MGMS and RSC MMG
Young Modellers' Forum 2026
PROGRAMME & ABSTRACTS

Programme of Oral Presentations

09:15-09:50	Registration and Coffee
09:50-10:00	Introduction/Welcome
10:00-10:20	Wojtek Treyde, University of Oxford **Computational Design of γ-Lactam Serine-β-Lactamase Inhibitors**
10:20-10:40	Sophia Ber, Queen Mary University of London **Insights into the Gating Mechanism of Two-Pore Domain Potassium Channel**
10:40-11:00	Bowen Deng, University of Bristol **Chemical Intuition Distiller: Turning Human Chemical Intuition into Collective Variables**
11:00-11:30	Flash Poster presentations
11:30-11:50	Break
11:50-12:10	Mauno Lius, University of Manchester **Development of a Theoretical Receptor to Model Covalent Inhibition of Autotaxin by Boron-containing Compounds**
12:10-12:30	Brian Wiley, University of Bath **Deep Learning to Enhance Probing Dioxygen Pathways in Multiple Dioxygenases**
12:30-12:50	Victoria Byelova, University of Leeds **Mesoscale Heterogeneity in Protein Hydrogels Induced by Dynamic Unfolding and Post-Gelation Rearrangements**
12:50-14:10	Lunch and Posters
14:10-14:30	Miles Pemberton, University of Bath **Learning Wheland Intermediate Energies for Mechanistically Interpretable Electrophilic Aromatic Substitution Regioselectivity Prediction**
14:30-14:50	Chenfeng Zhang, University of Edinburgh **Looks Open, Stays Dry: Simulations Reassign the Open State of TRPP Ion Channels**
14:50-15:10	Magdalena Kazmierczak, University of Bristol **Machine Learning Potentials for Accurate and Efficient Multiscale Enzymatic Diels-Alder Reaction Modelling**
15:10-15:30	Break
15:30-15:50	Ronald Cvek, University of Oxford / Diamond Light Source **MoCASSIn: 3D Similarity Scoring and Placement for Evaluating Generative Compound Designs Against Experimental Fragment Poses**
15:50-16:10	Amy Gill, University of Warwick **Uncovering the Molecular Mechanism of Undecaprenyl Phosphate Recycling Within Bacterial Membranes by DedA Family Proteins**
16:10-16:30	Stephen Harman, University of Nottingham **Not so Classical: Non-Classicality as the Rule in CHX8+ Chemical Space**
16:30-16:45	Deliberation and Prizes
16:45	Close

Talk 1

Computational Design of γ-Lactam Serine-β-Lactamase Inhibitors

Wojtek Treyde, Dylan John, Christine Arndt, Jiarui Zong, Liam A. Wilson, *Christopher J. Schofield, Fernanda Duarte*

Chemistry Research Laboratory, University of Oxford, Oxford, OX1 3TA, UK.

The most widely prescribed class of antibiotics are β-lactams, which inhibit penicillin- binding proteins (PBPs) essential to bacterial cell wall synthesis. Resistance to β-lactam antibiotics is increasingly driven by serine β-lactamases (SBLs), enzymes that hydrolyse the β-lactam ring. While development of β-lactamase inhibitors has extended the use of β-lactam antibiotics, there remains a need for novel SBL inhibitors, particularly against class C SBLs such as AmpC, for which there are fewer clinically approved inhibitors.

Here, we present a multistage computational approach for identifying covalent AmpC inhibitors, combining machine learning (ML), docking and molecular dynamics (MD) simulations. We trained an ML model for predicting AmpC activity using 12 quantum-mechanically derived molecular descriptors capturing intrinsic reactivity on internal and literature inhibition data. We designed and virtually screened a library of γ-lactams, and ran docking and MD simulations for around 250 compounds with predicted activity against AmpC. Top candidates were synthesised and tested experimentally, yielding the first monocyclic γ-lactam SBL inhibitor.

This work demonstrates that combining ligand-based ML with structure-based modelling enables efficient prioritisation of candidates for synthesis and experimental validation. The approach offers a generalisable framework for the development of novel inhibitors targeting SBLs, PBPs and serine proteases more broadly.

Talk 2

Insights into the Gating Mechanism of Two-Pore Domain Potassium Channel

Sophia Ber, *Wojciech Kopec*

Department of Chemistry, Queen Mary University of London, Mile End Road London E1 4NS, UK.

Two-pore domain potassium (K_{2P}) channels conduct background K^+ currents that stabilize the membrane potential and modulate cellular excitability. While all K_{2P} channels exhibit selectivity filter (SF)–based gating, some, including THIK-1, also possess an additional lower gate. Widely expressed across cell types, K_{2P} channels contribute to diverse physiological and pathological processes. Despite being well-characterized and recognized drug targets, their gating mechanisms remain poorly understood, limiting the development of selective therapeutics.

To address this knowledge gap, we investigated the mechanisms of THIK-1 activation and inhibition on the atomic-level using molecular dynamics simulations, employing the recently developed force field that incorporates electronic polarization via charge scaling[1]. This approach allowed us to quantify key gating events at both the SF and lower gates. Our findings advance the fundamental understanding of K_{2P} gating and provide a foundation for novel drug design strategies.

THIK-1 gating is controlled by the interplay between a leaky-type lower gate and the SF. A specific orientation of the two tyrosine residues forming the lower gate enables the formation of a Tyr–K^+–Tyr complex that permits ion permeation. SF inactivation occurs through disruption of the S1 or S2 ion-binding sites following backbone carbonyl flips of F113 and G112, respectively. Furthermore, our results suggest that the M2 helix contributes to coupling between the lower gate and SF conductivity.

Taken together, our findings advance our understanding of K_{2P} function and highlight structural features in THIK-1 as potential therapeutic targets.

References:

[1] C. Hui, R. de Vries, W. Kopec and B. L. de Groot, *Proc. Natl. Acad. Sci. U.S.A.*, (2025), **122**, e2423866122.

Talk 3

Chemical Intuition Distiller: Turning Human Chemical Intuition into Collective Variables

Bowen Deng[a,b], Kun Yang[b], Haowen Yu[c], Marko Hanzevacki[b], Elliot W. Chan[b], Kirill Zinovjev[d], Sofia Oliveira[b], M.Carmen Galan[a], Adrian Mulholland[b]

[a] School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[b] Centre for Computational Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[c] School of Psychology & Neuroscience, University of Glasgow, 62 Hillhead Street, Glasgow, G12 8QB, UK

[d] Departamento de Química Física, Universidad de Valencia, Burjassot, Spain, 46100

Machine-learned collective variables (MLCVs)[1] capture complex molecular processes in a few data-driven coordinates, facilitating enhanced sampling of rare transitions and free-energy reconstruction. However, their development faces a "chicken-and-egg" problem: learning effective CVs requires training data capturing key transitions, yet generating such data often depends on predefined CVs. Chemical intuition suggests molecular pathways, but translating it into coordinates suitable for enhanced sampling remains challenging. Interactive molecular dynamics in virtual reality (iMD-VR)[2] allows users to guide molecular transformations without predefined CVs. iMD-VR trajectories encode pathway information but are sparse, human-guided nonequilibrium demonstrations, insufficient alone to identify metastable states or transition-state ensembles, or reconstruct free-energy landscapes.

We developed Chemical Intuition Distiller, a teacher-student framework for learning CVs from iMD-VR demonstrations. An encoder-only Transformer teacher processes local basic-CV sequences and integrates reference representations from other demonstrations through cross-attention guided by dynamic time warping of frozen Uni-Mol structural embeddings. Teacher progress labels, hidden representations, cross-trajectory constraints and path-distance weights supervise an eight-member TabM student with a shared multilayer perceptron backbone. The student predicts a differentiable progress coordinate from instantaneous basic CVs alone, distilling the teacher's structural and temporal knowledge into a compact model for simulation. The learned progress coordinate and an analytic path-deviation coordinate enable enhanced sampling and subsequent free-energy reconstruction. Using the same basic CVs, the framework supports sampling along distinct folding pathways guided by different iMD-VR demonstrations.

