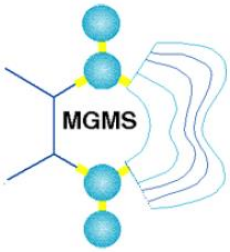

MGMS and RSC MMG
Young Modellers' Forum 2024
PROGRAMME & ABSTRACTS

Programme of Oral Presentations

09:15-09:50	Registration and Coffee
09:50-10:00	Introduction/Welcome
10:00-10:20	Phospholipid transport to the bacterial outer membrane through an envelope-spanning bridge Robbie Clark, University of Oxford
10:20-10:40	Broad-Spectrum Boronic Acid Inhibitors Targeting Serine β-Lactamases: Molecular Optimization and Computational Screening Using LSTM and DFT Shuang Chen, UCL
10:40-11:00	Automated Adaptive Absolute Binding Free Energy Calculations Finlay Clark, University of Edinburgh
11:00-11:30	Flash Poster presentations
11:30-11:50	Break
11:50-12:10	Theoretical tautomer prediction as a testbed for theory-based experimental uncertainty analysis Michael Strobl, Technische Universität Dortmund
12:10-12:30	Nanoscale Non-adiabatic Dynamics Simulation of Charge Generation in Organic Solar Cells Filip Ivanovic, UCL
12:30-12:50	Getting the Most from Low Resolution Data: A Semi-Automated Pipeline for the Generation of Atomistic Protein Ensembles in Accordance with SAXS Data Cameron Brown, University of Southampton
12:50-14:10	Lunch and Posters
14:10-14:30	Uncovering a dual mechanism controlled by stereoisomerism in alginate lyase William Houppy, Northumbria University
14:30-14:50	How Different PTMs Affect the Conformation of Peroxiredoxin 2? Zhicheng He, University of Nottingham
14:50-15:10	RhoB: an unstable G protein mutated in bladder cancer Morwenna Hall, University of Cambridge
15:10-15:30	Break
15:30-15:50	Structure-guided disulfide engineering restricts antibody conformation and flexibility to elicit TNFR agonism in anti-cancer therapeutics. Isabel Elliott, University of Southampton
15:50-16:10	Predicting the conformational flexibility of antibody CDRs Fabian Spoendlin, University of Oxford
16:10-16:30	Enhancing enzyme production yield using rational design Margaux Haon, University of Liverpool
16:30-16:45	Deliberation and Prizes
16:45	Close

Talk 1

Phospholipid transport to the bacterial outer membrane through an envelope-spanning bridge

Robert Clark[a], Benjamin F. Cooper[b], Anju Kudhail[b], Gira Bhabha[c], Damian C. Ekiert[c,d], Georgia L. Isom[b] and *Syma Khalid*[a]

[a]Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK.
[b]Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, UK.
[c]Department of Cell Biology, New York University School of Medicine, New York, NY, 10016.
[d]Department of Microbiology, New York University School of Medicine, New York, NY, 10016.

The outer membrane of Gram-negative bacteria is an essential barrier that promotes antimicrobial resistance and pathogenesis. Understanding how this barrier is maintained could open novel therapeutic targets. We have used a combination of atomistic and coarse-grained simulations to study YhdP, a protein involved in the maintenance of a functional outer membrane[1]. By combining experimental evidence from our collaborators and our own molecular dynamics results, we provide compelling evidence for YhdP being a bridge between the two membranes of *E. coli* via which phospholipids are moved.

We use Alphafold to model the structure of YhdP and observe a 60 β strands arranged in continuous sheet. Pore dimension analysis shows this extended β sheet forms a continuous, lipophilic channel. Simulations of an N-terminal fragment show the inner membrane interactions are largely mediated by a transmembrane helix and a helix that lies parallel to the membrane, that inserts into the hydrophobic core of the membrane. These interactions permit spontaneous entry of phospholipids from the inner membrane into the groove of YhdP, demonstrating a consistent phospholipid uptake pathway which we observe across systems. We then identify potential gating mechanisms within the N-terminus that could act to regulate phospholipid entry into the groove of YhdP. We study the C-terminus using a randomised bilayer assembly protocol to resolve unclear membrane interaction predictions, finding the C-terminus interacts primarily through two amphipathic helices. Our simulations of YhdP in a model double membrane system show these identified terminal interactions are sufficient to stably bind YhdP in a double membrane. We confirm spontaneous phospholipid entry still occurs through the same pathway in a double membrane system.

Our results support a model where YhdP bridges the two membranes of Gram-negative bacteria and provides a hydrophobic environment for the transport of phospholipids to the outer membrane. We propose a gating mechanism helps prevent unregulated anterograde lipid transport and that additional protein partners may not be required for lipid uptake.

References:
[1] Grimm J, Shi H, Wang W, Mitchell AM, Wingreen NS, Huang KC, Silhavy TJ. 2020 *PNAS* 117:26907–26914

Talk 2

Broad-Spectrum Boronic Acid Inhibitors Targeting Serine β-Lactamases: Molecular Optimization and Computational Screening Using LSTM and DFT

Shuang Chen[a], *Shozeb Haider*[a,b]

[a] Department of Pharmaceutical and Biological Chemistry, School of Pharmacy, University College London, London, United Kingdom
[b] UCL Centre for Advanced Research Computing, London UK.

Boronic acid inhibitors represent promising candidates for combating serine β-lactamases, addressing antibiotic resistance caused by the presence of β-lactamases. However, a diverse range of β-lactamase types exists. Moreover, bacteria, particularly *Gram*-negative pathogens, are constantly evolving, leading to the emergence of new variants of β-lactamases. To develop inhibitors with broad-spectrum capabilities and tripodal-targeting potential—capable of inhibiting Class A, C, and D β-lactamases—we utilized a long short-term memory (LSTM) network to construct a molecular generator. Additionally, transfer learning was applied using a curated set of experimentally validated inhibitors, thereby enabling the model to acquire the characteristics of efficacious inhibitors targeting various β-lactamases. The model was fine-tuned using different hyperparameters, and noise vectors sampled from a zero-centered normal distribution were added to the latent representations of seed molecules to generate new data. The generated molecules were evaluated based on validity, diversity, novelty, and similarity to seed molecules, leading to the selection of the optimal generator. As a result, 13,612 new molecules were produced, and after an initial screening, 2,845 were selected for further analysis. These molecules were subsequently docked to three representative proteins: KPC-2 (Class A), PDC-3 (Class C), and OXA-24/40 (Class D). GFN2-xTB was used for geometry optimization of the ligands, proteins, and complexes, while density functional theory (DFT) was employed to calculate binding free energies according to the equation $\Delta G_{bind} = \Delta G_{complex} - (\Delta G_{protein} + \Delta G_{ligand})$. Finally, the top 20 molecules were identified based on their binding affinities and selected for further study.

References:
[1] Zhou, J., Stapleton, P., Haider, S., Healy, J.: Boronic acid inhibitors of the class a β-lactamase KPC-2. *Bioorg. Med. Chem.*, (2018) **26**(11), 2921–2927.
[2] Arús-Pous, J., Johansson, S.V., Prykhodko, O., Bjerrum, E.J., Tyrchan, C., Reymond, J.L., Chen, H. and Engkvist, O.: Randomized SMILES strings improve the quality of molecular generative models. *J Cheminform*, (2019) **11**, 1-13.

Talk 3

Automated Adaptive Absolute Binding Free Energy Calculations

Finlay Clark[a], *Graeme R. Robb[b], Daniel J. Cole[c], Julien Michel[a]*

[a]EaStCHEM School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh EH9 3FJ, United Kingdom.
[b]Oncology R&D, AstraZeneca, Cambridge CB4 0WG, United Kingdom.
[C]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom.

