionic lipid membranes. These results lead to the hypothesis that hIAPP may exert toxic effects by perturbing membrane asymmetry, leading to the appearance of PS in the outer leaflet which in turn promotes further hIAPP membrane interactions.

Evolutionary Origins of Amyloidogenicity - Reconstruction and Analysis of Ancestral IAPP Sequences

Ivan **FILIPPOV**[1], Min Hu[1], Tangweina Yang[1], Aria Baghai[2], Alexander Zhyvoloup[1], Daniel P Raleigh[1,2,3]

1 Institute of Structural and Molecular Biology, Division of Biosciences, University College London, Gower Street, London WC1E 6BT, United Kingdom
2 The Francis Crick Institute, 1 Midland Road, London NW1 1AT, United Kingdom
3 Laufer Center for Quantitative Biology, Stony Brook University, Nicolls Road, Stony Brook, New York 11790, United States of America

Email: ivan.filippov.18@ucl.ac.uk

Abstract:
Human Islet Amyloid Polypeptide (hIAPP, also known as hAmylin) plays an important role in regulating blood glucose and controlling satiety. However, the polypeptide aggregates to form pancreatic islet amyloid in a process that is toxic to the insulin producing beta-cells and contributes to type-2 diabetes. hAmylin is among the most amyloidogenic peptides known, exhibiting higher aggregation rates than any other amylin from extant species studied to date. It is unclear why this seemingly undesirable trait that is known to contribute to type-2 diabetes has persisted and no explanation has been proposed for the high aggregation rate of the human sequence. Other members of the calcitonin peptide super-family, to which amylin belongs, show no signs of amyloid formation in vivo in any species, suggesting that this trait was acquired by amylin after the last gene duplication event. To study the evolutionary history of this phenomenon we employ ancestral sequence reconstruction, which provides sequence data for each intermediate node on a phylogenetic tree, allowing us to bridge the gaps between extant species. We characterise the ancestral nodes on the route from the human branch tip to the all-vertebrate ancestor for aggregation rate and cytotoxicity. The data reveals that early amylin ancestors are up to 30+-fold less amyloidogenic, as judged by their time to form amyloid than the human sequence, and this indicates a trend of increasing amyloidogenicity over evolutionary time. These results emphasise that hAmylin is not optimised to minimise amyloid formation. They suggest either a possible positive biological function of amyloid formation that could drive positive selection for this trait or lack of an evolutionary pressure to minimise amyloidogencity.

Addressing biases in large language models for variant impact prediction in macro proteins

Oriol **GRACIA I CARMONA**, Timir Weston, Mathias Gautel, Aleksej Zelezniak and Franca Fraternali

Institute of Structural and Molecular Biology
Research Department of Structural and Molecular Biology, University College London

Email: o.carmona@ucl.ac.uk

Abstract:
Large multi-domain proteins pose significant challenges for both experimental studies and computational modelling due to their extensive size, often spanning thousands of amino acids. While Large Language Models (LLMs) offer promise in studying variant effects, their token limit falls short of accommodating these macro proteins. One solution consists in dividing proteins into domain groups. However, since LLMs were trained on complete proteins, biases may arise when analysing protein fragments.
Our study delves deeper into these biases by assessing the performance of ESM2 across a diverse set of proteins containing prevalent domains found in macro proteins. Predictions performed on domain sequences displayed significantly lower confidence than those obtained using whole sequences as context, particularly near the N-terminal end positions.
To mitigate this bias, we fine-tuned ESM2 using a dataset of domain sequences. The resulting model exhibited no significant differences between full-length and domain sequences, yielding a narrower distribution of discrepancies compared to the original ESM2. Moreover, the final predictions from the fine-tuned model aligned closely with those of ESM2 using the whole sequence as context, indicating successful bias reduction without compromising prediction quality. Furthermore, we evaluated the fine-tuned model's efficacy in testing variant effects on Titin, a macro protein with over 35,000 amino acids. Our model outperformed both the original ESM2 and other state-of-the-art variant effect prediction methods across all tested metrics.
In summary, our findings highlight the effectiveness of fine-tuning LLMs on domain sequences to alleviate biases and improve accuracy in studying variant effects on large multi-domain proteins like Titin.

ImmunoMatch: Illuminating the design of antibody heavy and light chain pairs using deep learning approaches and structure analysis

Dongjun **GUO**, Joseph CF Ng, Deborah K Dunn-Walters, Franca Fraternali

Institute of Structural and Molecular Biology
Research Department of Structural and Molecular Biology, University College London

Email address: d.guo@ucl.ac.uk

Abstract:
Antibodies are composed of heavy (H) and light (L) chains. Sequence variations of H and L chains therefore combinatorially contribute to a diverse antibody "repertoire" for eliciting responses against a variety of antigens. How H chain chooses its L chain partner is still under debate. Little attention has been paid to the exact amino acid preferences and their relative importance in the H-L protein interface. Our results illustrate molecular rules governing antibody H-L chain pairing preferences.