Initial studies of sugar binding to a glucose receptor yielded a binding free-energy close to experiment. As a second benchmark, we applied the pipeline to Diels-Alderase-catalysed reactions, testing its ability to learn effective CVs for free-energy simulations of chemical reactions. Other applications include complex macromolecular folding such as G-quadruplex folding. These applications explore how expert chemical intuition can initiate quantitative molecular investigations while reducing the need to manually design and repeatedly test pathway-specific reaction coordinates.

References:

[1] Trizio E, et al. Advanced simulations with PLUMED: OPES and machine learning collective variables. arXiv preprint arXiv:2410.18019, 2024.

[2] O'Connor M, et al. Sampling molecular conformations and dynamics in a multiuser virtual reality framework. Science Advances, 2018, 4(6): eaat2731.

Talk 4

Development of a Theoretical Receptor to Model Covalent Inhibition of Autotaxin by Boron-containing Compounds

Mauno J. Lius[a], *Dr. Andrew G. Leach[a], Dr. Cristina Trujillo[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', a QM-cluster method[1] which systematically reduces the protein down to its binding site using proximity from bound co-crystallised ligand(s) to make it more amenable to quantum-mechanical methods in a high-throughput manner.

The method has been applied to the autotaxin (ATX) enzyme which is implicated in numerous diseases including cancers, pulmonary fibrosis, and neuropathy. This target has been documented to have boronic acid inhibitors in the literature, making it an ideal protein to probe boron-mediated covalent inhibition. Using GFN2-xTB[2] and DFT, accurate binding geometries and subsequent binding energies (ΔE) have been obtained for inhibitors across a wide range of potencies with seamless incorporation of covalent inhibition by BCCs. By correlating DE and computed AlogP against experimental measurements, a multilinear regression model has been generated to predict relative drug potencies.

Our regression model shows good predictions (N=55, R^2=0.69, RMSE=0.78) for both BCCs and other ligands while showing similar robustness to novel unseen BCC inhibitors and confirms the energetic preference for covalent inhibition by BCCs. Moreover, the model has enabled a consideration of chemically feasible alternative binding poses for a subset of ligands. This is most evident with PF-8380, an inhibitor with the highest literature-reported potency, where an alternative binding mode is supported by a combination of computed energies and re-interpretation of the deposited electron density.

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.

Talk 5

Deep Learning to Enhance Probing Dioxygen Pathways in Multiple Dioxygenases

Brian Wiley[a], *Georgios Exarchakis[b]*, *Carmen Domene[a]*

[a]Department of Chemistry University of Bath, Bath BA2 7AX, United Kingdom

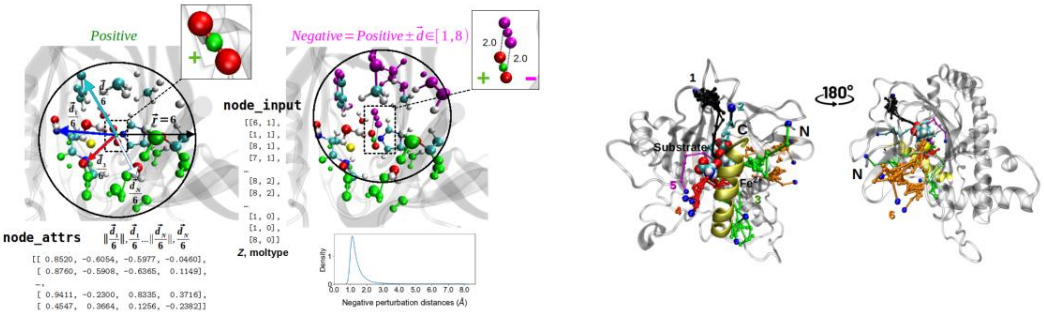
[b]Department of Computer Science University of Bath, Bath BA2 7AX, United Kingdom

Across all kingdoms of life, oxygen-dependent enzymes play fundamental roles in metabolism, signaling, and environmental adaptation. Early examples include heme-containing proteins such as hemoglobin and myoglobin, followed by oxygenases including TDO, IDO1/2, and the 2-oxoglutarate-dependent PHD enzymes that regulate the cellular response to hypoxia. These enzymes are important in immunity, cancer, and disease, while bacterial and fungal dioxygenases contribute to the biosynthesis and modification of diverse antibiotic scaffolds. Despite their sequence and functional diversity, many dioxygenases face a common mechanistic challenge: molecular oxygen must reach catalytic centers buried within the protein structure. Increasing evidence suggests that O_2 travels through transient pathways formed by interconnected hydrophobic cavities, which dynamically open and close through protein fluctuations. Molecular dynamics simulations can capture these processes, although identifying transient O_2-accessible states across long trajectories remains computationally demanding. Here, we investigate whether machine learning (ML), specifically SE(3)-equivariant graph neural networks, can learn the local structural and chemical environment associated with O_2 accessibility from MD simulations. By encoding interactions between protein atoms, solvent, and O_2 within a geometrically equivariant representation, we aim to identify features associated with transient gas-accessible cavities and tunnels. Ultimately, this approach seeks to use information learned from longer MD trajectories to characterise O_2 accessibility in shorter simulations of structurally related dioxygenases, providing a computational framework for probing dynamic gas-transport pathways at atomic resolution. We have discovered that, in enzymes that are perhaps crystallized in a more in-active state conformation, we can achieve determination of the same six to seven pathways per enzyme in two days where on average it could take five to six days of running long-timescale MD on high performance computers (HPCs) requiring more time and disk space.

References:

[1] Rabe, P.; Walla, C. C.; Goodyear, N. K.; Welsh, J.; Southwart, R.; Clifton, I.; Linyard, J. D. S.; Tumber, A.; Claridge, T. D. W.; Myers, W. K.; Schofield, C. J. Spectroscopic Studies Reveal Details of Substrate-Induced Conformational Changes Distant from the Active Site in Isopenicillin N Synthase. J. Biol. Chem. **2022**, 298 (9), 102249. https://doi.org/10.1016/j.jbc.2022.102249

[2] Sato, K.; Oide, M.; Nakasako, M. Prediction of Hydrophilic and Hydrophobic Hydration Structure of Protein by Neural Network Optimized Using Experimental Data. *Sci. Rep.* **2023**, *13* (1), 2183. https://doi.org/10.1038/s41598-023-29442-x



Talk 6

Mesoscale heterogeneity in protein hydrogels induced by dynamic unfolding and post-gelation rearrangements

Victoria Byelova[a], *Lorna Dougan[a], David Head[b]*

[a]School of Physics and Astronomy, University of Leeds, Woodhouse Lane, LS2 9JT

[b]School of Computer Science, University of Leeds, Woodhouse Lane, LS2 9JT

Protein gels are highly versatile systems that notably have multi-lengthscale behaviour and can be rationally designed. Through selection of their constituent protein building blocks they can be adapted to wound healing, tissue engineering and biosensing. There is a complex interplay between the behaviour of the individual building block and the bulk gel: when a protein in the gel undergoes conformational change, the bulk gel's structural and mechanical properties change. The converse is also true: bulk deformation can induce protein unfolding and presents a potentially complex mechanical response.

Though in-vitro experiments can map out individual protein and bulk gel characteristics well, the generic mesoscale formation and behaviour of a gel is less understood[1]. In-silico experiments have previously studied the force-lability of a few proteins or the network formation of a colloidal protein gel, but thus far there is no predictive, scalable mesoscale simulation that includes dynamic protein unfolding to our knowledge. We present a dynamic trimer model to represent protein unfolding and incorporate this into network formation simulations of over 10^4 proteins – this simple coarse-grained model bridges both scalability and dynamic unfolding in protein gel simulations. By varying energetic barrier for single-protein unfolding, we witness the formation of heterogeneous protein networks with controllable unfolding. Real-space analysis allows us to localise precisely where unfolding is taking place and what processes affect network rearrangement. This gives us unprecedented access to the physical mechanisms underlying protein hydrogel response and highlights the relevant parameters to control to produce bespoke hydrogels for applications.