Alchemical absolute binding free energy (ABFE) calculations offer a rigorous method to calculate the binding affinities of varied molecules. They can address problems such as optimising ligand selectivity or promiscuity, and predicting the functional response of ligand binding. They are also useful as a final, accurate filter in virtual screening. However, ABFE calculations are computationally expensive, demanding tens to hundreds of GPU hours for a single molecule. To become a routine tool in industrial drug discovery, ABFE calculations must become more efficient and automated.
I will discuss algorithms to improve the efficiency of ABFE calculations by reducing both user and simulation time.[1] In particular, I will discuss the automated selection of the number and spacing of simulations intermediate between the thermodynamic end states, the detection of equilibration, and the adaptive allocation of simulation time to achieve minimum standard error of the final free energy estimate. I will address the automated selection of restraints between the receptor and ligand, which can significantly influence the performance of ABFE calculations.[2] Finally, I will discuss the implementation of these algorithms in an open-source software package, A3FE (Automated Adaptive Absolute binding Free Energy calculator).[1] We hope these approaches reduce the barrier to the routine use of absolute binding free energy calculations in drug discovery.

References:
[1] F. Clark, G. R. Robb, D. J. Cole and J. Michel, *J. Chem. Theory Comput*., 2024, XXXX, DOI:10.1021/acs.jctc.4c00806.
[2] F. Clark, G. Robb, D. J. Cole and J. Michel, *J. Chem. Theory Comput.*, 2023, 19, 3686–3704.

Talk 4

Theoretical tautomer prediction as a testbed for theory-based experimental uncertainty analysis

Michael Strobl, Nicolas Tielker, Christian Chodun, Lukas Eberlein, Julia Jasper-Eberlein, *Stefan M. Kast*

Technische Universität Dortmund, Otto-Hahn-Str. 4a, 44227 Dortmund

Aqueous tautomer equilibria are highly relevant to a wide range of chemical and biological issues, e.g. DNA base-pairing or biological activity of drugs. Despite the importance, experimental data is scarce and potentially unreliable due to difficulties associated with measurements. Complementary theoretical approaches aim at reliability and predictive power, but developing models without reliable datasets is difficult. Here, different approaches to calculate free energies of tautomerization are applied to several tautomer equilibria datasets available, primarily the Tautobase.[1] The first ansatz is a direct approach derived solely from solution-phase properties provided by the Embedded Cluster Reference Interaction Site Model (EC-RISM).[2] In the second approach, an indirect thermodynamic route is chosen, which is complemented by gas phase free energies derived from either state-of-the-art DLPNO-CCSD(T) calculations or quantum-based machine learning models like ANI-1ccx. In this context, application of a machine learning-based optimization of EC-RISM is also investigated.

Subsequently, our best-performing model combinations are compared with the literature, facilitating construction of a consensus dataset and various statistical analyses for curating the reference datasets. Data points where the various theoretical approaches agree within a certain range but deviate strongly from the experimental reference are then analysed in more detail. Hence, we are able to identify suspicious database entries that may be based on problematic measurements or incorrect annotations. As a key result, an ordered set of tautomer pairs with increasing experimental uncertainty is produced, measured by increasing consensus prediction error. This curated dataset will more faithfully allow for training and evaluating novel computational methods.

References:

[1] O. Wahl, T. Sander, *J. Chem. Inf. Model.*, (2020), 60, 1085-1089.
[2] N. Tielker, L. Eberlein, G. Hessler, K.F. Schmidt, S. Güssregen, S.M. Kast, *J. Comput. Aided Mol. Des.*, (2021), **35**, 453-472.

Talk 5

Nanoscale Non-adiabatic Dynamics Simulation of Charge Generation in Organic Solar Cells

Filip Ivanovic[a], Samuele Giannini[b], Wei-Tao Peng[c], *Jochen Blumberger*[a]

[a]Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT
[b]Via Madonna del Piano 10, 50019 Sesto Fiorentino
[c]1727, Section 4, Taiwan Boulevard, Xitun District, Taichung City,

Organic solar cells (OSCs) have long been the subject of intensive research, as their easily processible constituent materials suggest a commercially viable alternative to their inorganic counterparts. In OSCs, excitation by light yields a tightly bound electron-hole pair, termed an exciton. Two-component cells, where two species are separated by an interface, constitute OSCs' greatest leap towards commercialisation, with their highest efficiencies exceeding 19%. Although such an interface is crucial for the dissociation of excitons into a sufficient yield of free charges, understanding the exact mechanism by which free charges are generated across the interface remains a formidable theoretical challenge, with contradictory results having been reported in the literature[1].

The complexity of such systems precludes a complete treatment by analytical theories, and instead requires first-principles quantum dynamics simulations, which must also be fast enough to access time scales comparable to experiment.

Here, we use an in-house non-adiabatic molecular dynamics package, termed X-SH[2], to simulate charge generation in an oligothiophene-perylene diimide interface, on experimentally relevant time and length scales. Our use of a DFT-parameterised Hamiltonian, which is updated on-the-fly, bypasses the need for explicit electronic structure calculations during the dynamics.

We elucidate the mechanism of charge generation in such an interface by modelling the excitons and charges with an explicit electronic wavefunction. This allows us to track their locations in real-time, on a microscopic scale that often cannot be resolved with photochemical experiments. We leverage the computational speed and flexibility of X-SH to identify key physical parameters affecting the efficiency of charge generation, and translate this into design rules to guide the synthesis of OSCs with yet higher efficiencies.

References:
[1] A. Armin, W. Li, O. J. Sandberg, Z. Xiao, L. Ding, J. Nelson, D. Neher, K. Vandewal, S. Shoaee, T. Wang, H. Ade, T. Heumüller, C. Brabec, *P. Meredith*, *Adv. Energy Mater*. 2021, **11**, 2003570
[2] WT. Peng, D.Brey, S.Giannini, D.Dell'Angelo, I.Burghardt, *J.Blumberger, J. Phys. Chem. Lett.* 2022, **13**, 7105-7712

Talk 6

Getting the Most from Low Resolution Data: A Semi-Automated Pipeline for the Generation of Atomistic Protein Ensembles in Accordance with SAXS Data

Cameron Brown, Josh McKeown, Arron Bale, Hayden Fisher, Thomas Bowden, Christopher Prior, *Robert Rambo, Jonathan Essex*

School of Chemistry and Chemical Engineering, University of Southampton, Southampton, SO17 1BJ[a]
Diamond Light Source, Harwell Science and Innovation Campus, Didcot, OX11 0DE[a]
School of Mathematical and Computer Science, Durham University, DH1 3LE[b]

Small Angle X-ray Scattering (SAXS) is a low-resolution biophysical technique that reveals the shapes, sizes, and conformations of biological macromolecules in near-native environments. In contrast, MD simulations provide atomic resolution, capturing the dynamic behaviour of these systems. Since the advent of Alphafold, it has never been easier to approach SAXS experiments with a structural hypothesis. Despite this, a large percentage of Alphafold and crystal structure predictions do not fit the corresponding solution SAXS profiles. To address this problem, we apply an integrative, semi-automated modelling protocol utilising SAXS, MD and a novel SAXS-driven structure optimisation algorithm developed by collaborators at Durham University, to bridge the gap between static protein structure prediction and the dynamics reality.[1]

We demonstrate applicability through two use cases. Firstly, demonstrating domain flexibility within the Rift Valley Fever Gn antigen, which is hypothesized to aid higher order assembly formation and viral cell entry. While an open conformation Gn crystal structure was not obtainable, MD simulations of the closed Gn structure show domain opening, supported by the experimental SAXS. Our second use-case builds on previous X-ray crystallographic and SAXS studies on hinge disulfide-engineered IgG2 $F(ab)_2$ fragments holding promise for cancer therapeutics.[2] MD simulations of crystal $F(ab)_2$ structures produced atomistic ensembles that did not cover the necessary radius of gyration (Rg) range from which the ensemble selection program GAJOE could reliably sample. Here, we recover the necessary ensemble Rg range through SAXS-driven structural optimization followed by MD. $F(ab)_2$ structures are restrained by interatomic distances between the C-alpha atoms of opposing disulfide bonding cysteine residues, while extending the F(ab) arms outwards in accordance with the experimental scattering data. Following all-atom modelling, secondary structure is largely maintained while the resultant MD ensembles, reweighted by protocols of maximum parsimony and Baysesian-maximum entropy, display a range of Rg values in line with that suggested by the experimental scattering.

