



MGMS and RSC MMG Young Modellers' Forum 2018 PROGRAMME & ABSTRACTS

Programme of Oral Presentations

9.15 – 9.50	Coffee and Registration
9.50 – 10.00	Welcome and Introduction
10.00 – 10.20	Christina Founti, Information School, University of Sheffield How low can you go? Assessing precision in models' predictions
10.20 – 10.40	Kieran Selvon, School of Chemistry, University of Southampton Simulations of drug permeation using a dual resolution membrane model
10.40 – 11.00	Jonathan Shearer, School of Chemistry, University of Southampton The Outer Membrane proteins OmpA, OmpX, OmpF, BtuB, FhuA and EstA have unique lipopolysaccharide fingerprints
11.00 – 11.30	Poster Presenters "Lightning Talks"
11.30 – 11.50	Tea/Coffee break
11.50 – 12.10	Damiano Spadoni, Department of Chemical and Environmental Engineering, University of Nottingham Investigating aerobiosis in radical SAM enzymes: molecular modelling of lysine 2,3-aminomutase
12.10 – 12.30	Francesca Vianello, Department of Chemistry, Imperial College London From Atomistic Descriptions to Protein-Protein Interaction Networks
12.30 – 12.50	Michael Connolly, School of Chemistry, University of Bristol Investigating enzyme temperature dependence and heat capacity using computer simulation
12.50 – 14.10	Lunch and Poster Session
14.10 – 14.30	Timothy Burd, Department of Chemistry, University of Oxford Quantum Dynamics in the Atmosphere
14.30 – 14.50	Kimberly Bowal, Department of Chemical Engineering and Biotechnology, University of Cambridge Soot particle structure: Insights from molecular dynamics and Monte Carlo simulations
14.50 – 15.10	Maeve Kavanagh, Department of Chemistry, Imperial College London Excited States of the Photosystem II Reaction Centre
15.10 – 15.30	Sehee Na, Institute of Physical Chemistry, University of Freiburg Thermodynamic Integration Network Study of Electron Transfer: from Proteins to Aggregates
15.30 – 15.30	Tea/Coffee break
15.50 – 16.10	Lisa Patrick, Department of Chemistry, University of Edinburgh Revealing allosteric pathways with Information Theory
16.10 – 16.30	Helen Deeks, Department of Chemistry, University of Bristol Molecular Tetris: Human-guided exploration of a protein-ligand system in Virtual Reality
16.30 – 17.00	Deliberations and Prizes
17.00	End

Talk 1

How low can you go? Assessing precision in models' predictions

Christina Founti^a, Val Gillet^a, Jonathan Vessey^b

^a The University of Sheffield, Information School, Regent Court, 211 Portobello, Sheffield

^b Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds

Estimates of experimental error are essential in model evaluation and are consulted to validate the accuracy of predictive models [1]. However, these are not easy to account for when evaluating the precision of the models' predictions, since repeated measurements are not often available in modelling data and the performance of common algorithms deteriorates. Nevertheless, it is understood that measurement error is variable across compounds. Precision in models is addressed by making assumptions regarding a model's error distribution and, thus, introducing confidence to prediction error estimates for individual predictions. Yet, for predictions to be useful, determining an optimal trade-off between confidence and precision is necessary as precision deteriorates with higher confidence. The present work evaluates the precision of models' predictions against the precision of the experimental measurements.

Ensemble models were trained on ADME (Absorption, Distribution, Metabolism and Excretion) data with repeated measurements and prediction error was estimated using prediction variation. Two methods are studied for the assessment of precision estimates. The first method requires that measurements and predictions are represented as Gaussian distributions and is based on distribution divergence. The second method evaluates prediction interval estimates, which are expected to enclose future measurements with a given probability. Conformal prediction uses repeated sampling to infer an empirical error distribution from existing data and then generates prediction intervals that are associated with a user-defined level of confidence. For this method, precision was evaluated by comparing the width of prediction intervals to the width of confidence intervals of measured data that were estimated from repeated measurements.

The results presented from the two methods demonstrate that taking into account the precision of measurements in the modelled data is useful in the validation of prediction error estimates.

[1] M. Wenlock and L. Carlsson, *JCIM*, (2015) **55**, 125 – 34.