References:

[1] A. Boroumand, M. D.G. Hughes, S. Cussons, N. Mahmoudi, D. A. Head, S. Peyman, A. I.I. Tyler, L. Dougan, *JOCAIS,* (2025), **699**, 138149.

Talk 7

Learning Wheland Intermediate Energies for Mechanistically Interpretable Electrophilic Aromatic Substitution Regioselectivity Prediction

Miles J. Pemberton,[ab] Samuel G. Espley,[a] Carlo Alberto Gaggioli,[c] Elliot H. E. Farrar,[c] and Matthew N. Grayson[a]

[a] Department of Chemistry, Faculty of Science, University of Bath, Bath, UK.

[b] Department of Computer Science, Faculty of Science, University of Bath, UK.

[c] Chemical Development, Process Technology & Development, Operations, AstraZeneca, Macclesfield, UK.

Electrophilic aromatic substitution (EAS) remains one of the most important classes of transformations in synthetic organic chemistry, yet predicting regioselectivity can be challenging when multiple substituents exert competing directing effects. Existing machine learning (ML) approaches can achieve high predictive accuracy but typically provide only qualitative rankings of reactive sites with limited mechanistic rationale. Quantum chemical methods provide a more interpretable alternative, particularly through analysis of Wheland intermediate stabilities,[1] but can require substantially greater computational resources and turnaround times compared to ML. Consequently, no practical approach currently combines the mechanistic rigour of quantum chemistry with the speed and scalability of ML.

Here, we present an ML framework that learns density functional theory (DFT)-derived Wheland intermediate energies and uses them to predict EAS regioselectivity. Using a dataset of 5,159 DFT-calculated EAS bromination intermediates generated with AstraZeneca's RegioEAS workflow, we evaluate classical ML methods, message-passing neural networks, and SE(3)-equivariant architectures. Rather than directly predicting the major reaction site, the models learn the underlying energetics governing regioselectivity, providing an explicit and chemically interpretable rationale for each prediction. The best-performing model achieves sub-kcal mol^{-1} accuracy relative to DFT-optimised intermediates and correctly identifies the major product-forming site in more than 95% of cases, with successful validation on experimentally relevant case studies spanning diverse functional motifs. Transfer learning further enables rapid extension to alternative electrophiles, achieving strong predictive performance with fewer than 200 additional training examples. Importantly, ensemble-based uncertainty quantification identifies reactions with competing regioisomeric outcomes and flags cases where ML predictions are insufficiently confident, enabling targeted escalation to higher-level DFT calculations.

Overall, this work demonstrates how physically informed ML can produce DFT-level regioselectivity predictions at a fraction of the computational cost while providing quantitative insight that is directly useful for synthetic chemists.

References:

[1] M. Liljenberg, T. Brinck, B. Herschend, T. Rein, G. Rockwell, M. Svensson, *J. Org. Chem.*, (2010), **75**, 4696-4705.

Talk 8

Looks Open, Stays Dry: Simulations Reassign the Open State of TRPP Ion Channels

Chenfeng Zhang[a], *Wojciech Kopec*[b], *Julien Michel*[a]

[a]EaStChem School of Chemistry, University of Edinburgh, Edinburgh, EH9 3FJ, UK.

[b]Department of Chemistry, Queen Mary University of London, London, E1 4NS, UK.

Autosomal dominant polycystic kidney disease (ADPKD) affects approximately 1 in 1000 people worldwide and can eventually lead to kidney failure, with no cure currently available.[1] Dysfunction of the PKD2 (TRPP1) ion channel is one of the main causes of the disease. Wild-type PKD2 structures show a π-helix bulge in the S6 pore-lining helix, whereas an α-helical S6 conformation was observed in the gain-of-function mutant F604P-PKD2. Based on pore radii from cryo-EM structures and experimental functional observations, the α-conformation was inferred to represent the channel open state, leading to the proposed π-to-α activation mechanism.[2] However, whether the α-conformation can support ion conduction has never been directly established.

Here, the functional capabilities of the α- and π-conformations of PKD2 were investigated using multiple molecular dynamics (MD) simulation techniques. The α-conformation observed in the cryo-EM structure was first shown to be unable to maintain the pore hydration required for function. Multiple pore-opening scans by steered MD then showed that the α-conformation could only maintain pore hydration when the lower gate was widened beyond the range observed in mammalian TRP channel structures. In contrast, the π-conformation was found to support pore hydration and ion conduction following a subtle lower-gate opening. Computational electrophysiology simulations with the electronic continuum correction (ECC) reproduced experimentally observed conductance, rectification, and ion selectivity in the open π-conformation. Further simulations of PKD2L1 (TRPP2) revealed the same conformation–function relationship.

Together, these findings challenge the assignment of the α-conformation as the conductive state, support the π-conformation as a functionally relevant state of TRPP channels, and suggest an α-to-π activation mechanism opposite to the previously proposed π-to-α mechanism. By revising the current model of PKD2 activation, this study provides a new framework for understanding PKD2 gating and identifying functionally relevant conformations for PKD2-targeted drug discovery in ADPKD.

References:

[1] E. Gall, A. Alam and R. D. Perrone, *Lancet*, (2019), **393**, 919–935,

[2] W. Zheng, X. Yang, R. Hu, R. Cai, L. Hofmann, Z. Wang, Q. Hu, X. Liu, D. Bulkley, Y. Yu, J. Tang, V. Flockerzi, Y. Cao, E. Cao and X. Z. Chen, *Nat. Commun.*, (2018), **9**, 2302.

Talk 9

Machine learning potentials for accurate and efficient multiscale enzymatic Diels – Alder reaction modelling.

Magdalena M. Kazmierczak[a], Elliot Chan[b], Chris Willis[b], *Marc van der Kamp*[a]

School of Biochemistry and Biological Sciences, University of Bristol, Bristol, BS8 1TD, UK.

School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom.

To perform efficient computational screening of reactions catalysed by enzymes, an efficient multiscale potential needs to be selected. Usually, this involves a trade-off between accuracy and efficiency: either substantial sampling with semi-empirical methods or limited sampling with more accurate DFT is applied in QM/MM. Here, we propose the use of machine learning potentials (MLPs) [1] applied together with electrostatic machine learning embedding (EMLE) [2] as an alternative to the traditional QM/MM calculations. This ML(EMLE)/MM approach has already been proven to reduce the computational cost to below the level of semi-empirical QM/MM and reproduce DFT-level QM/MM free energy profiles [2]. Initial benchmarking using ML(EMLE)/MM indicated that nanoseconds of simulations per day could be performed with the accuracy in energy and structures at the DFT level. The additional EMLE layer is required to accurately capture the influence of the enzyme electric field on the reaction.

After the initial research, which delves into broadly understood reactivity of three Diels–Alderases of catalytic interest (AbyU, Cyc15b, Cyc15b/S28K/F140Y), an in-depth understanding of the stereoselectivity of these three enzymes was gained. We currently focus our efforts on benchmarking and verifying efficient algorithms for MACE MLP training. A variety of experimentally verified molecules were considered for testing the applicability of MLPs for enzymatic Diels–Alder reactions, including enantiomers and intermolecular substrates, and their ML(EMLE)/MM US reproduced DFT QM/MM US with substantial accuracy. Currently, the initial Diels–Alder reaction to obtain training conformations for four, previously considered, stereoisomers was performed in water to minimise any environmental influences and generalise the training algorithm towards non-natural Diels–Alder products. A variety of methods were used for the structure selection process, and the MACE training itself took into account the quality of the produced machine learning potential as well as computational resources and time. The goal is a reproduction of DFT-level calculations of previously tested Diels–Alderases, with the ultimate aim of performing scanning of a library of putative Diels–Alderases and predicting their reactivity and stereoselectivity.