Here we present ImmunoMatch, a heavy-light chain pairing prediction tool taking advantage of recently published antibody language models. We capitalise on the increase in single-cell, paired H-L antibody repertoire data, and build the model to distinguish cognate H-L pairs from random synthetic pairs, with the AUC achieved 0.75. We assembled an antibody structure database (VCAb: https://fraternalilab.cs.ucl.ac.uk/VCAb/) for external validation and further structural interpretation of the pairing prediction. We show that our model, trained on human antibody repertoire, performs well on human and humanized antibody, while the performance dropped in detecting the cognitive pairs from mouse and chimera antibody. We took one therapeutic antibody (trastuzumab) for further analysis by searching through the potential mutation space using ImmunoMatch and extracting attention matrix and found positions which can increase/decrease the H-L pairing likelihood clustering around CDR loops and H-L interface. These results highlight the necessity of considering the entire antibody sequence in antibody design by pre-excluding unlikely H-L combinations in the pipeline for better developability.

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

Julian Diego **HERRERA BRAGA**

Institute of Structural and Molecular Biology, University College London

Email address: julian.braga.22@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. Recently, I have improved the integration pipeline and added more data sources, such as transcriptomics data from GTex, protein models from AlphaFold and protein residue contact networks. In parallel, I am exploring machine learning models applied to the multiomics database with the goal of classifying damaging variations.

The role of oxidation in primary nucleation and further prion aggregation

Stella **KYRIAKOU,** Laszlo L. P. Hosszu, Jan Bieschke

MRC Prion Unit, University College London

Email address: s.kyriakou@prion.ucl.ac.uk

Abstract:
Introduction: The cellular prion protein (PrPc) is a cell surface glycoprotein anchored by glycosyl-phosphatidyl inositol (GPI) (1). Misfolding of the PrPc leads to the formation of the disease isoform PrPSc, which is the cause of prion diseases (2). Prions propagate by templating of PrPSc (3). However, it is not known what causes the initial misfolding of PrPc and primary nucleation in sporadic prion disease. Oxidation of PrPC can induce PrP misfolding and aggregation (4,5,6,7). Hydrogen peroxide induced oxidation, promoted liquid to solid transition of PrP: Cu2+ condensates and PrP aggregation (8).

Objectives: The aim is to show that oxidation plays a key role in primary nucleation and further prion aggregation. Additionally to demonstrate the exact molecular mechanism in which prions are formed and aggregate via an oxidative pathway.

Methods: CD spectroscopy has been used to examine the extent of PrPC unfolding and EPR to detect whether free radical generation occurred under oxidation. Mass Spectroscopy has be used to establish whether oxidation took place and in-vitro ThT Aggregation assays to detect PrP amyloid formation using oxidised PrP as a seed. Chemical and thermal denaturation using CD spectroscopy aimed to compare the stability of PrPC and oxidised PrP (PrPox).

Results: CD spectroscopy of oxidised PrP revealed loss of α-helix and partial unfolding while EPR identified free radicals under oxidative conditions. Chemical and thermal denaturation demonstrated that PrPox has lower stability than PrPC therefore PrPox is more prone to aggregation than PrPC. Lastly, ThT aggregations assays at different pH revealed that pH affects aggregation.

Conclusion: Oxidation of PrPC has shown distinct changes in structure, stability and has induced aggregation leading to PrP amyloid formation. Future experiments such FCCS, RT-PMSA, ThT Aggregation Assay with spin trapping and identification of cross-link post oxidation via western blotting, will shed light to the insides of the molecular mechanism of prion aggregation.

Reference List:
(1) Linden R. The Biological Function of the Prion Protein: A Cell Surface Scaffold of Signaling Modules. Front Mol Neurosci. 2017;10:77
(2) Terry C, Wadsworth JDF. Recent Advances in Understanding Mammalian Prion Structure: A Mini Review. Front Mol Neurosci. 2019;12:169.
(3) Prusiner SB. Molecular Biology of Prion Diseases. Science. 1991;252(5012):1515-22
(4) Younan ND, Nadal RC, Davies P, Brown DR, Viles JH. Methionine oxidation perturbs the structural core of the prion protein and suggests a generic misfolding pathway. J Biol Chem. 2012;287(34):28263-75.
(5) Requena JR, Groth D, Legname G, Stadtman ER, Prusiner SB, Levine RL. Copper-catalyzed oxidation of the recombinant SHa(29-231) prion protein. Proc Natl Acad Sci U S A. 2001;98(13):7170-5.
(6) Gerum C, Silvers R, Wirmer-Bartoschek J, Schwalbe H. Unfolded-State Structure and Dynamics Influence the Fibril Formation of Human Prion Protein. Angewandte Chemie International Edition. 2009;48(50):9452-6.
(7) Redecke L, Bergen Mv, Clos J, Konarev PV, Svergun DI, Fittschen UEA, et al. Structural characterization of β-sheeted oligomers formed on the pathway of oxidative prion protein aggregation in vitro. Journal of Structural Biology. 2007;157(2):308-20
(8) do Amaral MJ, Mohapatra S, Passos AR, Lopes da Silva TS, Carvalho RS, da Silva Almeida M, et al. Copper drives prion protein phase separation and modulates aggregation. Science Advances. 2023;9(44):eadi7347.