References:
[1] C. Prior, O. Davies, D. Bruce and E. Pohl, *Journal of Chemical Theory and Computation*, (2020), **16**, 1985–2001.
[2] C. Orr, H. Fisher, X. Yu, C. Chan, Y. Gao, P. Duriez, S. Booth, I. Elliott, T. Inzhelevskaya, I. Mockridge, C. Penfold, A. Wagner, M. Glennie, A. White, J. Essex, A. Pearson, M. Cragg and I. Tews, *Science Immunology*, (2022), **7**, 73–84.

Talk 7

Uncovering a dual mechanism controlled by stereoisomerism in alginate lyase

William Houppy[1], *Warispreet Singh[1]*, Claire Morley[2,3], Hamish C L Yau[4], Neil J Lant[4], Gary Black[1], Jose Munoz-Munoz[1]

[1]Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK.
[2]Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, UK.
[3]School of Natural and Environmental Sciences, Newcastle University, Devonshire Building, Newcastle upon Tyne, NE1 7RU, UK.
[4]Procter and Gamble, Newcastle Innovation Centre, Whitley Road, Newcastle upon Tyne, NE12 9BZ, UK.

Alginate lyases (AL) are enzymes capable of depolymerizing alginate, a polysaccharide extracted from seaweed. These enzymes have gained increasing importance in various industries, including food production, laundry detergence, and biomass valorization. In this study, two efficient ALs from the CAZy PL7 family, identified in Flavobacterium species, were analyzed through a mechanistic study to explain the experimental variations in activity.

Alginate consists of two monomers, β-D-mannuronate (M) and its C5 epimer α-L-guluronate (G), connected through β-1,4 glycosidic bonds. Comparative analysis of one of the ALs (Aly30) with other previously described ALs revealed a conserved structure, using an AlphaFold prediction, and a similar active site formed by a histidine (HIS) and a tyrosine (TYR) residues.

ALs typically operate via a β-elimination mechanism, where HIS acts as a base and TYR as an acid (HIS/TYR) [1], attacking the α-hydrogen at the C5 position. However, initial docking studies of Aly30 with a poly-M substrate suggested that this mechanism may not be feasible. Instead, docking results pointed toward an alternative mechanism, where TYR, in its deprotonated state, alternates between acting as an acid and a base (TYR/TYR) [2]. This was further supported by molecular dynamics simulations.

Interestingly, when Aly30 was studied with a poly-G ligand, docking and simulation data suggested that the conventional HIS/TYR mechanism is more likely to occur. As the only difference between M and G is the configuration of the C5 position, the stereoisomerism appears to dictate the enzymatic mechanism and may explain the observed variation in alginate degradation activities.

We are currently validating our computational findings with additional wet-lab experiments and conducting quantum chemistry studies to draw the energy profiles of the proposed mechanisms.

References:
[1] F. Xu, X.-L. Chen, X.-H. Sun, F. Dong, C.-Y. Li, P.-Y. Li, H. Ding, Y. Chen, Y.-Z. Zhang and P. Wang, Structural and molecular basis for the substrate positioning mechanism of a new PL7 subfamily alginate lyase from the arctic, *Journal of Biological Chemistry*, 2020, **295**, 16380–16392.
[2] M. Vuillemin, B. Pilgaard, E. Kiehn, F. Fredslund, D. H. Welner, A. S. Meyer, F. L. Aachmann and C. Wilkens, Glucuronan lyases from family PL7 use a Tyr/Tyr syn β-elimination catalytic mechanism for glucuronan breakdown, *Chem. Commun*., 2024, **60**, 440–443.

Talk 8

How Different PTMs Affect the Conformation of Peroxiredoxin 2?

Zhicheng He, *Charles Laughton, Mattéa Finelli*

School of Pharmacy, Biodiscovery Institute, University Park, Nottingham, NG7 2RD

Peroxiredoxin 2 (Prdx2), an antioxidant enzyme, possesses two crucial cysteines: the peroxidative cysteine (C_P) which reacts with approaching peroxides and the resolving cysteine (C_R) potentially capable of forming a disulfide bond with C_P. This disulfide bond formation process requires a global conformational shift from the fully folded (FF) state to the local unfolding (LU) state. Not only disulfide bond but also other post-translational modifications (PTMs) have an impact on the conformational equilibrium between FF and LU.[1,2] To evaluate the influence, multiple MD simulation methods are used. Initially, string method was used to map the pathway between the two structures under different PTMs. With the pathway, weighted ensemble (WE) was applied to compute the kinetic profile of the pathway for each PTM. Among the intermediate states in WE trajectory, a 'meta stable state' where C_R is exposed to the solvent but not in the FF was observed. The knowledge of that state well explained the S-nitrosylation (SNO) process that C_P undergoes SNO first, which causes C_R to become solvent-exposed and subsequently undergo SNO as well. However, due to the large distance between FF and LU, WE was unable to facilitate FF to reach LU. Instead of computing the kinetic profile in the entire pathway, we are computing the relative free energy change between FF and LU for different PTMs to see whether a given PTM promotes or hinders the transformation from FF to LU by employing free energy perturbation (FEP).

References:
[1] Schroder, E., Isupov, M. N., *Structure,* (2000), **8**, 605-615.
[2] Wood, Z. A., Schroder, E., Robin Harris, J., & Poole, L. B., *Trends Biochem Sci,* (2003) **28**, 32-40.

Talk 9

RhoB: an unstable G protein mutated in bladder cancer

Morwenna L. Hall[a], Ari Karthikeyan[a], Frankie M. Houghton[a], *Tanadet Pipatpolkai[b]*, *Helen R. Mott[a], Daniel Nietlispach[a] and Richard W. Broadhurst[a]*

[a]Department of Biochemistry, University of Cambridge, 80 Tennis Court Rd, Cambridge, CB2 1GA
[b]Division of Physics and Applied Physics, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

RhoB, a Ras superfamily small G protein, is involved in actin cytoskeleton regulation, and its expression is tightly regulated in cells.[1] Unlike other small GTPases, RhoB is understudied in terms of its structure, dynamics, and role in physiology and disease.[1] Recent genome sequencing projects have identified *RHOB* mutations in bladder cancer that could affect GTPase function.[2] This work aims to provide molecular insights into RhoB and these newly found bladder cancer variants using molecular dynamics (MD) simulations, nuclear magnetic resonance (NMR) spectroscopy and other biochemical techniques.

Our *in vitro* thermal stability studies have shown RhoB is much less stable than related members of the Ras GTPase superfamily, and frequently observed mutations in bladder cancer altered its stability further. We employed MD simulations with GROMACS to predict how different Rho proteins (RhoB, Cdc42 and RhoA) unfold *in silico*, to complement the experimental NMR dynamics and biophysical data collected on RhoB. MD simulation trajectory analysis revealed RhoB was more mobile and reached higher root mean square deviation (RMSD) values than other Rho GTPases at multiple temperatures. Accordingly, RhoB exhibited higher root mean square fluctuation (RMSF) values at different small GTPase functional regions – these sequence differences could be influencing this increased mobility. Moreover, RhoB NMR backbone assignment has enabled NMR dynamics experiments to be conducted; further analysis will verify RhoB protein flexibility and/or decreased stability through reporting the motion of individual residues at different timescales.