References:

[1] T. A. Young et al., *Chem. Chem. Phys.,* (2022), **24**, 20820 — 20827

[2] K. Zinovjev et al., *J. Chem. Theory Comput.*, (2024), **20**, 11, 4514 — 4522

[3] M. J. Byrne et al., *J Am Chem Soc*, (2016), **138**, 6095 — 6098

[4] L. Maschio et al., *Chem. Sci.*, (2025)

Talk 10

MoCASSIn: 3D similarity scoring and placement for evaluating generative compound designs against experimental fragment poses

Ronald Cvek[a,b], *Max Winokan*[c], *Carlos Perez*[d], *Lewis Vidler*[c], *Frank von Delft*[a,b]

[a]*Centre for Medicines Discovery, Nuffield Department of Medicine, NDM Research Building, Old Road Campus, Oxford, OX3 7FZ*

[b]*Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, Oxfordshire, UK*

[c]*Eli Lilly and Company, 8 Arlington Square West, Downshire Way, Bracknell, Berkshire RG12 1PU, UK*

[d]*Lilly Research Laboratories, Alcobendas, Spain*

Fragment-based drug discovery (FBDD) yields rich structural hypotheses that encode how small molecules engage a target, yet this signal is often lost when downstream algorithmic and generative design takes over. A key reason is the lack of robust tools to score and place proposed compounds against experimentally observed binding poses[1]. We present MoCASSIn, a framework that scores and places compound designs by 3D similarity against one or many observed crystallographic poses.

MoCASSIn evaluates the joint overlap of 3D shape and chemical features ("colour"), using a similarity measure selected to suit the comparison at hand. Beyond scoring, it performs placement: superposing each candidate to its references. As references are generally experimentally determined structures, the assessment is anchored in observed geometry rather than an estimated interaction energy, and is inherently directional, quantifying a candidate against references rather than treating the pair symmetrically. Evaluated against a pooled fragment screen, this yields answers to concrete design questions: whether a compound recapitulates its reference fragments, which fragments most informed it, and whether the elaboration is a productive one. The framework is efficient, parallelisable, and provides an interactive, interrogable Jupyter viewer.

We validate MoCASSIn on three fronts. First, that its placement recovers known crystallographic poses. Second, that the poses it produces carry favourable relative energy ("strain") comparable to published thresholds for protein-ligand complexes[2]. Third, on a focused set of real crystallographic fragment-to-lead elaborations, observed conformations are reproduced. Together these show that placement anchored to experimental structures provides a reliable, interpretable filter to triage large numbers of generative designs into ordered candidates.

References:

1. M. Thomas, R. Astudillo, Y. S. Raouf, C. Navarro, G. De Fabritiis and A. Bender, *Digital Discovery*, 2026, **5**, 3459–3473.
2. E. R. S. Wallace, N. C. Frey and J. A. Rackers, *J. Chem. Inf. Model.*, 2025, **65**, 6613–6620.

Talk 11

Uncovering the Molecular Mechanism of Undecaprenyl Phosphate Recycling within Bacterial Membranes by DedA Family Proteins

Amy Gill[a], Robin A. Corey[b], *Phillip J. Stansfeld*[a,c]

[a] School of Life Sciences, University of Warwick, Coventry, UK, CV4 7AL

[b] School of Physiology, Pharmacology, and Neuroscience, University of Bristol, UK, BS8 1TD

[c] Department of Chemistry, University of Warwick, Coventry, UK, CV4 7AL

DedA family proteins are a widespread group of bacterial inner membrane proteins whose molecular functions and mechanisms remain poorly understood. *Bacillus subtilis* UptA is implicated in recycling the essential lipid carrier undecaprenyl phosphate (UndP), which is required for bacterial cell-envelope biosynthesis[1,2]. Understanding how UptA transports UndP could therefore provide new insight into an important and potentially targetable aspect of bacterial physiology.

Here, we use a combination of coarse-grained and atomistic molecular dynamics simulations to investigate the molecular mechanism by which UptA transports UndP between membrane leaflets. Our simulations suggest that UndP flipping may be driven by proton-motive force, involving a network of conserved membrane-embedded acidic residues and arginine residues located within the core of the DedA fold. Simulations further reveal a potential role for the flexible C-terminal helix in facilitating the transport process. Preliminary simulations also suggest that *Escherichia coli* YabI and YghB and *Staphylococcus aureus* UptA may also support UndP transport in their respective pathogens. Together, our findings provide a potential model for DedA-mediated lipid transport and demonstrate how multiscale molecular dynamics can be used to investigate the mechanisms of poorly characterised membrane proteins.

References:

[1] I. J. Roney and D. Z. Rudner, *Nature*. (2023), **613**, 7945

[2] A. O. Oluwole, N. V. Kalmankar, M. Guida, J. L. Bennet, G. Poce, J. R. Bola and C. V. Robinson, *PNAS.* (2024), **121**, 41

Talk 12

Not so Classical: Non-Classicality as the Rule in CHX8+ Chemical Space

Stephen J. Harman, *Kristaps Ermanis.*

School of Chemistry, University of Nottingham, Nottingham, NG7 2RD, United Kingdom

Carbocations are fundamental intermediates in organic chemistry, invoked in mechanisms since the 1930s. In classical carbocations, the positive charge is formally located in a vacant p-orbital on the cationic carbon or delocalised through adjacent π-bonds. In non-classical carbocations (NCCs), this electron deficiency is stabilised through delocalised σ-bonds and three-centre two-electron (3c-2e) bonding. The 2-norbornyl cation is the defining example of non-classicality, its 3c-2e bridged structure the subject of a 60-year-long dispute over whether it was a single species or a pair of equilibrating classical structures. Computation supported a non-classical characterisation of the structure before a crystal structure confirmed it in 2013.[1] Currently, no systematic exploration of carbocation space exists, and non-classicality in cations must be established on a case-by-case basis using the same computational methods that demonstrated non-classicality in the norbornyl debate.

Here we present CHX8+, the first systematically generated dataset of hydrocarbon cations with up to and including eight carbon atoms. From an existing enumerative database of 31,497 optimised neutral hydrocarbons,[3] 95,546 carbocation structures have been generated. Following DFT-based geometry optimisations, we identify 37,936 unique carbocations which make up CHX8+. A tree-based machine-learning classifier has been developed. Trained on the geometries of these structures and validated against expert assignments it achieved a binary classification accuracy of 95.3±1.2%. This classifier therefore provides an automated route to cation assignment at scale that would be impractical for manual analysis.

Application of the machine learning model demonstrates that NCCs are the primary NCCs are the primary outcome of cation formation. Across all unique carbocations, 72.2±2.7% are non-classical, reframing NCCs as a general structural feature of carbocation chemistry, rather than as mere curiosities and historical exceptions.

References:

1. G. A. Olah, Hypercarbon Chemistry, J. Wiley & Sons, 2011
2. F. Scholz, D. Himmel, F. W. Heinemann, P. v. R. Schleyer, K. Meyer and I. Krossing, *Science,* 2013, **341**, 62–64.