ShareYourCloning: towards a FAIR standard for sequence provenance

Manuel **LERA-RAMIREZ**[1], Bahler, J[1]

[1]Univerisity College London, United Kingdom

E-mail address: manuel.lera-ramirez@ucl.ac.uk

Abstract:
Recombinant DNA technology is used in diverse research fields to generate new DNA sequences by combining fragments from existing DNA molecules. These engineered molecules are introduced into cells as self-replicating plasmids and can be used to modify their genomes. Currently, no stablished standard exists to describe how DNA sequences are generated from its ancestors. Therefore, researchers document the provenance of plasmids, cell lines and strains using spreadsheets or text-based systems, which are necessarily inconsistent and differ between collections.
While some proprietary tools keep track of the provenance of sequences, they don't allow users to export this information in an open format. ShareYourCloning aims to:

1. Develop a data model to describe sequence provenance
2. Develop web tools that leverage the model and allow researchers to plan and document their experiments.

In this way, the data is captured through an intuitive interface in a FAIR (Findable, Accessible, Interoperable, Reusable) format at the beginning of the data lifecycle.
A prototype of this tool can be accessed at http://www.shareyourcloning.org/

Antigen-specific Antibody Design – Preliminary Progress

ChuNan **LIU**[1], Lilian Denzler[1], Oliver Hood[1], Andrew Martin[1]

[1]Institute of Structural Molecular Biology, University College London

Email address: chunan.liu@ucl.ac.uk

Abstract:
Antibodies have increasingly been developed as drugs with over 100 now licensed in the US or EU. During development, it is often necessary to increase or reduce the affinity of an antibody and rational attempts to do so rely on having a structure of the antibody-antigen complex often obtained by modeling. The antigen-binding site consists primarily of six loops known as complementarity-determining regions (CDRs), and an open question has been whether these loops change their conformation when they bind to an antigen. Existing surveys of antibody-antigen complex structures have only examined CDR conformational change in case studies or small-scale surveys. With an increasing number of antibodies where both free and complexed structures have been deposited in the Protein Data Bank, a large-scale survey of CDR conformational change during binding is now possible. To this end, we built a dataset, AbAgDb, that currently includes 177 antibodies with high-quality CDRs, each of which has at least one bound and one unbound structure. We analyzed the conformational change of the Cα backbone of each CDR upon binding and found that, in most cases, the CDRs (other than CDR-H3) show minimal movement, while 70.6% and 87% of CDR-H3s showed global Cα RMSD ≤ 1.0Å and ≤ 2.0Å, respectively. We also compared bound CDR conformations with the conformational space of unbound CDRs and found most of the bound conformations are included in the unbound conformational space. In future, our results will contribute to developing insights into antibodies and new methods for modeling and docking.

Reference:
This works has been published
Liu, C., Denzler, L. M., Hood, O. E. C., & Martin, A. C. R. (2024). Do antibody CDR loops change conformation upon binding? mAbs, 16(1). https://doi.org/10.1080/19420862.2024.2322533

Structural comparison of Arp2/3 iso-complexes

Tianyang **LIU**[1], Luyan Cao[2], Miroslav Mladenov[2], Michael Way[2,3], Carolyn A. Moores[1]

1 Institute of Structural and Molecular Biology, Birkbeck College, London WC1E 7HX, UK.
2 The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK.
3Department of Infectious Disease, Imperial College, London W2 1PG, UK.