These modelling results specifically help pinpoint and improve our understanding of the sources of instability in RhoB and its potential physiological relevance. New information about non-redundant roles of this central signalling protein implicated in cancer will be useful for further therapeutic research.

References:
[1] F. M. Vega and A. J. Ridley, *Small GTPases*, 2016, **9**, 384–393.
[2] The Cancer Genome Atlas Research Network, *Nature*, 2014, **507**, 315–322.

Talk 10

Structure-guided disulfide engineering restricts antibody conformation and flexibility to elicit TNFR agonism in anti-cancer therapeutics.

Isabel G. Elliott[a,b,c,d], Hayden Fisher[a,b,c,d,e], Christian M. Orr[f], H. T. Claude Chan[c], Tatyana Inzhelevskaya[c], C. Ian Mockridge[c], Christine A. Penfold[c], Patrick J. Duriez[c], Julie Herniman[a], Kri T. J. Müller[c], *Ivo Tews[b,d], Mark S. Cragg[c,d], Jonathan W. Essex[a,d]*

[a] School of Chemistry, University of Southampton; Southampton SO17 1BJ, UK.
[b] School of Biological Sciences, University of Southampton; Southampton SO17 1BJ, UK.
[c] Centre for Cancer Immunology, University of Southampton; Southampton, SO16 6YD, UK.
[d] Institute for Life Sciences, University of Southampton; Southampton SO17 1BJ, UK.
[e] European Synchrotron Radiation Facility, Grenoble, Cedex 9, 38043, France
[f] Diamond Light Source, Didcot, OX11 0FA, UK.

Immunostimulatory antibodies (ISAs) represent a promising strategy for cancer immunotherapy. By activating co-stimulatory molecules expressed on immune cells, such as tumour necrosis factor receptors (TNFRs), ISAs can enhance the immune response towards tumours, resulting in powerful anti-cancer effects. Antibodies comprise two antigen-binding domains linked to an effector domain through a disulfide-containing hinge. Of the four isotypes of human (h)IgG, previous work has shown that the hIgG2 isotype can deliver strong agonistic activity for ISAs, due to its unique hinge disulfide arrangement[1].

Here we employ an integrative approach to understand how structure and conformational dynamics affect agonistic activity of hIgG2 antibodies. We use cellular assays to ascertain agonistic activity, small angle X-ray scattering (SAXS) to assess flexibility and conformation, X-ray crystallography to determine protein structure and disulfide position, and molecular dynamics simulations with SAXS-guided ensemble reweighting to probe antibody conformational dynamics.

By modifying hinge disulfide patterns using cysteine-to-serine exchange mutations, we show that strong agonistic activity is associated with restricted global antibody flexibility and reduced conformational dynamics due to the presence of a cross-over of disulfides in the hinge[2]. We demonstrate this for antibodies targeting two different co-stimulatory receptors, CD40 and 4-1BB. We then use structure-guided approaches to design new hIgG2 antibody variants with novel disulfide patterns, to further restrict flexibility and enhance biological activity.

Together, these results demonstrate the importance of structure and conformational dynamics in developing efficacious ISAs and provide a strategy for the rational design of more powerful antibody therapeutics, and thus more effective anti-cancer treatments.

References:
[1] A. L. White, H. T. C. Chan, R. R. French, J. Willoughby, C. I. Mockridge, A. Roghanian, C. A. Penfold, S. G. Booth, A. Dodhy, M. E. Polak, E. A. Potter, M. R. Ardern-Jones, J. S. Verbeek, P. W. M. Johnson, A. Al-Shamkhani, M. S. Cragg, S. A. Beers, and M. J. Glennie, *Cancer Cell,* (2015), 27, 138-148.
[2] C. M. Orr, H. Fisher, X. Yu, C. H. Chan, Y. Gao, P. J. Duriez, S. G. Booth, I. G. Elliott, T. Inzhelevskaya, C. I. Mockridge, C. A. Penfold, A. Wagner, M. J. Glennie, A. L. White, J. W. Essex, A. R. Pearson, M. S. Cragg, and I. Tews, *Sci. Immunol.,* (2022), 7, eabm3723.

Talk 11

Predicting the conformational flexibility of antibody CDRs

Fabian Spoendlin[a], *Wing K. Wong[b], Guy Georges[b], Alexander Bujotzek[b], Charlotte M. Deane[a]*

[a] Department of Statistics, University of Oxford, 24-29 St Giles', Oxford, OX1 3LB.
[b] Large Molecule Research, Roche Pharma Research and Early Development, Roche Innovation Center Munich, Penzberg, Germany

Conformational changes give rise to functional properties of many classes of proteins. While a single, `static' protein structure can be predicted at high accuracy with current machine learning tools (e.g. AlphaFold), methods are severely limited at predicting structural flexibility [1]. A major factor limiting such predictions is the absence of large datasets to train such methods.

Here, we focus on the conformational flexibility of antibody CDRs, a functionally highly important protein motif. We mine the PDB and the Structural Antibody Database (SAbDab) for crystallographic evidence of the conformational flexibility of CDRs and structurally related loops across all classes of proteins. Through this approach we create a large dataset set of more than 20,000 loop motifs with determined flexibility.

Building on this dataset, we develop AbFlex, a model which shows strong predictive power for classifying if antibody CDRs are able to transition between multiple conformational states or adopt a single, stable conformation. Our method substantially outperforms AlphaFold2-based alternatives which have previously been described as predictors of protein flexibility [2]. Ablation studies highlight biophysical factors that influence the conformational flexibility of antibody CDRs. While sequence affects the tendency to adopt multiple conformations, we identify the arrangement of residues in the surrounding (structural context) as a key factor that drives loop flexibility.

The conformational flexibility of CDRs affects functional properties of antibodies including affinity and specificity, which are key properties that need to be optimise in therapeutic drugs. AbFlex, therefore, adds a valuable tool to investigate antibody function and assist the drug discovery process.

References:
[1] J. R. Riccabona, F. C. Spoendlin, A.-L. M. Fischer, J. R. Loeffler, P. K. Quoika, T. P. Jenkins, J. A. Ferguson, E. Smorodina, A. H. Laustsen, V. Greiff, S. Forli, A. B. Ward, C. M. Deane and M. L. Fernández-Quintero, *Structure*, 2024, S0969212624003708.
[2] D. del Alamo, D. Sala, H. S. Mchaourab and J. Meiler, *eLife*, 2022, **11**, e75751.

Talk 12

Enhancing enzyme production yield using rational design

Margaux Haon[a,b], *Daniel J. Rigden[b], Andrew J. Carnell[a]*

[a]Department of Chemistry, University of Liverpool, Crown Street, Liverpool, L69 7ZD, UK
[b]Institute of Systems, Molecular and Integrative Biology, Crown Street, Liverpool, L69 7ZB, UK

A large amount of everyday pharmaceutical drugs present at least one amide bond in their structure and 16% of pharmaceutical reactions for drug production are amide bond synthesis[1]. The chemical synthesis of this bond is done starting with an amine and a carboxylic acid that needs to be activated prior to the reaction. This chemical reaction is easy and common but often uses toxic or hazardous reagents to activate the acid and thus generates a lot of waste. In contrast, enzymatic methods for amide bond production are becoming more attractive due to their requirement for milder reaction conditions. However, for biocatalysis to be competitive with chemistry, the enzymes used need to be highly active and easy to produce.