Talk 2

Simulations of drug permeation using a dual resolution membrane model

Kieran Selvon^a, Richard Ward ^b, Jonathan W. Essex ^a

^a The School of Chemistry, University of Southampton, Highfield, SO171BJ.

^b AstraZeneca, IMED Oncology, Cambridge, UK.

The permeation of molecules through the cell membrane is the mechanism by which the majority of exogenous drugs enter into the general circulation of the body. A molecule's permeability is therefore of great importance in the field of drug design. From a commercial standpoint, it would be highly desirable to have a way to reliably and quickly calculate a molecule's permeability. Pharmaceutical companies require the ability to test large numbers of compounds, a chemical series for a newly developed drug, for instance. It would be very costly and highly impractical to synthesize a large number of compounds to test experimentally without any notion of their potential efficacy. Multi-scale computational methods may hold the key to solving this problem.

ELBA¹ is a multi-scale model for water and lipid membrane systems developed primarily at the University of Southampton, in which water and lipid molecules are coarse grained, but interact naturally through electrostatics with atomistic solutes embedded in the system. Ideally, simulations can be performed with speed comparable to that of fully coarse grained models, but with accuracy comparable to that of atomistic models.

In this project, ELBA has been used to calculate free energy profiles across a model bilayer membrane on a drug-molecule test set provided by AstraZeneca. These free energy calculations can be used to compute the relative bilayer permeabilities² of the test set. In this presentation, these results will be compared against experimental permeability data, to assess the ability of the simulations to reproduce cell permeabilities. In addition, the simulations have been analysed extensively to understand the behaviour of the molecules as they cross the bilayer, including their conformational changes, hydrogen bonding and orientations, providing valuable insight into mechanisms of passive permeation.

References:

- [1] M. Orsi and J.W. Essex, *PLoS ONE*, (2011) **6**, 1-22.
- [2] S. Genheden and J.W. Essex, *JCTC*, (2015) **11**, 4749-4759.

Talk 3

The Outer Membrane proteins OmpA, OmpX, OmpF, BtuB, FhuA and EstA have unique lipopolysaccharide fingerprints

Jonathan Shearer, Damien Jefferies, Syma Khalid

School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ.

The outer membrane of Gram-negative bacteria has a highly complex asymmetric architecture, containing a mixture of phospholipids in the inner leaflet and in the outer leaflet almost exclusively lipopolysaccharide molecules. In *E. coli*, the outer membrane contains a wide range of membrane proteins with beta barrel architectures, that vary in size from the smallest having eight strands to larger barrels composed of twenty-two strands. Lipid-protein interactions are significant in determining the function and organisation of membrane proteins, through macroscopic bilayer properties or specific protein-lipid interactions. Computational and experimental studies have revealed a significant amount with respect to protein-lipid interactions. A majority of computational studies of outer membrane proteins have focussed on protein-lipid interactions in symmetric systems, which are not representative of the *in vivo* environments of these proteins.

Here we report coarse-grained molecular dynamics simulations of five proteins from the *E. coli* outer membrane OmpA, OmpX, OmpF, BtuB, FhuA and EstA from *Pseudomonas aeruginosa* in a range of lipopolysaccharide-containing membrane environments. We show that each protein has a unique lipopolysaccharide fingerprint, that often varies with the smoothness of the lipopolysaccharide environment. Overall we present analyses from over 200 microseconds of simulation for each protein.

Talk 4

Investigating aerobiosis in radical SAM enzymes: molecular modelling of lysine 2,3-aminomutase

Damiano Spadoni,^a Dr. Anna K. Croft^a, Dr. Christof M. Jäger^a and Dr. Charles A. Laughton^b

^aDepartment of Chemical and Environmental Engineering, University of Nottingham, Nottingham, NG7 2RD, UK

^bSchool of Pharmacy, University of Nottingham, Nottingham, NG7 2RD, UK

The radical S-adenosyl methionine (SAM) enzyme superfamily represents an ensemble of proteins that are able to catalyse biochemical reactions involving organic radical intermediates¹. The intermediate radical abstracts an H-atom from the substrate which then undergoes a range of reactions which are difficult to accomplish in the laboratory with traditional chemical techniques. Members of this class of enzymes could then be used as a novel biocatalyst for the production of a wide variety of chemical compounds. Virtually every SAM enzyme is known to be unstable in air due to the requirement of a catalytic, oxygen sensitive [4Fe-4S] cluster cofactor which results in a loss of catalytic activity under oxygen exposure making the study of these enzymes technically challenging.