Email address: t.liu@bbk.ac.uk

Abstract:
The Arp2/3 complex nucleates branched actin networks, which is important for many cellular processes. The Arp2/3 complex is a sevensubunit complex containing Arp2, Arp3, and ArpC1-ArpC5. In mammals, Arp3, ArpC1, and ArpC5 have two isoforms, and these isoform pairs share 91%, 67%, and 67% sequence identity, respectively. Thus, eight iso-complexes can exist in mammals. Cellular data and in vitro biochemical data have shown that the iso-complexes containing ArpC1B and ArpC5L are significantly more efficient in promoting actin assembly than those containing ArpC1A and ArpC5. To understand what leads to the difference in their branch nucleating ability, we determined the structure of the ArpC1BC5LArp3 nucleated branch junction and the ArpC1AC5Arp3 nucleated branch junction using cryo-electron microscopy.
We found that in both structures the seven subunits are stably present, in contrast to previous cellular cryo-ET data, which shows that the ArpC1 density is weaker in ArpC5-containing branches than in ArpC5Lcontaining branches. We also found that although the overall structure is not altered, there are local changes between them. For example, the loop after the ArpC1 protrusion helix in the ArpC1AC5 branch junction is stable, while in the ArpC1BC5 branch junction, this loop is highly flexible and can't be modelled. The N-terminus of ArpC5 shows relatively large movement between these two structures. Loops at the ArpC1 and ArpC5 interfaces adopt different conformations. We inferred that these structural changes may contribute to the Arp2/3 isoform-specific actin polymerization activity. We also found that crossed branches can form along one mother filament with two adjacent Arp2/3 complexes present in close proximity. ArpC3 and ArpC5 from the nearby two branches respectively may form electrostatic interactions. It will be important in the future to explore the functional significance of these cross branches.

The structure-function relationship of virus lasers

Anka **LUCIC**, Finn Porter, Jaime Teneb Lobos and John Hales.

Department of Biochemical Engineering, University College London

Email address: a.lucic@ucl.ac.uk

Abstract:
Understanding the structure function relationship of any biological probe is key to future probe development and optimisation. Virus lasers are an emerging analytical technology based on dye-labelled viruses which act as the active component of a biological laser system. Virus lasers are highly sensitive, and the greatest responsivity of any assay developed to the best of our knowledge. This project aims to understand how the structure of a virus laser can affect the biological structure laser function link by testing multiple conjugation sites and using different biological probe scaffolds. This knowledge will be used to optimise key characteristics of the virus laser, including the minimum concentration needed to generate a laser signal, thus improving sensitivity.

Reference:
Hales, J., Matmon, G., Dalby, P. A., Ward, J. and Aeppli, G., Virus lasers for biological detection, Nat. Comms, 2019, Vol 10, article no. 3594.

Quantum Mechanics for High-Throughput Drug Discovery

Reuben **MARTIN**

Department of Structural and Molecular Biology, University College London

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

Abstract:
Structure-based drug design (SBDD) is rapidly evolving to be a fundamental tool for faster and more cost-effective methods of lead drug discovery. SBDD aims to offer a computational replacement to traditional high throughput screening (HTS) methods of drug discovery. This 'virtual screening' technique utilises the structural data of a target protein in conjunction with large databases of potential drug candidates and then applies a range of different computational techniques to determine which potential candidates are likely to bind with high affinity and efficacy. It is proposed that high throughput SBDD (HT-SBDD) will significantly enrich the success rate of HTS methods, which currently fluctuates around ~1%. This project is centered around the design and implementation of quantum mechanics into the current drug discovery pathway. Here I plan to investigate and develop automated tools for the application of Fragment Molecular Orbital (FMO) techniques for early-stage high throughput Structure-based drug design (SBDD) as an alternative to traditional HTS methods.

During the FMO process, a computational replica of the target biological system is fragmented into groups of atoms which are treated as one, due to the nature of the calculations being performed this fragmentation is the only way that allows quantum mechanical calculations to be done on a system that is large enough to be biologically interesting. For protein:drug complexes, the drug candidate is typically treated as one fragment, and individual residues of the target protein as other fragments. The output is a series of interaction energies, where the total interaction energy for any given drug candidate is equal to the sum of interactions between the individual fragments. I have also applied Molecular Dynamics simulations to study the dynamic relationship of biological systems. Here I present results of implementing these techniques with the output of identifying significant binding patches and an explanation of a mutation study first identified in 1994.

BHive+: An integrated computational toolkit for B-cell systems immunology

Joseph **NG**, Franca Fraternali

Institute of Structural and Molecular Biology
Research Department of Structural & Molecular Biology, Division of Biosciences, UCL

Email address: joseph.ng@ucl.ac.uk

Abstract:
Antibodies, produced by specialised immune cells known as B-cells, play a crucial role in our immune response by neutralising pathogens and recruiting effector cells to clear these pathogens, thereby protecting us from infections. The immunology community has unmet data analysis needs on the annotation of antibody sequences, structures, and B-cell subtypes. The community also lacks easy-to-use software tools to extract novel insights from high-dimensional data, and apply cutting-edge artificial intelligence and machine learning methods to devise B-cell therapies and the design of antibody therpaeutics. This is particularly timely as the COVID-19 pandemic and recent advancement in cancer immunotherapy demonstrate that therapeutic antibodies are a vital pharmaceutical resource, with thousands in development, primarily IgG1, but none IgA or IgE. Understanding the functions of different classes of antibodies could expand their potential utility for treating different disorders.