By studying various hydrolases, it appears that some of them have a promiscuous N-acyltransferase activity in bulk water. Hydrolysis and acyl transfer mechanisms are closely related and hydrolysis can be viewed as acyl transfer to water. Accordingly, competing hydrolysis often hampers efforts to use enzymes for acyltransferase reactions. This problem can sometimes be overcome by, for example, using excess acyl donor[2].

Our team discovered a promising β-lactamase able to form amide bonds in bulk water from an ester and amine. However, the production yield of the enzyme is too small to be used in large scale experiments, closer to industrial amide production. In this presentation, we will show how we enhanced enzyme solubility by rational enzyme design. For this aim, we developed a pipeline to: 1) generate several new enzyme sequences using proteinMPNN, 2) build the structure associated with these new sequences and 3) select *in silico* the best mutants using different scores such as mutation stability or aggregation propensity. Some of the new mutants show a better production yield while retaining the wild type activity. In the future, we also want to broaden the substrate scope of our enzyme and reduce the hydrolysis propensity by applying rational design methods.

References:
[1] Petchey, M. R.; Grogan, G. Enzyme-Catalysed Synthesis of Secondary and Tertiary Amides. *Adv. Synth. Catal.* 2019, ***361*** (17), 3895–3914. https://doi.org/10.1002/adsc.201900694.
[2] Müller, H.; Terholsen, H.; Godehard, S. P.; Badenhorst, C. P. S.; Bornscheuer, U. T. Recent Insights and Future Perspectives on Promiscuous Hydrolases/Acyltransferases. *ACS Catal.* 2021, ***11*** (24), 14906–14915. https://doi.org/10.1021/acscatal.1c04543.

Programme of Poster Presentations

1	De Novo Design of GALK1 Inhibitors in a Flexible Binding Pocket Asma Feriel Khoualdi, Newcastle University
2	Using tiered computational screening to discover small molecule inhibitors of the SARS-CoV-2 nsp3 protein Mac1 domain Shubhangi Kandwal, Trinity College Dublin
3	Using multiple receptor structures in generative models for de novo drug design Wei Dai, Queen Mary University of London
4	Software toolkits for in silico screening of polymer excipients used in small molecule formulation and drug delivery Hannah Turney, King's College London,
5	Is TC2559 an allosteric super-agonist at high sensitivity (α4)2(β2)3 nicotinic acetylcholine receptors? Franco Viscarra, University of Oxford
6	Druggability-enhanced structural chemogenomics approaches to navigate the dopamine GPCR-ligand interaction space Stylianos Iliadis, Queen Mary University of London
7	Developing and Applying a Theoretical Receptor Approach to Novel Boron-containing Autotaxin Inhibitors Mauno Lius, University of Manchester
8	macrostrain: Automated Calculation of Macrocycle Strain Aleksy Kwiatkowski, University of Oxford
9	Investigating origins of differential carbapenem hydrolysis by OXA-48 variants through multiscale simulations Daojiong Wang, University of Bristol
10	Capturing enzymes in action: molecular movies of indoleamine 2,3-dioxygenase (IDO1) by time-resolved serial X-ray crystallography and molecular modelling Abilasha Balakrishnan, University of Leicester
11	Using Markov state model and Deep Learning to explore the dynamic difference of Class A β-Lactamases Jing Gu, UCL
12	Machine Learning Methods for Quantitative Biophysics-Ensemble Integration: Applications to HDX-MS Alexander Hussain-Siddiqui, University of Oxford

Poster presenters will give their 2 minute "Lightning Talks" in the above order starting at 11:00 and will be available at the poster session to answer questions.

Poster 1

De Novo Design of GALK1 Inhibitors in a Flexible Binding Pocket

Asma Feriel Khoualdi[1]; *Matteo Degiacomi[2]; Wyatt Yue[3]; Daniel J. Cole[1]*

[1]School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, NE1 7RU, UK
[2]Department of Physics, Durham University, DH1 3LE, UK
[3]Newcastle University Medical School, Newcastle-upon-Tyne, NE2 4HH, UK

Galactosemia is a rare autosomal recessive disease caused by a defect in galactose metabolism. GALK1 is a kinase enzyme involved in galactose metabolism and whose inhibition can help alleviate galactosemia[1]. By making use of the GALK1's allosteric inhibitors which are crystalized[1], molecules with better binding properties can be designed, modified and optimized within the binding pocket using FEgrow tool[2]. Molecules can be designed interactively or automatically using substructures from databases, built and scored using FEgrow.

Pocket flexibility is addressed with the conformation prediction tool Molearn[3] which uses a convolutional neural network to learn from relatively short, example MD simulation trajectories to predict the different conformations the protein will have. Designing inhibitors that fit the predicted ensemble of pocket conformations means more potent inhibitors can be generated and tested against not only a static, rigid receptor structure as is currently done in docking, for example, but against a flexible form of the receptor target. The active learning feature of FEgrow, in particular, ensures efficiency in selecting potential inhibitors from the very large chemical space.

References:
[1] Mackinnon, S.R., et al., Fragment Screening Reveals Starting Points for Rational Design of Galactokinase 1 Inhibitors to Treat Classic Galactosemia. ACS Chemical Biology, 2021. **16**(4): p.586-595.
[2] Bieniek, M.K., et al., An open-source molecular builder and free energy preparation workflow. Communications Chemistry, 2022. 5(1): p. 136.
[3] S.C. Musson and M.T. Degiacomi (2023). Molearn: a Python package streamlining the design of generative models of biomolecular dynamics. Journal of Open Source Software, **8**(89), 5523

Poster 2

Using tiered computational screening to discover small molecule inhibitors of the SARS-CoV-2 nsp3 protein Mac1 domain

Shubhangi Kandwal[1,2,3], *Prof. Andrew Bowie[4] and Dr. Darren Fayne[2,3]*

[1]Molecular Design Group, School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, 152-160 Pearse St, Dublin 2, D02 R59, Ireland
[2]Molecular Design Group, School of Chemical Sciences, Dublin City University, Glasnevin, Dublin, D09 V209, Ireland
[3]DCU Life Sciences Institute, Dublin City University, Glasnevin, Dublin, D09 V209, Dublin, Ireland
[4]School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, 152-160 Pearse St, Dublin 2, D02 R59, Ireland

As new medications are used to treat COVID-19, many studies have reported that proteins such as spike, polymerase and proteases are prone to high levels of mutation that can create resistance to therapy over time [1]. Thus, it becomes necessary to, not only target other viral proteins such as the non-structural proteins (nsp's), but to also target the most conserved residues of these proteins. A synergistic combination of bioinformatics, computer-aided drug-design and *in-vitro* studies can feed into better understanding of SARS-CoV-2 (SC-2) and therefore help in the development of small molecule inhibitors against the nsp's [2]. As part of our initial anti-viral work, a pharmacophore study on nsp15 found a hit molecule (INS316) that made interactions with Ser293, Lys344 and Leu345 residues [3] which are highly conserved across SC-2.

Our group was selected to enter an international challenge organized by CACHE to find inhibitors for the Mac1 domain of SC-2 nsp3. Our MSA alignment results of ~1 million nsp3 sequences indicated that the Mac1 domain is a highly conserved pocket that can be targeted for developing promising SC-2 inhibitors. We used a tiered screening workflow which included the use of volume/shape information of the binding pockets (fastROCS), use of in-house pharmacophore generation software (MoPBS [4]/MOE) and performed docking in the binding pocket (FRED) to rank compounds for subsequent clustering and to identify hits that bind to these conserved pockets. The primary experimental validation results provided by CACHE found that two of our predicted hits show activity in HTRF and SPR assays.