Lysine 2,3-aminomutase (LAM) is a widely studied radical SAM enzyme which has a naturally existing oxygen-tolerant variant² from a different microorganism but it is low in activity compared to the air-sensitive LAM and its structure is also unknown. The oxygen-tolerant LAM structure has been predicted here through homology modelling using the existing oxygen sensitive crystal structure as the template. The model of the oxygen-tolerant LAM is evaluated and its dynamics studied and compared to the known enzyme using MD simulations. A QM approach has also been used to study in detail the [4Fe-4S] cluster-interacting residues that may play a role in the oxygen tolerance. The goal is to explore and understand the source of the exceptional oxygen tolerance in the SAM enzyme, which should allow us to transfer and exploit this feature to other members of the SAM family, enabling easier access to radical SAM enzymes in the lab and larger-scale biotechnological applications.

References:

- [1] C. J. Walsby, D. O. Ortillo, J. Yang, M. R. Nnyepi, W. E. Broderick, B. M. Hoffman, J. B. Broderick, *J. Am. Chem. Soc.*, (2002), 124, 3143.
- [2] B. W. Lepore, F. J. Ruzicka, P. A. Frey, D. Ringe, *Proc. Natl. Acad. Sci. U. S. A.*, (2005), 102, (39), 13819.

Talk 5

From Atomistic Descriptions to Protein-Protein Interaction Networks

Francesca Vianello^a, Sophia N. Yaliraki^a, Mauricio Barahona^b

^a Department of Chemistry, Imperial College London, 80 Wood Ln, London W12 0BZ

^b Department of Mathematics, Imperial College London, Exhibition Road, London SW7 2AZ

Protein-protein interactions (PPIs) are central to all regulatory processes in nature, from metabolic interactions to DNA replication. Therefore, they are thought as especially important nodes in disease pathways and considered particularly attractive drug targets. Many experimental methods (and some computational approaches) focus on the detection of potential interactions that involve one or a family of proteins but stop short of characterising the structural details of these interactions. However, targeted drug discovery efforts require a knowledge of the interactions at the atomistic and bond-specific level for an effective design of small molecules aimed at interfering with specific regulatory processes.

We put forward a novel theoretical framework (originated from the investigation of intra-protein allosteric communication¹) for the structural characterisation of PPI sites, based on the exploration of the edge space of a protein network constructed from the atomic coordinates and chemical interactions between atoms. This approach allows not only for the direct integration with drug discovery approaches, but also for the implementation to multiple protein structures (such as an entire signalling cascade) without unreasonable computational expenditure.

We have applied our methodology to the Glucagon-like peptide 1 (GLP1) receptor protein, a transmembrane structure involved in the control of blood sugar levels by enhancing insulin secretion. Our results identify a possible signal propagation pathway through the receptor and allow us to quantify the likelihood for specific residues to be interacting with a separate protein structure (this receptor forms complexes with G-proteins, a family of intracellular molecular switches). PPI interfaces on GLP1 have previously been estimated only by space filling contact maps.² It is worth noting that the implementation of our methodological framework does not require any knowledge of the interacting protein complexes. Pathways are found on one structure and, in cases such as the GLP1 receptor where there exists a stable structure of both interacting partners, subsequent superimposition validates predictions. We have also expanded our study to include a selection of receptors from different families, all with promising results, opening the way towards a larger scale systematic study of G-protein dynamics.

References:

[1] B. R. C. Amor *et al.*, *Nat. Commun.*, (2016), **7**, 1–30

[2] Y. Zhang *et al.*, *Nature*, (2017), **546**, 248-253

Talk 6

Investigating enzyme temperature dependence and heat capacity using computer simulation

Michael S Connolly^a, Marc W Van der Kamp^a, Erica J Prentice^b, Vickery L Arcus^b and Adrian J Mulholland^a

^aCentre for Computational Chemistry, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom

^bSchool of Science, University of Waikato, Hamilton 3240, New Zealand

Heat capacity is emerging as a fundamentally important property in understanding and explaining the evolution and thermoadaptation of enzyme catalysis. Enzymes lose activity above their optimum temperature, but the textbook explanation that this is due to protein unfolding alone does not explain this behaviour.¹ Arcus *et. al* have shown that the change in heat capacity of activation can account for the observed temperature dependence of rates for many enzymes, and developed macromolecular rate theory (MMRT) to formalize this concept.^{1,2} This unique view suggests that changes in protein dynamics are central to enzyme thermoadaptation. It also helps to explain why enzymes are such large proteins, why they have such complex and varied structures, and how the optimum temperature of each enzyme is 'tuned' by evolution. MMRT suggests that there is a negative ΔC_p^\ddagger for many enzyme-catalysed reactions, which determines the optimum temperature of enzyme activity and accounts for the curvature of their temperature-rate profiles. We have carried out molecular dynamics (MD) simulations of two enzymes (ketosteroid isomerase (KSI) and an α -glucosidase (MalL)) that give activation heat capacities in good agreement with experiment.² The simulations also identify, in detail, the dynamical changes that give rise to the observed heat capacity changes. This shows that MD simulations can be used to identify and investigate the physical origins of thermoadaptation. Partial ΔC_p^\ddagger values have been calculated in order to analyse the contributions to heat capacity change from different protein domains, revealing unexpected behaviour large changes in heat capacity from regions of the enzyme far removed from the active site. The results are in good agreement with the experimentally determined values, and show that simulation can give insight into the importance of ΔC_p^\ddagger in enzyme chemistry.

References:

[1] Arcus, V. L.; Prentice, E. J.; Hobbs, J. K.; Mulholland, A. J.; Van der Kamp, M. W.; Pudney, C. R.; Parker, E. J.; Schipper, L. A. On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry* 2016, 55 (12), 1681–1688 DOI: 10.1021/acs.biochem.5b01094.

[2] Kamp, M. W.; Prentice, E. J.; Kraakman, K. L.; Connolly, M.; Mulholland, A. J.; Arcus, V. L. Dynamical Origins of Heat Capacity Changes in Enzyme-Catalysed Reactions. *Nat. Commun.* 2018, 9 (1), 1177.

Talk 7

Quantum Dynamics in the Atmosphere

Timothy A. H. Burd, Xiao Shan, David C. Clary

Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, OX1 3QZ

Using *ab initio* quantum modeling to predict the rate constants of chemical reactions can help build better kinetic models of the atmosphere, to interpret experimental results and to predict novel reaction pathways.

For reactions with many atoms, traditional ways of calculating these rate constants are either far too computationally expensive (for example, *ab initio* quantum scattering calculations have only been performed for reactions with six atoms or fewer) or fail to account for quantum tunneling effects, which can be large (for example, transition state theory). In between these two extremes lie semi-classical rate theories which aim to give results close to the full quantum result, with only slightly more expense than simple transition state theory. One key challenge is to model the potential energy surface of the reaction as accurately as possible, whilst requiring only a few *ab initio* calculations [1].

In this talk, we show how one such semi-classical method, Semi-Classical Transition State Theory (SCTST), can be applied to a key atmospheric process - the unimolecular decay of simple Criegee intermediates - to give accurate rate constants across a range of temperatures [2]. We also study the catalytic potential of a single water molecule on the reaction, and the effect of vibrational and rotational anharmonicity. We show that this process is significantly catalysed by atmospheric water molecules, and so the catalytic pathway should not be neglected in atmospheric models.

References:

- [1] S. Greene, X. Shan and D. C. Clary, *J. Chem. Phys.*, (2016), **144(24)**, 244116:
- [2] T. A. H. Burd, X. Shan and D. C. Clary, *PCCP*, (2018), Advance Article

Talk 8

Soot particle structure: Insights from molecular dynamics and Monte Carlo simulations

Kimberly Bowal^a, Jacob W. Martin^{a,b}, Peter Grančič^c, Markus Kraft^{a,b,d}

^aDepartment of Chemical Engineering and Biotechnology, University of Cambridge, West Site Philippa Fawcett Drive, Cambridge, CB3 0AS UK

^bCambridge Centre for Advanced Research and Education in Singapore (CARES), CREATE Tower, 1 Create Way, 138602 Singapore

^cInstitute of Soil Research, University of Natural Resources and Life Sciences, Peter-Jordan-Straße 82, 1190 Vienna

^dSchool of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, 637459 Singapore

Incomplete combustion produces carbonaceous particulate matter, known as soot, which negatively affects combustion devices, human health and the environment. The development of soot prediction models and mitigation strategies requires an accurate knowledge of soot particle formation and structure. Electron microscopy has revealed that mature soot particles often have a core-shell structure, in which the molecules near the centre of the particle differ in size from those in the surface layers [1]. However, it is unknown what interactions cause this partitioning and to what extent it is present in nascent particles.