Programme of Poster Presentations

1	Fernanda Souza, University of Warwick **Integrating Machine Learning, AlphaFold 3 Competitive Modeling, and Membrane Simulations to Identify Selective CB2R Agonists for Osteoarthritis**
2	Benedict Tan, University of Edinburgh **A Physics-Based Structural Prediction Pipeline for Molecular Glue Degraders**
3	Rianne van Diest, University of Basel / University College London **Thermodynamic Integration with Enhanced Sampling Enables Accurate Prediction of Bisphenol Binding to Estrogen Receptors**
4	Jack Hewitt, University of Liverpool **Artificial Intelligence-Driven Design of Macrocycle Molecules Against Drug Resistant *Pseudomonas aeruginosa***
5	Keyi Wei, University of Warwick **Investigating Protonation-Dependent Conformational Transitions in the Bacterial Membrane Flippase WzxE Using Integrated Computational Approaches**
6	Long-Hung Pham, Imperial College London **Towards Proteome-Wide Prioritisation of Functional Allosteric Residues by Atomistic Graph Perturbations**
7	Minjoon Seo, University of Oxford **Molecular Dynamics Reveals Mutation-Dependent Communication Networks in the Voltage-Sensing Domain of TPC2**
8	Vidhi Singhai, University of Bristol **Switching Cofactor Preference Disrupts Inter-Subunit Communication in Formate Dehydrogenase**
9	Kingsley Oguma, University of Oxford **Search Engines for Drug Discovery: Investigating Ranking Algorithms for Virtual Screening**
10	Kin Chao, University of Southampton **Force Field Benchmarking Reveals Limitations in NMR-Derived PROTAC Permeability Models**
11	Amelia Francis, Cardiff University **Tackling Norovirus Outbreaks in Our Hospitals: The Potential of Biosensors**
12	Ryan Demel, University of Oxford **Target-Conditioned Non-Canonical Amino Acid Incorporation via Objective-Based Substitutions of Template Peptides**

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

Integrating Machine Learning, AlphaFold 3 Competitive Modeling, and Membrane Simulations to Identify Selective CB2R Agonists for Osteoarthritis

Fernanda F. Souza[a,b], Juliana F. Vilachã[b], Phillip Stansfeld[b,c], *Kathia Maria Honorio[a,d]*

[a]Postgraduate Program in Science and Technology Chemistry, Federal University of ABC, Santo André, 09280-560, São Paulo, Brazil.

[b]School of Life Sciences, University of Warwick, Gibbet Hill Campus, Coventry, CV4 7AL, UK.

[c]Department of Chemistry, University of Warwick, Coventry, CV4 7SH, UK.

[d]School of Arts, Sciences and Humanities, University of São Paulo, 03828-000, São Paulo, Brazil.

Osteoarthritis (OA) is a prevalent joint disease with chronic pain and inflammation, where conventional palliative therapies carry severe long-term risks. The Cannabinoid Receptor 2 (CB2R) represents a compelling non-psychoactive target for peripheral analgesia, though developing selective agonists remains challenging due to high sequence homology with CB1R. Here, we integrate supervised machine learning (ML), 500 ns membrane molecular dynamics (MD), and AlphaFold 3 (AF3) modeling to discover selective CB2R candidates and probe conformational binding preferences.

A curated dataset of 2,750 compounds was generated from ChEMBL and BindingDB. Supervised ML classification (Random Forest, AUC = 0.97) and regression models (KNN R^2 = 0.78, MAE = 0.19) were trained on Morgan and PubChem fingerprints. Virtual screening across the Latin American Natural Products Database (LANaPDB) and ZINC20 identified selective, BBB-permeable, non-P-glycoprotein substrate hits. All-atom 500 ns MD simulations in POPC bilayers confirmed conformational stability for prioritized lead candidates, designated A) FL-01, B) FL-02, C) FL-03, and synthetic candidate D) FZ-04, reproducing the active-state binding pose, key orthosteric interactions, and thermodynamic binding energy of reference agonist E) WIN55,212-2.

To probe ligand-induced conformational preferences, 75 AF3 predictions were systematically evaluated across receptor constructs (cryo-EM 6PT0 vs. UniProt wild-type), heterotrimeric Gi-protein coupling, binary competitive pairs, and explicit multi-lipid mixtures. AF3 accurately reproduced experimental binding poses for WIN55,212-2 (RS = 0.94, ipTM = 0.94). In competitive pair runs, AF3 demonstrated intrinsic pocket selectivity, ranking binding preference as E > D > C > A > B and displacing lower-affinity compounds to peripheral loops. Assemblies in complex lipid membranes retained active-state structural fidelity (ipTM = 0.64). This integrated workflow establishes a robust strategy for GPCR ligand discovery, laying the ground for downstream coarse-grained simulations and experimental validation.

References:

[1] F. F. De Souza, J. F. Vilacha, H. De Paula, et al., *Front. Drug Discov., (2026), 6, 1809420.*

[2] J. Abramson, J. Adler, J. Dunger, et al., Nature, (2024), 630, 493–500.

Poster 2

A physics-based structural prediction pipeline for molecular glue degraders

Benedict Tan[a], Alessio Ciulli[b], *Julien Michel[a]*

a) School of Chemistry, University of Edinburgh, Joseph Black Building, Edinburgh, U.K., EH9 3FJ.

b) Centre for Targeted Protein Degradation, School of Life Sciences, University of Dundee, 1 James Lindsay Place, U.K., DD1 5JJ

Molecular glue induced targeted protein degradation is a powerful emerging therapeutic modality that exploits the ubiquitin-proteasome degradation pathway to break down proteins associated with disease proliferation. At the structural level, degradation proceeds through the formation of a ternary complex induced by small molecules between the target protein and an E3 ligase. In the absence of an experimental structure, modelling of the ternary complex through methods such as co-folding or docking is challenging. This primarily stems from the complexity and high degree of cooperativity associated with the ternary complex interface, as well as the limited number of public ternary complex structures.

We recently presented a Molecular Dynamics-based simulation protocol that uses Umbrella Sampling in the geometric route to calculate the ternary complex binding free energy[1]. Here, we apply this protocol within a structural prediction pipeline for CRBN-based molecular glues. An initial selection of target protein-E3 ligase complexes is generated using the popular open-source HADDOCK docking method. Initial filtering is performed using MM/GBSA calculations, followed by screening by the separation Potential of Mean Force (PMF). Finally, promising protein-protein interfaces are selected by the full calculated binding free energy. The binding of the glue at the protein-protein interface is modelled by pharmacophore-constrained docking. We demonstrate the end-to-end structural prediction pipeline for the CRBN-based GSPT1 degrader CC-885[2], and successfully reproduce the key interactions driving ternary complex formation.

Overall, our protocol shows promise as a means of predicting the mechanism driving molecular glue-induced degradation in the absence of expensive cryo-EM or X-ray crystallography experiments.

References:

[1] B. Tan, A. Ciulli and J. Michel, *ChemRxiv*, (2026).

[2] M. E. Matyskiela, G. Lu, T. Ito, et al., *Nature*, (2016), **535**, 252–257.

Poster 3

Thermodynamic Integration with Enhanced Sampling Enables Accurate Prediction of Bisphenol Binding to Estrogen Receptors

Rianne E. van Diest[a,b,c], Shunzhou Wan[c], *Martin Smieško[a,d], Peter V. Coveney[c], Markus A. Lill[a,b]*

a. Computational Pharmacy, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland

b. Swiss Centre for Applied Human Toxicology, University of Basel, Basel, Switzerland

c. Centre for Computational Science, Department of Chemistry, University College London, London, United Kingdom

d. Current affiliation: Société des Produits Nestlé, Nestlé Research, Lausanne, Switzerland

Endocrine-disrupting chemicals (EDCs) are an increasing public health concern due to their ability to interfere with hormonal signaling pathways. Bisphenol A is a well-known EDC that can mimic endogenous hormones by binding to nuclear receptors (NRs), including estrogen receptors. Although its use has been restricted, numerous structural analogues have been introduced as alternatives in plastic manufacturing, several of which also display endocrine-disrupting activity. Assessing bisphenol interactions with NRs remains challenging, as experimental outcomes depend strongly on biological context, resulting in inconsistent findings across studies. The rapid emergence of new analogues further highlights the need for efficient and reliable safety assessment methods.

Here, we present an *in silico* workflow based on Thermodynamic Integration with Enhanced Sampling (TIES) to estimate bisphenol binding affinities to NRs with near-experimental accuracy.[1] TIES combines thermodynamic integration with ensemble simulations to calculate relative binding free energies for structurally related compounds. While thermodynamic integration is widely applied in drug discovery, its use with enhanced sampling approaches for toxicological assessment remains limited, highlighting the potential of this strategy for endocrine disruption research.