References:
[1] M. Pachetti, B. Marini, F. Benedetti, F. Giudici, *et al., J. Transl. Med.,* (2020) ***18***(1):179.
[2] S. Kandwal and D. Fayne, *Virology*, (2023) **581**, 97-115.
[3] S. Kandwal and D. Fayne*, J. Biomol. Struct. Dyn.*, (2022) ***40***(3), 1316–1330.
[4] J. Braun, and D. Fayne, *J. Mol. Graphics Modell*, (2022) **115**, 108228.

Poster 3

Using multiple receptor structures in generative models for de novo drug design

Wei Dai,[a] *Arianna Fornili,[a] Noel O'Boyle,[b] Chris de Graaf [b]*

[a] School of Physical and Chemical Sciences, Queen Mary University of London, E1 4NS, United Kingdom
[b] Nxera Pharma, Steinmetz Building, Granta Park, Great Abington, Cambridge, CB21 6DG, United Kingdom

Artificial Intelligence (AI) has recently demonstrated considerable promise in computer-aided drug design, particularly through the use of generative models. This project focuses on developing and applying generative AI in structure-based drug design. Specifically, a new scoring approach based on multiple receptor structures (MultiReceptor), is being constructed, building on an existing *de novo* generative model by Morgan et al.,[1] which employs a recurrent neural network (RNN) with a reinforcement learning (RL) agent. During training, randomly generated SMILES strings from the RNN model are docked with user-provided receptor structures. The RL agent then refines the RNN model based on the docking scores, aiming to produce molecules with improved binding affinity.

MultiReceptor has been integrated within MolScore,[2] allowing the simultaneous use of multiple receptor structures during training. A rank-by-vote scoring function is applied, where generated molecules are evaluated based on docking results across all input receptor structures. Molecules are ranked and scored by the number of votes they receive, with the model rewarding those with higher votes, thereby encouraging the generation of molecules with favourable properties across multiple receptors.

The performance of MultiReceptor is currently being assessed on the dopamine D2 G protein-coupled receptor. Three receptor structures were selected to train single-receptor models, as well as a multi-receptor model incorporating all three structures. Molecules sampled from the single-receptor models were compared with those generated from the multi-receptor model using metrics including docking scores and interaction fingerprint similarities. Preliminary results show that molecules generated by the multi-receptor model are predicted to have more favorable binding affinity values compared to those from the single-receptor models, as well as a higher interaction similarity with co-crystallized ligands. This suggests that the inclusion of multiple receptor structures can enhance the quality of molecules produced by the *de novo* generative model.

References:
[1] M. Thomas, N. M. O'Boyle, A. Bender and C. De Graaf, *Journal of cheminformatics*, (2022) **14**, 68.
[2] M. Thomas, N. M. O'Boyle, A. Bender and C. De Graaf, *Journal of cheminformatics*, (2024) **16**, 64.

Poster 4

Software toolkits for *in silico* screening of polymer excipients used in small molecule formulation and drug delivery

Hannah Turney, *Micaela Matta*

Department of Chemistry, King's College London, Southwark, London

The use of polymers as excipients in small molecule pharmaceutical formulations is an established approach for the controlled delivery of drugs[1]. However, designing safe and effective formulations is resource-intensive and delays product delivery to the clinic, primarily due to the sensitivity of polymer substructure to delivery properties.

Molecular dynamics (MD) simulations can reveal information about the conformational changes, binding interactions, and dynamical properties of molecules. However, limitations remain in the routine application of MD simulations to polymer excipients, such as the difficulty in parameterizing larger polymers and ensuring the transferability of force fields across different polymer chemistries.

In this project, we create a robust and scalable building and parameterization workflow for polymer excipients. This research uses advanced molecular dynamics techniques from organizations such as the Open Force Field Consortium to create a robust *in silico* polymer parameterization methodology[2]. With our established workflow, we perform systematic molecular dynamics simulations of different polymer:drug systems to yield kinetic and mechanical parameters predicting excipient suitability for drug delivery. This enables fast, accurate, and reproducible profiling of polymers in the context of drug formulation design. We collaborate with experimental formulation development teams at Johnson&Johnson Innovative Medicine to validate our models and drive the design of our polymer:drug candidate systems.

By incorporating existing open-source software, and sharing any resulting tools developed, we aim to promote reproducibility and collaboration throughout the project.

References:

[1] H.K. Makadia and S.J. Siegel, *Polymers (Basel),* (2011), **3**, 1377-1397.
[2] L. Wang, P. Kumar Behara, M.W. Thompson, T. Gokey, Y. Wang, J.R. Wagner, D.J. Cole, M.K. Wilson, M.R. Shirts, D.L. Mobley, *The Journal of Physical Chemistry,* (2024) **128**, 7043-7067.

Poster 5

Is TC2559 an allosteric super-agonist at high sensitivity (α4)2(β2)3 nicotinic acetylcholine receptors?

Franco Viscarra, Teresa Minguez, *Isabel Bermudez-Diaz, Philip C. Biggin*.

Structural Bioinformatics and Computational Biochemistry, Department of Biochemistry, University of Oxford, South Parks Road, Oxford. UK. OX1 3QU

The most abundant nicotinic acetylcholine receptor (nAChR) in the brain is the α4β2 subtype. It is involved in drug and nicotine addiction and a range of severe central nervous system (CNS) disorders. Two alternate stoichiometries of this receptor exist, the high sensitivity (α4)2(β2)3 (HS) and the low sensitivity (α4)3(β2)2 (LS), which have different functional and pharmacological properties. In the HS stoichiometry, the ligand TC2559 exhibits a "super-agonist" profile, with a maximal response of ~3 fold higher than the maximal response of acetylcholine while being a partial agonist of poor efficacy in the LS receptor (~0.2 fold of acetylcholine maximal response) [1]. This low efficacy on the LS stoichiometry is explained by the inability of TC2559 to bind the α4-α4 interface present in this receptor [2], but its ability to act as a super-agonist in the HS receptor remains unexplained. Here, we investigate the mechanism of action of TC2559 on the α4β2 nAChR using a combination of molecular dynamics and electrophysiology. Exploratory simulations revealed a binding pattern among α4β2 agonists that correlated with their efficacies but TC2559 does not follow. Based on the existence of non-canonical binding sites for modulators, such as the benzodiazepines in the GABA receptors, morantel in the α3β2 nAChR, and NS9283 in the LS α4β2 nAChR, we hypothesised that TC2559 binds at an additional interface not present in the LS stoichiometry, namely the β2-β2 interface. Exploratory simulations of TC2559 in the β2-β2 interface revealed two key aspartates, present in loop C of the β2(+) and loop F of the β2(-) subunit, which appeared important for binding. Both aspartates were subsequently validated by site-directed mutagenesis, suggesting that the β2-β2 interface may play a role in the mechanism of action of TC2559 as an allosteric super-agonist on the HS (α4)2(β2)3 receptor.

References:

[1] Moroni M, Zwart R, Sher E, Cassels BK, Bermudez I, *Mol Pharmacol,* (2006),**70**(2),755-68.
[2] Mazzaferro S, Gasparri F, New K, Alcaino C, Faundez M, Iturriaga Vasquez P, Vijayan R, Biggin PC, Bermudez I, *J Biol Chem*, (2014), ***289****(31)*, 21795-806.