We are building BHive+, an integrated software ecosystem dedicated for systems immunology of antibodies and B-cells. This growing repository of tools include a user-friendly webserver (BRepertoire) for statistical analyses of large-scale antibody sequence “repertoires” across individuals facing different immune challenges, as well as software packages (sciCSR, BrepPhylo, BrepConvert) for detailed analyses of how these antibody repertoires were developed in vivo, as well as the dynamics of B-cell maturation. Going forward, we are evolving BHive into a central hub of computational resources dedicated to B-cell immunology, adding software to (i) enable full-length antibody structure modelling and energetic evaluation of these models; (ii) facilitate high-dimensional analysis of cell surface proteome for deep characterisation of B-cell subsets; (iii) democratisie the application of state-of-the-art protein language models in antibody engineering and specificity prediction. Leveraging our expertise in computational immunology, we bridge the research domains of cellular immunology and antibody discovery, and enable seamless integration of multimodal data types for B-cell and antibody-related research.

An integrative characterisation of proline cis and trans conformers in a disordered peptide

Alice J. **PETTITT**[1,3,4], Vaibhav Kumar Shukla[1,4], Angelo Miguel Figueiredo[1], Lydia S. Newton[1], Stephen McCarthy[2], Alethea B. Tabor[2], Gabriella T. Heller[1], Christian D. Lorenz[3], D. Flemming Hansen[1,4]

1) Department of Structural and Molecular Biology, Division of Biosciences, Gower Street, London, WC1E 6BT, United Kingdom
2) Department of Chemistry, Faculty of Mathematical and Physical Sciences, 20 Gordon Street, London, WC1H 0AJ, United Kingdom
3) Department of Physics, Faculty of Natural, Mathematical and Engineering Sciences, King's College London, Strand, London, WC2R 2LS, United Kingdom
4. The Francis Crick Institute, London, NW1 1BF, United Kingdom

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

Abstract:
Intrinsically disordered proteins (IDPs) often contain proline residues, which undergo cis/trans isomerisation. While molecular dynamics (MD) simulations have the potential to fully characterise the proline cis and trans sub-ensembles, they are limited by the slow timescales of isomerisation and force field inaccuracies. Nuclear magnetic resonance spectroscopy (NMR) can report on ensemble-averaged observables for both the cis and trans proline states, but a full atomistic characterisation of these sub-ensembles is challenging. Given the importance of proline cis/trans isomerisation for influencing the conformational sampling of disordered proteins, we employed a combination of all-atom MD simulations with enhanced sampling (metadynamics), NMR, and small-angle X-ray scattering (SAXS) to characterise the two sub-ensembles of the ORF6 C-terminal region (ORF6CTR) from SARS-CoV-2 corresponding to the proline-57 (P57) cis and trans states. A good agreement between the cis-P57 populations predicted from MD simulations in AMBER03ws was observed with populations obtained from experimental NMR data. Moreover, we observed good agreement between the radius of gyration predicted from the MD simulations in the AMBER03ws force field and that measured using SAXS. Our findings suggest that both the cis-P57 and trans-P57 conformations of ORF6CTR are extremely dynamic and that interdisciplinary approaches combining both multi-scale computations and experiments offer avenues to explore highly dynamic states that cannot be reliably characterised by either approach in isolation.

Characterisation of monomeric isoforms of amyloid-β using solvation-based Nuclear Magnetic Resonance techniques

Kirsten **SILVEY**[1], Angelo Mendes Pinto De Figueiredo, D. Flemming Hansen[1, 2] and Gabriella T. Heller[1]

[1] Department of Structural and Molecular Biology, Division of Biosciences, UCL Darwin Building, London.
[2] The Francis Crick Institute, London UK.

Email address: Kirsten.silvey.22@ucl.ac.uk

Abstract:
Alzheimer's disease is a fatal neurological condition characterised by plaques of aggregated amyloid-β (Aβ) found in patient's brains. Aβ aggregates in a complex, multistage process from an intrinsically disordered monomeric peptide to highly ordered fibril. The two most common forms of Aβ (Aβ40 and Aβ42) have strikingly different aggregation rates, seen in the earliest stages of aggregation, with Aβ42 more commonly associated with toxicity. The differences in these early-stage aggregation rates hint at structural differences in the monomeric forms of the peptides. Despite extensive structural studies into monomeric Aβ, there have been no significant differences identified in their monomeric states, suggesting the differences in aggregation rates are driven by the two additional C-terminal hydrophobic amino acids in Aβ42 or in early aggregated species. Using CLEANEX Nuclear Magnetic Resonance (NMR) to monitor residue-specific hydrogen exchange rates, I have identified differences between the monomeric forms of Aβ40 versus Aβ42 upstream of the hydrophobic C-terminal region, hinting at differences in solvent protection between the two peptides. I have implemented key controls for pH and aggregation into my experimental workflow, making it a more reproducible and reliable platform. These results highlight the sensitivity of using CLEANEX as a method to study challenging intrinsically disordered systems in contrast to more traditional NMR approaches such as chemical shift perturbations.