The protocol was validated on estrogen receptor alpha (ERα) and estrogen receptor beta (ERβ) using a curated dataset of compounds with experimentally determined affinities. Accurate predictions were obtained for both receptors, with five out of nine compounds predicted within 1 kcal/mol of experimental values and all compounds within 2 kcal/mol. Binding mode analysis revealed multiple thermodynamically accessible ligand poses separated by low energy barriers, emphasizing the importance of adequately sampling alternative binding modes.

The validated workflow was subsequently applied to 17 bisphenol analogues with unknown affinities to enable affinity ranking. Ongoing work focuses on expanding the chemical space of bisphenols and evaluating transferability to additional NRs. This approach aims to support rapid and accurate screening of potential endocrine disruptors and contribute to more efficient hazard assessment.

References:

[1] A.P. Bhati, S. Wan et al. *J. Chem. Theory Comput.* (2017) **13**, 210–222.

Poster 4

Artificial Intelligence-driven Design of Macrocycle Molecules Against Drug Resistant *Pseudomonas aeruginosa*

Jack Hewitt, *Prof. Neil Berry, Dr. David W. Hong*

Department of Chemistry, University of Liverpool, Crown Street, Liverpool, L69 7ZD

The growing thread of multidrug-resistant *Pseudomonas aeruginosa*, as well as a dwindling and expensive pipeline of novel antibiotics, has created an urgent need for strategies that enhance the efficacy of existing antibiotics. Macrocyclic compounds are promising antibiotic adjuvants, offering conformational rigidity, three-dimensional chemical diversity, and the ability to selectively bind to targets often considered 'undruggable'. This study presents an integrated machine learning, deep learning, and future generative design framework for the discovery of macrocyclic adjuvants that potentiate clinically relevant antibiotics against P. aeruginosa.

A curated dataset of 6,000 macrocycles with reported adjuvant activity was analysed and used within the QSARtuna platform to construct two quantitative structure-activity relationship (QSAR) classification models: one trained on compounds that had potentiating activity with tobramycin, and another with colistin. QSARtuna's automated Bayesian optimisation workflow performed systematic descriptor selection, model building, and cross-validation across multiple machine learning algorithms.[1] The resulting random forest models achieved moderate predictive performance (AUROC=0.71, 0.81; Matthew's correlation coefficient=0.16, 0.34, and $EF_{1\%}$=6.49, 4.69, for tobramycin and colistin, respectively). We have also established a deep learning pipeline. By implementing various data engineering methods, loss functions, and architectures, we then developed a multitude of neural network models; these were designed to try and look past assay noise to identify subtle relationships that correspond to rare active compounds. Evaluated under the Polaris framework, a multilayer-perceptron employing the focal loss function consistently outperformed graph-based models. Using Matthews Correlation Coefficient and Enrichment Factors, eight champion architectures were isolated. Further statistical validation, uncertainty quantification, and benchmarking confirms that these models function optimally as ranking engines rather than absolute predictors.

Our in-house modification extended REINVENT4's generative capabilities to accommodate macrocyclic ring closures.[2] Using Linkinvent, we can guide REINVENT4's priors to splice a warhead compound, fill in the specified gap with a generated linker and then perform a macrocyclic ring closure. Reinforcement learning will guide molecular generation using multi-objective optimisation to balance predicted activity, synthetic accessibility, and drug-likeness.

These results demonstrate the power of combining predictive QSAR modelling with generative reinforcement learning for targeted macrocycle discovery. The integration of optimised QSAR models and a macrocycle-compatible REINVENT4 framework provides a scalable, automated approach for designing next-generation antibiotic adjuvants capable of restoring tobramycin and colistin potency against resistant P. aeruginosa.

1. L. Mervin, A. Voronov, M. Kabeshov and O. Engkvist, J. Chem. Inf. Model., 2024, **64**, 5365-5374.

2. H. H. Loeffler, J. He, A. Tibo, J. P. Janet, A. Voronov, L. H. Mervin and O. Engkvist, Journal of Cheminformatics, 2024, **16**, 20.

Poster 5

Investigating protonation-dependent conformational transitions in the bacterial membrane flippase WzxE using integrated computational approaches

Keyi Wei, *Phillip J. Stansfeld*

School of Life Science, University of Warwick, Gibbet Hill Campus, Coventry, CV4 7AL

The outer membrane of Gram-negative bacteria forms a major permeability barrier to antibiotics, making its biogenesis an important target for antimicrobial development. Enterobacterial common antigen (ECA) is a conserved surface polysaccharide in Enterobacterales that contributes to bacterial envelope integrity. Its biosynthesis requires the flipping of lipid-linked trisaccharide repeating units across the inner membrane, mediated by the membrane flippase WzxE. However, how WzxE undergoes conformational transitions to facilitate substrate flipping remains poorly understood.

Recent cryo-electron microscopy (cryo-EM) structures have revealed inward-facing and outward-facing conformations of WzxE [1], highlighting the potential importance of conformational transitions in facilitating substrate flipping. However, these structures represent apo states, leaving the roles of substrate binding, protonation, and the energetic driving forces underlying conformational transitions unresolved.

Here, we apply an integrated computational approach to investigate the conformational dynamics and substrate flipping mechanism of WzxE. A WzxE–trisaccharide complex was constructed using AlphaFold3, followed by unbiased molecular dynamics (MD) simulations to examine substrate–protein interactions and conformational flexibility. Constant-pH molecular dynamics (CpHMD) simulations were employed to investigate how changes in protonation states influence WzxE conformational dynamics and potentially facilitate transitions between inward-facing and outward-facing states. Preliminary CpHMD results suggest that substrate binding and conformational changes modulate the proton affinity of Glu367, indicating potential coupling between protonation and the substrate flipping cycle. Umbrella sampling simulations are being performed to characterise the free-energy landscapes of WzxE in its apo and substrate-bound states under different protonation conditions, providing insights into the energetic barriers associated with conformational rearrangements. In parallel, generative AI approaches are being explored to identify potential intermediate conformations and transition pathways between experimentally resolved states.

By integrating molecular dynamics, enhanced sampling, and generative AI-based modelling, this study aims to elucidate how substrate binding and protonation influence the conformational landscape of WzxE, providing mechanistic insights into membrane flipping and demonstrating the potential of complementary computational approaches for investigating complex membrane transport processes.

References:

[1] A. Le Bas, B. R. Clarke, T. Teelucksingh et al., Open Biol., (2025), 15, 240310.

Poster 6

Towards Proteome-wide Prioritisation of Functional Allosteric Residues by Atomistic Graph Perturbations

Long-Hung Pham[a], *Mauricio Barahona[b], Sophia Yaliraki[a]*

[a]Department of Chemistry, Imperial College London, United Kingdom

[b]Department of Mathematics, Imperial College London, United Kingdom

Despite the major advancement of deep learning-based protein structure prediction, the characterization of the underlying biophysics and functional significance of individual residues remains poorly understood, hindering the prioritisation of new non-catalytic, allosteric sites.

At the same time, the emergence of new chemical modalities, such as reactive covalent ligands and cyclic peptides, has opened up vast new opportunities to tackle traditionally challenging protein surfaces. Despite the wealth of data on potentially ligandable sites, it also remains challenging to identify which binding events can truly modulate protein functions. Thus, mapping the functional landscape of non-catalytic, allosteric residues is a major bottleneck to expanding the druggable proteome.

In this poster, building on protein atomistic graphs and the bond-to-bond propensity framework[1], we developed new measures that quantify the positional perturbation (PosPert) importance of a residue within the protein biophysical network. Computation on more than 10,000 CATH proteins establishes a reference distribution of PosPert values in diverse proteins and reveals amino-acid-specific enrichment of functional motifs in highly scored residues. Retrospective case studies of challenging allosteric functional sites (WRN helicase conformational lock, AKT1 neo-zinc, CFTR corrector sites) showed that PosPert outperforms deep-learning-based pocket prediction (AF2Bind), or simple graph betweenness centrality methods. To demonstrate the scalability of the computation, we conducted proteome-wide prediction on human proteins with available covalent ligandability data and identified numerous allosteric pockets prioritised by our measures, presenting new opportunities to tackle understudied proteins. Finally, we prospectively predicted potentially functional residues for essential proteins in the HAP1 haploid cell line, in an ongoing experimental validation with CRISPR-based mutagenesis fitness assay[2]. Taken together, the atomistic graph approach holds great prospects for efficient exploration of functional residues at the proteome scale.