Poster 6

Druggability-enhanced structural chemogenomics approaches to navigate the dopamine GPCR-ligand interaction space

Stylianos Iliadis[1], Georges Chalhoub[2], *Arianna Fornili*[1], *Peter McCormick*[2], *Francesca Deflorian*[3], *Chris de Graaf* [3*]

1. School of Physical and Chemical Sciences, Queen Mary University of London, E1 4NS, United Kingdom
2. Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Biosciences Building, Liverpool, L69 7BE, United Kingdom
3. Nxera Pharma, Steinmetz Building, Granta Park, Great Abington, Cambridge, CB21 6DG, United Kingdom
*current address: Structure Therapeutics, 601 Gateway Blvd, Suite 900, South San Francisco, CA

G protein-coupled receptors (GPCRs) constitute important drug targets, offering opportunities for therapeutic interventions. 3D structure determination of GPCR-ligand complexes accelerates drug discovery by revealing ligand-binding modes. To bridge the gap between the increasing number of ligands and the limited experimentally known GPCR-ligand binding modes, structural chemogenomics methods have proven useful by enabling ligand binding mode elucidation[1].

In this work, we incorporate binding site druggability analysis (physicochemical features and explicit water network)[2] into structural chemogenomics methods to enhance binding mode predictions and implement it on dopamine GPCRs, a subfamily of the well-studied aminergic GPCR family.

From the available dopamine receptor structures, a subset was selected based on structure quality, redocking performance and/or binding site druggability criteria, considering separately orthosteric and allosteric binding sites. Additionally, dopamine GPCR ligands from bioactivity databases and manual data extraction were collected, curated, and annotated. The ligands were split per receptor subtype and orthosteric/allosteric binding site and docked at the binding sites of the selected structures. One binding pose per ligand was selected based on druggability insights (occupied lipophilic hotspots and displaced high-energy waters)[2] which were annotated, resulting in a dataset of druggability-annotated docked ligands and binding pockets.

To showcase the value of the method and dataset, a binding mode elucidation case was pursued for dopamine D1 allosteric ligands, for which structure-activity relationships (SAR) and mutation data were collected. The dataset docked poses were evaluated or refined against SAR, mutation data mapping, and systematic docking. Guided by the druggability defined pockets and human dopamine receptor orthologue sequences, mutation studies were designed to validate the binding hypotheses. Experimental work is still underway, with initial mutagenesis study results presented.

References:

[1] B. Briels, C. Graaf and A. Bender, *Structural Biology in Drug Discovery*, Wiley, (2020), 53–77.
[2] J. S. Mason, A. Bortolato, D. R. Weiss, F. Deflorian, B. Tehan and F. H. Marshall, *In Silico Pharmacol.*, (2013), **1**, 23.

Poster 7

Developing and Applying a Theoretical Receptor Approach to Novel Boron-containing Autotaxin Inhibitors

Mauno J. Lius[a], *Dr. Andrew G. Leach[a], Dr. Cristina Trujillo del valle[b], Dr. Jonathan M. Skelton[b]*

[a] Division of Pharmacy and Optometry, University of Manchester, Oxford Rd, Manchester M13 9PL.
[b] Department of Chemistry, University of Manchester, Manchester M13 9PL

The use of boron-containing compounds (BCCs) in the pharmaceutical industry has been steadily increasing in the past decade. The interest arises because of boron's Lewis acidity and tendency to form reversible covalent interactions with proteins. However, computational methods to further investigate the energetic consequences of such binding is scarce and made difficult by the lack of efficient methods to model drug-protein complexes at the quantum-mechanical level. The current work has employed the use of a 'theoretical receptor' method[1] that aims to provide the best compromise between speed and accuracy in determining the binding geometry and affinity of drug-like compounds in a target protein by systematically extracting a sub-region of the protein (typically an active site) whose size is more amenable to traditional quantum-mechanical methods. In doing so, more accurate computational methods can be employed to best model the interaction of BCCs and their protein targets.

The method has been applied to the autotaxin (ATX) enzyme which is implicated in numerous diseases including cancers, pulmonary fibrosis, and neuropathy. This target was chosen because a range of boronic acid inhibitors have been identified in the literature making it an ideal protein to test our approach. Using GFN2-xTB[2], accurate binding geometries and associated binding energies (ΔE) have been obtained for a range of inhibitors including many BCCs. The theoretical receptor method reveals that covalently-bound boronic acids have a better binding energy than their non-covalent equivalents. Our multilinear regression model (including AlogP) suggests that we can predict pIC_{50} values (N=37, R^2=0.612, RMSE=0.825) with good overall accuracy, especially for boronic acids.

Future work on the ATX theoretical receptor will include conformational analysis of the active site to gauge whether it may improve pIC_{50} predictions for outlier ligands in the dataset.

References:
[1] I. Lukac, H. Abdelhakim, A. R. Ward, S. A. St-Gallay, J. C. Madden, A. G. Leach, *A. Chem. Sci*, (2019) **10**, 2218-2227.
[2] C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.,* (2019) **15**, 1652–1671.

Poster 8

macrostrain: **Automated Calculation of Macrocycle Strain**

Aleksy Kwiatkowski[a], *Fernanda Duarte*[a]

[a]Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA.

Strain is a key concept in organic chemistry, used to describe how molecules deviate from an 'ideal' set of geometric parameters. For macrocycles—ring-shaped molecules with potential in drug discovery for targeting difficult, "undruggable" proteins—ring strain is particularly significant. This strain influences not only their synthetic accessibility but also their reactivity and functional properties [1, 2]. However, quantifying strain remains challenging, often involving extensive manual work.

Here, we present *macrostrain*, an open-source, user-friendly tool that automates the calculation of ring strain, with a focus on macrocyclic structures. This tool takes only a SMILES string of the molecule as an input. By representing the molecule as molecular graph, *macrostrain* automatically identifies the atoms involved in the ring and the bonds at which to break the ring to generate unstrained linear references. From this, homodesmotic reactions are generated and their reactions energies, corresponding to strain, are computed.

We demonstrate *macrostrain* on a diverse set of pharmaceutically relevant macrocycles, spanning a range of sizes and chemical compositions. Additionally, we evaluate how different conformer generation algorithms influence the computed strains. The tool's ease of use makes it accessible to non-experts, facilitating the calculation of, and therefore the application of, strain as a metric in the assessment of synthetic accessibility, reactivity, and other essential properties.

References:

[1] Villar, E. A., Beglov, D., Chennamadhavuni, S., Porco Jr, J. A., Kozakov, D., Vajda, S., & Whitty, A. *Nat. Chem. Biol.*, (2014), **10**, 723-731.

[2] Bogdan, A. R., Jerome, S. V., Houk, K. N., & James, K. (2012). *JACS*, ***134***(4), 2127-2138.

Poster 9

Investigating origins of differential carbapenem hydrolysis by OXA-48 variants through multiscale simulations

Daojiong Wang [a], *Adrian J. Mulholland [b], James Spencer [c], Marc W. van der Kamp [a]*

[a] School of Biochemistry, University of Bristol, Bristol, BS8 1TD
[b] School of Chemistry, University of Bristol, Bristol, BS8 1TS
[c] School of Cellular and Molecular Medicine, University of Bristol, BS8 1TD

β-lactam antibiotics are the most widely used treatment of bacterial infections and have saved numerous human lives. Bacteria can produce β-lactamases (BLs), which are enzymes that first acylate β-lactams, forming an acyl-enzyme complex, and then release the hydrolysed antibiotic through deacylation, leading to inactivation of the drug.

The common BL OXA-48 has a strong hydrolysis activity towards the popular antibiotic imipenem. Both OXA-163 and -405 are very similar to OXA-48, but feature a four-residue deletion β5-β6 loop, which leads to a reduced imipenem hydrolysis activity (as well as introducing ceftazidime hydrolysis). Notably, a newly discovered OXA-48 variant, OXA-517, has just a two-residue deletion in the same loop and partially rescues imipenem hydrolysis activity compared to OXA-163 and -405, whilst maintaining ceftazidime actvity.[1] The molecular basis of the different imipenem hydrolysis activities and the role of β5-β6 loop for carbapenem hydrolysis remain unclear, and are thus investigated here.