Making 'antisense' of the PhoPR two-component system in the Mycobacterium tuberculosis complex

Jennifer **STIENS**[1], Matthew W. Scurlock[2], Sharon L. Kendall[2], Irene Nobeli[1]

[1]Institute of Structural and Molecular Biology, Biological Sciences, Birkbeck, University of London, UK
[2]Centre for Emerging, Endemic and Exotic Diseases, Pathobiology and Population Sciences, Royal Veterinary College, Hatfield, UK

Email address: jstien01@student.bbk.ac.uk

Abstract:
Mycobacterium tuberculosis (*M.tb*) and *Mycobacterium bovis* (*M.bovis*) have nearly identical genomes but different host specificities and pathogenic profiles. Both lineages of the *Mycobacterium tuberculosis complex* (MTBC) demonstrate pervasive antisense transcription which increases in stress conditions; however, comparative transcriptomics reveal differences in the use of transcription start sites and expression of regions outside of protein-coding genes, including antisense RNA. We propose that these understudied non-coding elements have an influence on transcription and/or translation in a host-specific manner. We used computational methods with publicly-available *M.tb* RNA-seq data to identify and cluster all expressed transcripts in a range of conditions and identified an antisense transcript found opposite the *phoR* gene which is highly expressed in acid and stationary growth conditions. The PhoR sensor-kinase is part of a two-component system that controls the cell response to acid stress by activating the PhoP transcription factor and is active and essential for virulence in both *M.tb* and *M.bovis* despite a potentially deleterious SNP in *M.bovis phoR*. We used a CRISPR-interference system to create a *M.tb* strain with inhibited expression of the *phoR*-antisense and evaluated the effects of this inhibition on the transcriptome using RNA-seq.

Nanoscale mapping of biosynthesis and molecular insertion in the Escherichia coli outer membrane

W. **TREWBY**[1,2], C. Kleanthous[3] and B. W. Hoogenboom[1,2]

[1]London Centre for Nanotechnology, University College London, London, WC1H 0AH, UK
[2]Department of Physics and Astronomy, University College London, London, WC1E 6BT, UK
[3]Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK

Email address: william.trewby.10@ucl.ac.uk

Abstract
The outer membrane (OM) of Gram-negative bacteria is a complex, asymmetric cocktail of lipopolysaccharides (LPS), outer membrane proteins (OMPs) and phospholipids. These components provide a thick diffusion barrier that represents a formidable first line of defence against antibiotic compounds [1], while bearing mechanical loads [2] and allowing for rapid cell growth and division.
The OM must thus form a robust barricade whilst being sufficiently fluid to allow for biosynthesis and rapid adaptation of the membrane components. A fundamental understanding of how these two apparently conflicting functions arise and are maintained is crucial in the escalating fight against antimicrobial resistance, but is challenging to unpick, especially at the nanoscale and in growing and dividing cells.
Here, we use high-resolution atomic force microscopy (AFM) to directly map the organisation and dynamics of the outer membrane at the scale of single OMP trimers in live Escherichia coli cells [3]. We demonstrate that the heterogeneous interactions of OM components lead to proteinaceous networks across the cell interspersed with LPS-rich domains of ~10s nm length, each with distinct dynamic and biophysical properties. These domains appear unevenly distributed across the membrane, with a higher density close to the division plane, indicating an intriguing functional relevance to biosynthesis.
Further work will extend this nanoscale domain mapping with physical labels that allow the unambiguous location of monomeric proteins and point towards the action of OMP insertion machinery such as the BAM complex.

References
[1] H. Nikaido, Microbiol. Mol. Biol. Rev. 67, 593 (2003).
[2] E. R. Rojas et al., Nature 559, 7715 (2018).
[3] G. Benn et al., Proc. Natl. Acad. Sci. 118, e2112237118 (2021).

Using purified Xenopus tubulin as a model to decipher the tubulin code

Luca A **TROMAN***[1], Ella de Gaulejac[2] , Simone Reber[2], Carolyn A. Moores[1]

[1] School of Natural Sciences, Birkbeck University of London, London, United Kingdom, [2] MPI for Infection Biology, Virchowweg 12, 10117 Berlin, Germany.