References:

[1] B. R. C. Amor, M. T. Schaub, S. N. Yaliraki, M. Barahona, *Nat. Commun.*, (2016), **7**, 1–13.

[2] H. Benns, M. Storch, J.A. Falco, et.al., *Nat. Microbiol.*, (2022), **7**, 1891-1905.

Poster 7

Molecular dynamics reveals mutation-dependent communication networks in the voltage-sensing domain of TPC2

Minjoon Seo[a], Yu Yuan[b], Dawid Jaślan[c], Sandip Patel[b], *Philip C. Biggin[a]*

[a]Department of Biochemistry, University of Oxford, Oxford, UK

[b]Department of Cell and Developmental Biology, University College London, London, UK

[c]Walther Straub Institute of Pharmacology and Toxicology, Faculty of Medicine, Ludwig-Maximilian University, Munich, Germany

Two-pore channel 2 (TPC2) is an endolysosomal cation channel that regulates $Ca^{2+,}$ Na^{+} and H^{+} homeostasis. It is activated by endogenous signalling molecules, including NAADP and $PI(3,5)P_2$, which differentially regulate its ion selectivity. Although TPC2 is conventionally considered voltage-independent, its second voltage-sensing domain (VSD2) partially retains structural features of canonical voltage sensors. However, the role of this 'cryptic' voltage sensor to ligand-induced channel activation remains unclear. Lysosomal electrophysiology results demonstrate that charge mutations within the IIS4 helix of VSD2 differentially affect ligand-induced ion permeation. While I551R, R554A and R557A impair ligand-induced Ca^{2+} permeation, R554A uniquely preserves H^{+} permeability.

Here, we performed all-atom molecular dynamics simulations and network analysis to investigate how VSD2 mutations alter allosteric communication within TPC2. Wildtype (WT) TPC2 and three IIS4 mutants (I551R, R554A, R557A) were each simulated in three independent repeats of 500 ns. Residue-residue coupling was quantified using normalised mutual information calculated from phi (ϕ) dihedral angle dynamics, and communication pathways were identified using MDPath[1]. Pathway clustering and residue- and region-specific participation analyses were then used to characterise mutation-dependent network redistribution.

Surprisingly, the analyses identified VSD2 as a major component of communication network in both WT and mutant TPC2, suggesting that this domain contributes to allosteric communication despite the channel's conventional voltage insensitive gating mechanism. WT, I551R and R557A each exhibited a dominant pathway cluster comprising at least 97% of their identified pathways, whereas R554A exhibited greater pathway cluster heterogeneity. Furthermore, individual mutants showed distinct patterns of residue-level pathway participation involving the IIS4-S5 linker and the pore-forming regions.

Together, our findings suggest that VSD2 charge mutations reshape dynamic communication networks connecting the cryptic voltage sensor to the pore. This work identifies candidate 'hidden' allosteric pathways that may contribute to the differential ion-permeation phenotypes associated with VSD2 mutants.

References:

1. N. P. Doering, M. Taterra, M. Bermúdez and G. Wolber, *Journal of Chemical Information and Modeling*, 2025, **65**, 11123-11135.

Poster 8

Switching cofactor preference disrupts inter-subunit communication in formate dehydrogenase

Vidhi Singhai, Dan Pike, *Prof. Sofia Oliveira, Prof. Adrian Mulholland*

School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS

Nicotinamide cofactor regeneration is central to industrial biocatalysis. Many industrially relevant oxidoreductases are NADPH-dependent, so engineering $NADP^+$ acceptance into regeneration enzymes broadens their use, though often at the price of catalytic performance. Formate dehydrogenase from *Pseudomonas sp. 101* (PseFDH) is widely used for NAD^+ regeneration, and directed evolution produced V9, a five-mutation variant that prefers $NADP^+$ but has reduced catalytic efficiency. Changes to the cofactor-binding pocket do not fully account for this loss,[1] and since PseFDH is a homodimer, altered communication between its active sites may also contribute.

Wild-type and V9 PseFDH were compared using two complementary approaches. Equilibrium molecular dynamics, analysed by dynamical cross-correlation and shortest-path maps, identifies the pathways connecting the active sites. Dynamical non-equilibrium molecular dynamics (D-NEMD)[2] then probes those pathways directly: removing formate from one or both active sites perturbs an equilibrium ensemble instantaneously, and the time-resolved structural response shows how that local change propagates through the protein.

In V9, inter-subunit communication is altered rather than lost. At equilibrium, correlated pathways between the active sites survive in both enzymes, with only the mutation sites decoupling from the partner subunit. The difference appears under perturbation: removing formate from one V9 active site redirects the response towards the cofactor-binding cleft and reverses the dominant direction of inter-subunit communication. Additionally, perturbing both sites reinforces the signal reaching the second active site in wild-type PseFDH but suppresses it in V9. This points to a contribution to the activity loss beyond the cofactor-binding pocket.

This raises the question of which substitutions could restore activity without sacrificing $NADP^+$ selectivity. Deep mutational scanning (DMS) by experimental collaborators has quantified the fitness effects of individual substitutions in V9, but cannot explain how distant mutations influence function. D-NEMD is therefore being extended to perturb individual residues and trace their effects to the active site, linking experimentally identified mutations to mechanism and identifying candidates for restoring inter-subunit communication.

References:

[1] L. Calzadiaz-Ramirez, C. Calvó-Tusell, G. M. M. Stoffel, S. N. Lindner, S. Osuna, T. J. Erb, M. Garcia-Borràs, A. Bar-Even and C. G. Acevedo-Rocha, *ACS Catalysis*, 2020, **10**, 7512–7525.

[2] G. Ciccotti and M. Ferrario, *Entropy*, 2013, **16**, 233–257.

Poster 9

Search Engines for Drug Discovery: Investigating Ranking Algorithms for Virtual Screening

Kingsley O. Oguma, *Prof. Garrett M. Morris*

Department of Statistics, Oxford Protein Informatics Group, *24-29 St Giles', Oxford OX1 3LB*

Structure-based virtual screening aims to prioritise active compounds from large chemical libraries for experimental testing. We propose viewing virtual screening as a Search Engine, where the target protein acts as a query, compounds as documents, and the objective is to rank compounds by their likelihood of activity. Although screening performance is ultimately determined by the quality of this ranked list, most scoring functions evaluate protein–ligand complexes independently and obtain a ranking only by sorting their predictions.

We introduce DrugRank, a LambdaMART-based virtual screening Search Engine that learns to rank compounds by combining heterogeneous relevance signals from GNINA structure-based scores, FCFP4 ligand fingerprints, ESM-2 protein embeddings, and drug-target interaction network features. The DTI score is inspired by PageRank, where information from connected nodes contributes to an item's ranking. DrugRank is trained on leakage-filtered BigBind data and evaluated against GNINA, with particular emphasis on the seven biochemical assays in LIT-PCBA.

On the biochemical LIT-PCBA subset, DrugRank achieves statistically significant improvements in early enrichment across several targets. At NEF1%, DrugRank statistically significantly outperforms GNINA on FEN1 and KAT2A, with the largest improvement observed for FEN1, where NEF1% increases from 0.022 to 0.081. At NEF10%, DrugRank statistically significantly outperforms GNINA on IDH1, VDR, and FEN1, including an increase from 0.051 to 0.205 for IDH1 and from 0.228 to 0.397 for FEN1. DrugRank also achieves statistically significant improvements in BEDROC and PR-AUC on FEN1 and KAT2A, and FEN1 also shows a statistically significant improvement in ROC-AUC.

These results show that framing virtual screening as a Search Engine and learning directly from heterogeneous relevance signals can improve early enrichment on biochemical LIT-PCBA targets. Next, we will test whether this generalises to other benchmarks and prospective screening.