Reaction simulations (quantum mechanics/molecular mechanics, QM/MM) of the deacylation step showed that the reaction is most efficient when the carboxylated Lys73 (acting as the base) is least hydrated and the deacylating water acts as hydrogen donor to imipenem 6α-hydroxyethyl group. Molecular dynamics simulations (MM) of acyl-enzyme complexes shows that the structure of OXA-48 and -517 is more stable, with the biggest structural difference between variants being around the β5-β6 loop, especially near Thr213. Mutations in the β5-β6 loop change the H-bond network around Thr213, causing a different preference of the deacylating water H-bonding patterns. The H-bonding pattern that leads to a lower energy barrier is easier for OXA-48 and -517 to achieve than for OXA-163 and -405. Our work captures a subtle structural change and identifies the molecular basis for changes in imipenem hydrolysis activity between OXA-48 variants, revealing a deeper insight into the dynamics and reaction mechanisms related to antibiotic resistance.

References:
[1] L. Dabos, J. E. Raczynska, P. Bogaerts, A. Zavala, D. Girlich, R. A. Bonnin, L. Dortet, A. Peyrat, P. Retailleau, B. I. Iorga, M. Jaskolski, Y. Glupczynski and T. Naas, *Antimicrob Agents Chemother*, (2023), **67**, e0109522.

Poster 10

Capturing enzymes in action: molecular movies of indoleamine 2,3-dioxygenase (IDO1) by time-resolved serial X-ray crystallography and molecular modelling

Abilasha Balakrishnan, *Dr Hanna Kwon*, *Dr Ralf Schmid*

Molecular and cell biology, Leicester Institute of Structural and Chemical Biology, Henry Wellcome Building – University of Leicester, Lancaster road, Leicester, LE1 7HB

Human Indoleamine 2,3- dioxygenase 1 (hIDO1) is an intracellular heme enzyme having a pivotal role at the first-rate limiting step of the metabolism of L-tryptophan (*L*-Trp) to *N*-formyl-kynurenine (NFK) in the kynurenine pathway. It is widely expressed in both immune and non-immune tissue and is integral to the host's innate immune defence and control. By depleting *L*-Trp from the local tissue microenvironment, hIDO1 promotes the formation of kynurenine metabolites. Due to its key immunomodulatory function, hIDO1 has been implicated in various diseases, including chronic inflammatory syndrome, inflammatory bowel diseases, diverse cancers, neurodegeneration, HIV, and depression.

Research into hIDO1's fundamental biochemistry, encompassing the catalytic reaction mechanism, substrate interactions, identification of enzyme inhibitors and the regulatory mechanisms governing hIDO1 expression, has become a prominent focus. Despite numerous crystal structures of hIDO1, the mechanism of *L*-Trp turnover remains elusive. Time-resolved structural studies at the XFEL facility aim to shed light on this mechanism. While several inhibitors for hIDO1 have entered clinical trials, understanding the mechanism of apo inhibitors remains incomplete. Molecular dynamic simulations of available X-ray structures of hIDO1 in the apo form with heme-displacing inhibitors bound, alongside inactive ferric hIDO1, will help characterize the pathways and energetics involved in binding and unbinding of heme and heme-displacing inhibitors to apo hIDO1.

Poster 11

Using Markov state model and Deep Learning to explore the dynamic difference of Class A β-Lactamases

Jing Gu, *Shozeb Haider*

Pharmaceutical and Biological Chemistry, University College London School of Pharmacy, London, United Kingdom

The β-lactams maintains to hold a prominent position in the armament of antibacterial drugs. However, antimicrobial resistance is a growing threat to global public health. Some bacteria can now produce enzymes called β-lactamases, which destroy some of the most commonly used β-lactam antibiotics such as penicillin. Class A β-Lactamases are known to be the most common enzymes to confer a high level of resistance to β-lactam antibiotics.[1] Understanding the dynamic difference and active site similarities of these enzymes is important to reveal the resistance mechanisms and inspire future drug development.

PER-1 and PER-2, two Class A β-Lactamases that exhibit extremely close sequence identities, were first compared in detail. Markov State Models were built with the performed molecular dynamics simulations using hydrophobic nodes, active site residues and all other existed different residues as features. BindSiteS-CNN model was further applied to observe the local similarities of active site features between PER-1, PER-2 and four other representative class A β-Lactamases (KPC-2, SME-1, TEM-1, and SHV-1).[1,2]

The difference between the dynamic of the α9 motif of PER-1 and PER-2 was highlighted during the comparison of their RMSF plots. The Ω-loop of the two enzymes were observed to be different with the distances between some key residue pairs calculated. The local similarities of active site features between Non-Carbapenemases (PER-1, PER-2, TEM-1 and SHV-1) were revealed with the BindSiteS-CNN based Deep learning analysis. While Non-Carbapenemases clustered together with each other, KPC-2 and SME-1 were distinct as two independent clusters, which suggests the connection between the substrate selectivity and the active site features.

The application of Markov state model and Deep learning (BindSiteS-CNN) has provided a better understanding of Class A β-Lactamases. Recognizing the differences and revealing the local similarities will be helpful in understanding antibiotic resistance and advancing drug development.

References:

[1] E. Olehnovics, J. Yin, A. Perez, G.D. Fabritiis, R.A. Bonomo, D. Bhowmik and S. Haider, *Front Microbiol*, (2021) **12**, 720991.
[2] O.B. Scott, J. Gu, and A.W.E. Chan, *J. Chem. Inf. Model*., (2022) **62**, 5383-5396.

Poster 12

Machine Learning Methods for Quantitative Biophysics-Ensemble Integration: Applications to HDX-MS

Alexander Hussain-Siddiqui[a], *Oliver Crook[a], Rachael Skyner[b], Maria Musgaard[b], Srinath Krishnamurthy[b] and Charlotte Deane[a]*

[a] Department of Statistics, 24-29 St Giles', Oxford OX1 3LB
[b] *OMass Therapeutics,* ARC Oxford, OX4 2GX

Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) is an experimental technique that provides rich insights into protein behaviour, measuring conformational dynamics in solution. While offering an unparalleled view into protein dynamics, current approaches using HDX-MS investigations remain qualitative. Like many biophysical experiments, challenges revolve around solving the inverse problem of transforming biophysical data into a structure or ensemble of which there is no tractable link. Instead, we must propose structural solutions using a forward predictive model in an attempt to find an optimal fit. These empirical forward models are inaccurate and current ensemble-optimisation approaches across biophysics do not robustly provide an approach to quantify uncertainty in the optimisation from a rigorous basis.

In this work we explore further the integration of machine learning (ML) and biophysics by applying statistical methods commonly used in ML-optimisation. We propose the use of careful data splitting and replicates to generate validation datasets and, crucially, multiple perspectives. Combining this with appropriately defined metrics suitable for ensembles from both thermodynamics and ML, we demonstrate how this framework can be used to quantify certainty and interpret the optimisation process. We do this over a wide case studies and experiments in the context of Maximum Entropy (MaxEnt) optimisation with HDX-MS data using HDXer. We find that this approach enables intuitive interpretation of ensembles without resorting to qualitative structure comparison, avoiding human bias that is widespread in structural methods.

We suggest that this framework is suitable for any forward model, not just HDX-MS. We believe this will facilitate further integration of biophysics with bio-molecular simulations across all experimental domains.

Useful Information

Travel to Lady Margaret Hall

https://www.lmh.ox.ac.uk/sites/default/files/documents/2017-01/LMH%20Travel%20Information.pdf

YMF will be located in the Simpkins Lee Theatre and Monson Room