Email address: l.troman@bbk.ac.uk

Abstract:
Microtubules are a core component of the eukaryotic cytoskeleton and rely on the dynamic assembly and disassembly of microtubules from heterodimers of α- and β- tubulin. In many vertebrates there are multiple isoforms of tubulin, which whilst highly conserved, have minor differences in amino acid sequence as well as post-translational modifications (PTMs); often referred to as the tubulin code. Recent progress has allowed better understanding of the importance of isoform differences, but it remains challenging to physiologically contextualise these. Xenopus oocytes express fewer tubulin isoforms offering the simplicity necessary to interrogate the effect of isoform distribution, whilst remaining physiologically relevant to cellular processes such as spindle assembly. To investigate the structural basis for observed differences in the tubulin dynamics between species we used cryo-electron microscopy. We determined <4 Å resolution reconstructions of the microtubule lattices from each species. We generated models for each and analysed the longitudinal and latitudinal interactions between tubulin dimers. Mapping species-specific differences onto these models allowed us to correlate them with variations in tubulin dynamic properties.

Different modes of interbacterial competition elicit distinct host responses and health outcomes

Mollie **VIRGO**[1,2], Serge Mostowy[2,4], Brian Ho[1,3,4]

[1] Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck College, London, UK
[2] Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK
[3] Institute of Structural and Molecular Biology, Division of Biosciences, University College London, London, UK
[4] Co-corresponding authors

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

Abstract:
Interbacterial competition is known to shape the microbial communities found in the host, however the interplay between this competition and host defense are less clear. Here, we use the zebrafish hindbrain ventricle (HBV) as an in vivo platform to investigate host responses to defined bacterial communities with distinct forms of interbacterial competition. We found that antibacterial activity of the type VI secretion system (T6SS) from both Vibrio cholerae and Acinetobacter baylyi can induce host inflammation and sensitize the host to infection independent of any individual effector. Chemical suppression of inflammation could resolve T6SS-dependent differences in host survival, but the mechanism by which this occurred differed between the two bacterial species. By contrast, colicin-mediated antagonism induced a negligible host response despite being a more potent bacterial killer. Altogether, these results provide insight into how different modes of interbacterial competition in vivo affect the host in distinct ways.

The Encyclopaedia of protein Domains (TED) from AlphaFold2-predicted structures: expansion of CATH domain space and insights into structural diversity

Lau AM[1$], Bordin N[2$], Kandathil SM[1], Sillitoe I[2], **WAMAN** VP[2&], Wells J[2,3], Orengo CO[2*] and Jones DT[1,2*]

[1]Department of Computer Science, University College London, London, WC1E 6BT, UK
[2]Institute of Structural and Molecular Biology, University College London, London, WC1E 6BT, UK
[3]Centre for Artificial Intelligence, University College London, London, WC1V 6BH, UK

[&]Presenting author
$ Authors contributed equally
*Corresponding authors: c.orengo@ucl.ac.uk , dtj@cs.ucl.ac.uk

Email address: v.waman@ucl.ac.uk

Abstract:
The AlphaFold Structure Database (AFDB) currently provides predictions for > 214 million three-dimensional structures of full-length proteins. Protein structure is composed of one or more domains i.e. independent folding units that can be found in multiple structural and functional contexts. We present The Encyclopedia of Domains (TED) which is the first structural-based resource to systematically classify domains from the AFDB [1,2].
TED is being developed, as a collaboration between the Orengo and Jones group. TED combines state-of-the-art deep learning-based domain detection, structure comparison (Foldseek[3]) and fold recognition (FoldClass) algorithms to identify and classify domains across the structures from AFDB. We used novel deep learning-based domain detection method by the Jones group (Merizo[4]), the Orengo group (Chainsaw [5]) and UniDoc [6], to segment AFDB protein structures into protein domains.
TED contains over 370 million protein domains, providing domain structure coverage to over 1 million taxa/species. Notably, 80% of the TED domains exhibit similarities with known superfamilies in the CATH database [7], expanding the CATH resource by over 600-fold. We identified over 6,000 domains with potentially new folds, some of which have unique symmetric protein architectures not seen previously (e.g. 11-helix propeller and 11-bladed beta-propeller).
TED as a unique domain resource derived from AFDB, will aid in a multitude of downstream analyses including identification of remote homologues and understanding of structural and functional diversity across protein superfamilies. We demonstrate an analysis of TED data

providing structural insights into the differences and commonalities in substrate-binding specificities between CoA-dependent acyltransferases in the context of different pathogens.
TED will be a valuable resource for the bioscience community that provides a functional interface to the AFDB. TED data will also be made available via 3D-Beacons at the EBI, as well as a dedicated resource, significantly enriching CATH superfamilies.