Poster 10

Force Field Benchmarking Reveals Limitations in NMR-Derived PROTAC Permeability Models

Kin W. Chao[a], Ileana Guzzetti[b], Karolina Kwapień[b], Tomas Leek[b], Leonardo De Maria[c], Christian Tyrchan[b], Werngard Czechtizky[b], *Jonathan W. Essex*[a]

[a]School of Chemistry and Chemical Engineering, University of Southampton, Southampton, UK
[b]BioPharma Chemistry, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden
[c]Hit Discovery, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

Proteolysis-targeting chimeras (PROTACs) are large beyond-Rule-of-5 (bRo5) molecules with physicochemical properties apparently unfavourable for passive membrane permeation, yet several achieve good oral bioavailability. How these molecules cross cell membranes remains unclear. Two perspectives have been proposed in the current literature to relate PROTAC structure to permeability: (1) molecular chameleonicity,[1] where environment-dependent conformational change reduces the exposed polar surface area (PSA) in membrane-like apolar environment, and (2) intramolecular hydrogen bonding or steric shielding, that reduces the number of solvent-exposed H-bond donors (eHBD) in an apolar environment, with an upper limit of eHBD ≤ 2 for good bioavailability for PROTACs.[2]

Here we use molecular dynamics (MD) simulations to study three literature PROTACs with reported experimental NOE data in both aqueous- and membrane-like environments that underpin the two models above. (1) PROTAC1 (alkyl linker, low reported PAMPA permeability) and PROTAC2 (PEG linker, high reported PAMPA permeability) are a matched pair with identical warheads,[1] and (2) ARV-766 is a clinical androgen receptor PROTAC whose three NH groups span the full range of eHBD shielding classifications (shielded in chloroform only, shielded in both solvents, and exposed in both).[2] Using four modern force fields (GAFF2, OpenFF 2.2.1, XFF, and CGenFF) we perform long replica exchange with solute tempering 2 (REST2) MD simulations in explicit water and chloroform, together with supplementary GAFF2 DMSO simulations and B3LYP/6-31G(d) QM torsion scans.

We found that CGenFF suffered from severe replica trapping in chloroform and was therefore excluded from further analysis here. The remaining three force fields broadly agree on general conformational trends but show moderate inter-force-field variability. Detailed comparison with published NOE distance constraints reveals substantial discrepancies: for PROTAC1 and PROTAC2, the short long-range NOE contacts between warheads that underpin the chameleonicity model are largely not reproduced by any force field, and for ARV-766, a key NOE distance central to the HBD shielding classification maps to a torsional local minimum that is energetically inaccessible in the QM torsion profile and rarely sampled by all three force fields. These findings highlight potential limitations in both the computational and experimental methods used to study large PROTACs, and raise questions about the current perspectives derived from these data.

References:
[1] V. Poongavanam, S. Peintner, Y. Abeje, F. Kölling, D. Meibom, M. Erdelyi and J. Kihlberg, *ACS Med. Chem. Lett.*, (2025), **16**, 681.
[2] M. Schade, J. S. Scott, T. G. Hayhow, A. Pike, I. Terstiege, M. Ahlqvist, J. R. Johansson, C. R. Diene, C. Fallan, A. Y. S. Balazs, E. Chiarparin and D. Wilson, *J. Med. Chem.*, (2024), **67**, 13106–13116.

Poster 11

Tackling norovirus outbreaks in our hospitals: The potential of biosensors

Amelia R. Francis[a], Professor Sébastien Côté[b], *Dr Georgina E. Menzies[a], Professor D. Dafydd Jones*[a]

[a] School of Biosciences, Cardiff University, Sir Martin Evans Building, Museum Ave, Cardiff, CF10 3AX

[b] Department of Physics, Université de Montréal, Complexe des sciences, 1375, ave Thérèse-Lavoie-Roux, Montréal, QC H2V 0B3

Costing the NHS an estimated £107 million pounds every year, norovirus outbreaks result in a significant loss of hospital resources. Mitigation of outbreaks is reliant on rapid, sensitive detection systems, which is not addressed by current diagnostic approaches [1]. One solution is the development of portable, fast-acting biosensors, based upon the binding of a small antibody fragment (nanobody) to the P domain of the norovirus capsid protein. Through incorporation of a photoreactive nonnatural amino acid (nnAA) into the protein sequence the nanobody can be attached in a reproducible manner to a graphene-field effect transistor. This can detect the analyte binding and convert it to a measurable signal.

Computational modelling of the mutated nanobody has been undertaken to assess the viability of different mutation positions, with post-simulation analysis of the root mean squared deviation (RMSD) of the nnAA side chain revealing the common side chain configurations [2]. Alignment of these configurations to the graphene plane reveals the likelihood of successful conjugation to the nanocarbon, as well as the potential electrostatic effect of the bound norovirus P domain on graphene.

Initial modelling at three different positions (position A, B and C) suggests that the RMSD analysis can successfully predict the rotamer states of the nnAA, with downstream processing indicating that position A could result in successful binding of the nanobody and the P domain close to the graphene plane. However, this method was unsuitable for high-throughput analysis of mutation sites. Custom code was written to use MD trajectory frames from simulations of the wildtype nanobody as an input. The code assembles the individual nanobody frames, P domain, nnAA and graphene for all possible mutation positions. The electrostatic effect of P domain binding on graphene for each ensemble is then predicted. Position A was again identified as a promising candidate. This variant will therefore be taken forward to protein production and biosensor construction.

This work has established a replicable workflow for exploring nnAA mutation positions, which can be applied to other nanobody-analyte pairs for biosensor development.

References:

[1] N. Wang, G. Pan, P. Liu, S. Rong, Z. and Gao, Q. Li, *Pathogens,* (2021), **10**, 1383

[2] R. Gwyther, S. Côté, CS. Lee, H. Miao, K. Ramakrishnan, M. Palma and D. Jones, *Nat Commun*, (2024), **15**, 7482

Poster 12

Target-conditioned non-canonical amino acid incorporation via objective-based substitutions of template peptides

Ryan Demel [a], *Annette von Delft* [a], Anthony Bradley [b], Paul E Brennan [a]

[a] Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford,

[b] Department of Chemistry, University of Liverpool

Peptides provide strong scaffolds for targeting protein interaction sites, but proteolytic degradation, poor permeability and weak binding affinity can limit their therapeutic utility.[1] Non-canonical amino acids expand the chemistry available to address these limitations. Computational support for this chemistry, however, does not necessarily establish which residues are suitable at a given position or how they will behave in combination. Especially in generative/co-folding models. A plausible predicted structure alone is not reliable for establishing favourable binding.[2]

We developed a computational workflow for optimising existing peptides in peptide-protein complexes through non-canonical amino acid substitution. Selected amino acids were parameterised and checked for geometry, charge and rotamers before being substituted into peptide designs. We evaluated single and multiple substitutions and compared positional selection strategies to examine how residue chemistry and location influenced design outcomes. Shortlisted designs were then assessed against native peptides under matched computational conditions.

Modified peptides achieved more favourable calculated interface scores across multiple targets, with selected designs retaining local binding poses under the applied refinement conditions. The location of substitutions influenced the recovery of favourable designs. Their combined effects also frequently differed from the sum of their individual contributions, showing why modifications need to be evaluated together within the surrounding protein environment.

By connecting chemical selection with structural and energetic assessment, this work provides a basis for prioritising peptide modifications for further testing.

References:

[1] Y. Wang*, Front. Chem*., 2022, **10**, 1–19.

[2] Petsalaki E, Stark A, García-Urdiales E, Russell RB. PLoS Comput Biol. 2009 Mar;5(3):e1000335. doi: 10.1371/journal.pcbi.1000335. Epub 2009 Mar 27. PMID: 19325869; PMCID: PMC2653190.

Useful Information

Travel to Lady Margaret Hall

https://core.web.ox.ac.uk/sites/default/files/core/documents/media/lmh_travel_information.pdf

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