In this work, the behaviour of heterogeneous polycyclic aromatic hydrocarbon (PAH) clusters was investigated using two independent molecular modelling methods to understand the structure of nascent soot particles. Clusters of up to 100 molecules containing equal proportions of two different sized PAHs, circumcoronene and coronene or ovalene and pyrene, were chosen as model particles. Replica exchange molecular dynamics (REMD) simulations sampled many configurations at high and low temperatures to determine stable low energy structures. These results were compared with a novel Sphere Encapsulated Monte Carlo (SEMC) method, developed here to extend basin-hopping minimisation methods for use with planar aromatic molecules.

Both the REMD and SEMC simulations showed that the most stable cluster structures consist of parallel columns of stacked PAHs in a core-shell structure, in which the larger PAHs are located in the core, and the smaller PAHs are found in a surrounding shell. This is an inverse partitioning compared to that seen in mature soot particles, proposing a unique structure for nascent soot particles and suggesting that the molecular organisation determined by intermolecular interactions is not responsible for the core-shell structure of mature soot. These results present the first simulations of soot-sized heterogeneous PAH clusters and can help further understanding of soot formation and growth processes.

[1] T. Ishiguro, Y. Takatori, and K. Akihama. *Combust. Flame*, (1997), **108**, 231–234.

Talk 9

Excited States of the Photosystem II Reaction Centre

Maeve Kavanagh, Laura Barter and Ian Gould

Department of Chemistry, Imperial College London, White City Campus, 80 Wood Lane, London W2 0BZ.

Photosystem II oxidises water during photosynthesis, via a powerfully oxidising, charge-separated state generated in its reaction centre. The reaction centre comprises 6 pigments, arranged in two branches and embedded in a protein matrix. Following initial excitation of the reaction centre, electron transfer occurs preferentially down one branch resulting in a highly-oxidising, radical pair state.

The excited states and subsequent electron-transfer intermediates of the reaction centre are difficult to study spectroscopically due to large spectral overlap. In order to try and overcome this, many models of the reaction centre have been made. However, existing models use large approximations for the reaction centre pigments, or have treated too small a region of the reaction centre, yielding an incomplete picture of the excited states.

We present the results of robust time-dependent density functional theory calculations on the largest QM model of the reaction centre to date. Our model was based on the crystal structure of *T. Vulcanus*¹ and long-range corrected functionals wB97X-D and CAM-b3LYP were used. Both wild-type and mutant reaction centres were modelled to test the validity of the method through comparison to mutant – wild-type difference spectroscopy.

We find that our model mimics the behaviour of the reaction centre in which charge transfer is directed preferentially down the ‘active’ branch. With our model we are able to observe excited states with charge-transfer character which match experimentally determined precursors to charge separation.² Trends in the experimental mutant – wild-type difference spectra are also successfully reproduced.

These results provide a benchmark for future models of the Photosystem II reaction centre. A better understanding of this natural process will aid in the design of future solar technology, and bio-engineered reaction centres.

References:

[1] Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. *Nature* **473**, 55–60 (2011).

[2] Romero, E., Novoderezhkin, V. I. & Van Grondelle, R. *Nature* **543**, 355–365 (2017).