References:
[1]. Lau et al. Exploring structural diversity across the protein universe with The Encyclopedia of Domains. bioRxiv (2024).03.18.585509; doi: https://doi.org/10.1101/2024.03.18.585509
[2]. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 50, D439–D444 (2022).
[3]. van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. (2023)
[4]. Lau, et al. & Jones, D.T. Merizo: a rapid and accurate protein domain segmentation method using invariant point attention. Nat Commun 14, 8445 (2023).
[5]. Wells, J.et al. Chainsaw: protein domain segmentation with fully convolutional neural networks. bioRxiv 2023.07.19.549732 (2023)
[6]. Zhu K, et al. A unified approach to protein domain parsing with inter-residue distance matrix. Bioinformatics ;39(2):btad070 (2023)
[7]. Waman VP, et al. CATH 2024: CATH-AlphaFlow Doubles the Number of Structures in CATH and Reveals Nearly 200 New Folds. J Mol Biol. Mar 27:168551. doi: 10.1016/j.jmb.2024.168551 (2024)

Acknowledgements:
This work was funded by BBSRC grants: BB/T019409/1 (D.T.J.; A.M.L.); BB/W008556/1 (D.T.J.; S.M.K.); BB/W018802/1 (I.S) and Wellcome Trust grants 221327/Z/20/Z (V.P.W; N.B.). J.W. acknowledges the studentship awards from the Health Data Research UK-The Alan Turing Institute Wellcome PhD Programme in Health Data Science (218529/Z/19/Z).

On the importance of being amidated: Analysis of the role of the conserved C-terminal amide of amylin in amyloid formation and cytotoxicity

Tangweina **YANG** [1] , Ivan Filippov [1] , Lakshan Manathunga [2,3], Aria Baghai [1] , Amandine Maréchal [1,4] , Daniel P. Raleigh [1,2,3] , Alexander Zhyvoloup [1]

1 Institute of Structural and Molecular Biology, Division of Biosciences, University College London, Gower Street, London WC1E 6BT, United Kingdom
2 Laufer Center for Quantitative Biology, Stony Brook University, Nicolls Road, Stony Brook, NY 11790, United States
3 Department of Chemistry, Stony Brook University, Nicolls Road, Stony Brook, NY 11790, United States
4 Institute of Structural and Molecular Biology, Division of Biosciences, Birkbeck College, London WC1E 7HX, United Kingdom

Email address: tangweina.yang.18@ucl.ac.uk

Abstract:
The polypeptide hormone Amylin (also known as islet amyloid polypeptide) plays a role in regulation of glucose metabolism, but forms pancreatic islet amyloid deposits in type 2 diabetes. The process of islet amyloid formation contributes to β-cell dysfunction and the development of the disease. Amylin is produced as a pro-from and undergoes processing prior to secretion. The mature hormone contains an amidated C-terminus. Analysis of an alignment of vertebrate amylin sequences reveals that the processing signal for amidation is strictly conserved. Furthermore, the enzyme responsible for C-terminal amidation is found in all of these organisms. Comparison of the physiologically relevant amidated form to a variant with a free C-terminus (Amylin-COO−) shows that replacement of the C-terminal amide with a carboxylate slows, but does not prevent amyloid formation. Pre-fibrillar species produced by both variants are toxic to cultured β-cells, although hAmylin-COO− is moderately less so. Amyloid fibrils produced by either peptide are not toxic. Prior work (ACS Pharmacol. Translational. Sci. 1, 132–49 (2018)) shows that Amylin-COO− exhibits a 58-fold reduction in activation of the Amylin1 receptor and 20-fold reduction in activation of the Amylin3 receptor. Thus, hAmylin-COO− exhibits significant toxicity, but significantly reduced activity and offers a reagent for studies which aim to decouple hAmylin's toxic effects from its activity. The different behaviours of free and C-terminal amidated Amylin should be considered when designing systems to produce the polypeptide recombinantly.

Comparative in-situ characterisation of respiratory supercomplexes in human mitochondria

Ines **ZOUHAIR** [1,2], Dr Cristina Ugalde[3], Dr Amandine Maréchal[1] and Dr 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:
The mitochondrial respiratory chain (MRC) is composed of four protein complexes, complexes I to IV, which through a series of coupled redox reactions build a proton gradient across the inner mitochondrial membrane that drives ATP synthesis by the ATP synthase. The MRC protein complexes assemble into supercomplexes (SC), which differ between species and in mammals tissue types. Changes in SC population and stoichiometry have been as well observed in several pathologies. However, the biological significance of SC in relation to energy production still remains unclear.

One of the proteins involved in SC formation is COX7A, a subunit of complex IV or cytochrome c oxidase. COX7A exists as 3 isoforms, COX7A1, COX7A2 and COX7A2L, which are differentially expressed depending on tissue type. 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 determine the role of the COX7A isoforms and different MRC organisations in energy production, by gaining an understanding of the molecular, ultrastructural and bioenergetic changes in the mitochondrial network caused by the COX7A isoforms. We will be applying a range of high-end techniques, including cryo Electron Tomography, Cross-linking Mass Spectrometry, cryo Soft-X-Ray Tomography and different growth conditions to obtain an in situ and multi-scale characterisation of those COX7A mutant cell lines.