Talk 10

Thermodynamic Integration Network Study of Electron Transfer: from Proteins to Aggregates ¹

Sehee Na ^a, Anna Bauß ^a, Michael Langenmaier ^b, Thorsten Koslowski ^a

a. Institut für Physikalische Chemie, Universität Freiburg, Albertstrasse 23a, D-79104, Freiburg, Germany

b. Institut für Anorganische und Analytische Chemie, Universität Freiburg, Albertstrasse 21, D-79104 Freiburg, Germany

We describe electron transfer through the NrfHA nitrite reductase heterodimer ² using a thermodynamic integration scheme based upon molecular dynamics simulations. From the simulation data, we estimate two of the characteristic energies of electron transfer, the thermodynamic driving forces, ΔG , and the reorganization energies, λ . Using a thermodynamic network analysis, the statistical accuracy of the ΔG values can be enhanced significantly: for example, in the case of the NrfA monomer the input data show an error of 4.9 kcal/mol⁻¹, whereas after the network analysis it is reduced to 1.0 kcal/mol⁻¹. This significantly improved statistical accuracy achieved by our network analysis allows us to examine subtle changes in energetic properties of electron transfer reactions within protein complex and to gain profound knowledge of the electron transfer in the multiheme case.

Although the reaction free energies and activation barriers are hardly affected by protein aggregation, the complete reaction mechanism only emerges from the simulations of the dimer rather than focussing on the individual protein chains: it involves an isoenergetic transprotein element of electron storage and conductivity. Furthermore, we suggest a barrier-wire(storage)-barrier model to explain the function of this large NrfHA nitrite reductase heterodimer regarding the energy landscape obtained by our simulation data.¹ Finally, our study exhibits how the thermodynamic integration methods combined with the network analysis can be applied for studies of electron transfer reactions within large protein systems such as NrfHA protein complex which has been considered hardly possible in the MD studies because of their system size.

References:

[1] S. Na, A. Bauß, M. Langenmaier, T. Koslowski, *PCCP*, (2017), **19**, 18938–18947.

[2] M. Rodrigues, K. Scott, M. Sansom, I. Pereira, M. Archer, *J. Mol. Biol.*, (2008), **381**, 341–350.

Talk 11

Revealing allosteric pathways with Information Theory

Lisa Patrick^a, Ben Cossins^b, Julien Michel^a

^aThe University of Edinburgh; Department of Chemistry; Kings Buildings, EH9 3FJ.

^bUCB; 216 Bath Rd, Slough, SL1 3WE.

Traditionally, drug design has focused on targeting the active site of a protein, usually with a small molecule, in order to inhibit or alter function for therapeutic effect. A promising alternative to this approach is to target an allosteric site: a binding site located distant to the active site, but where binding affects function at the active site. While some systems undergo fairly well understood structural changes, there is no overall model that satisfactorily describes how allostery works.¹

Molecular dynamics (MD) simulations provide a tool to study protein dynamics at the atomistic level, however traditionally employed analysis methods have proven inadequate to deliver a mechanistic description of allostery. In this work, we present our approach to tailor MD simulation analysis methods to identify motions which may be significant to allosteric signal transmission. These methods utilise two concepts from information theory: Kullback-Leibler (KL) divergence and Mutual Information (MI). KL-divergence allows structural descriptors for different simulations to be compared in order to highlight regions of the protein which change on binding of an allosteric ligand. In addition, determining which structural changes relate to changes in activity can be achieved by comparing different motions with interaction energies using MI, in order to determine if two variables are correlated.

Initial development of the methodology used phosphoinositide-dependent kinase-1 (PDK1) as test case.² Long MD trajectories were run for PDK1 in complex with covalent activator and inhibitor small molecules, using the software SOMD.³ Applying our approach, we were able to identify differences between the activated and inhibited simulations, and identify dominant motions leading to these structural changes. Subsequently, an energetic comparison was performed using a per-residue decomposition of the interaction energy between different components of the system. Calculating MI of these energies relative to structural features obtained from a principal component analysis, highlighted that the first principal component correlates with the interaction energy of ATP with the protein.

A second, more challenging test case was posed by protein-tyrosine phosphatase 1B (PTP1B) where steered-MD simulations were needed in order to investigate a key conformational change linked to the allosteric inhibition of this protein.^{4,5} Future work will include a “swarm of trajectories” approach, seeded from intermediate structures gathered during the steered-MD, in order to generate a Markov State Model description of the conformational changes involved. Overall we aim to validate a generally applicable toolkit for analysis of equilibrium and biased MD simulations to detect and characterise allosteric coupling in protein structures.

References:

- [1] R. Nussinov and C. J. Tsai, *Cell*, (2013), **153**, 293-305.
- [2] J. D. Sadowsky et al., *Proc. Natl. Acad. Sci.*, (2011), **108(15)**, 6056-6061.
- [3] C. J. Woods and J. Michel, *Sire/OpenMM*. (2014).
- [4] C. Wiesmann et al., *Nat. Struct. Mol. Biol.*, (2004), **11(8)**, 730-737.
- [5] D. A. Keedy et al., *eLife*, (2018), **7**, e36307.

Talk 12

Molecular Tetris: Human-guided exploration of a protein-ligand system in Virtual Reality

Helen M. Deeks^{a,b}, Rebecca K. Walters^{a,b}, Mike O'Connor^{a,b}, *David R. Glowacki^{a,b}, Adrian J. Mulholland^a*

^aDepartment of Chemistry, University of Bristol, Cantocks Close, Bristol, BS8 1TS.

^bDepartment of Computer Science, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol, BS8 1UB

Recent advances in virtual reality technology has allowed us to create and interact with fully three-dimensional representations of molecular dynamics simulations¹, an emerging use case for which is drug docking. Exploring protein-ligand systems using VR is akin to playing a game of 'molecular tetris'; like moving and rotating a tetris piece so that it fits into place, those trying to dock a small molecule into a protein must orientate the ligand to fit correctly in the active site.

Human spatial intuition has been successfully used to solve protein folding puzzles. Players of the game FoldIt discovered a folded protein configuration with higher activity than previously discovered², largely because players are able to traverse the high energy pathways that computational algorithms avoid. In a similar sense, human-guided docking allows key atomic contacts between the protein and ligand to be recreated, regardless of any associated energy penalties, allowing greater exploration of protein-ligand binding dynamics.

Nano Simbox allows users to interactively manipulate molecular dynamics simulations using commercial-grade virtual reality hardware. Recently, Nano Simbox has been developed to allow full molecular dynamics simulations of proteins and thus has been used to recreate binding poses for several targets.

References

- [1] M. O'Connor, H. M. Deeks, E. Dawn, O. Metatla, A. Roudaut, M. Sutton, L. M. Thomas, B. R. Glowacki, R. Sage, P. Tew, M. Wonnacott, P. Bates, A. J. Mulholland and D. R. Glowacki, *Sci Adv*, (2018) **4**, 6.
- [2] C. B. Eiben, J. B. Siegel, J. B. Bale, S. Cooper, F. Khatib, B. W. Shen, B. L. Stoddard, Z. Popovic and D. Baker, *Nature biotechnology*, (2012) **30**, 190.

Programme of Poster Presentations

Poster 1	Rosemond Mensah, School of Life and Medical Sciences, University of Hertfordshire Optimization of the process parameters for PLGA microparticle formulation based on Taguchi design
Poster 2	Shanlin Rao, Department of Biochemistry, University of Oxford Functional annotation of ion channel structures: Predicting pore hydration state based on local radius and hydrophobicity
Poster 3	David Ormrod Morley, Department of Chemistry, University of Oxford Computational Modelling of Two-Dimensional Atomic Networks
Poster 4	William Glass, Department of Biochemistry, University of Oxford Multiscale molecular dynamics to explore voltage-gated sodium channel oligomerisation
Poster 5	Joshua Carter, Nuffield Department of Medicine, University of Oxford <i>De Novo</i> Prediction of Pyrazinamide Resistance in <i>Mycobacterium tuberculosis</i> from Protein Structure
Poster 6	Joe Bluck, Department of Biochemistry, University of Oxford In silico free energy calculations to guide bromodomain inhibitor optimisation
Poster 7	Irene Marzuoli, Randall Division & Physics Department, King's College London Multi scale simulations of de novo designed self-assembling peptides for antimicrobial and gene-delivery strategies
Poster 8	Silvia Amabilino, Department of Chemistry, University of Bristol & NovaData Solutions The <i>in silico</i> Medicinal Chemist: Neural Networks for Kinase Inhibitor Design
Poster 9	Rowan Howell, The Francis Crick Institute & Randall Division of Cell and Molecular Biophysics, King's College London – <i>WITHDRAWN</i> Training a Boolean model of the Mitotic Exit Network
Poster 10	Sam Johns, School of Biochemistry, University of Bristol Simulation-led redesign and evaluation of a natural Diels-Alderase
Poster 11	Alex Wade, Department of Physics, University of Cambridge Assimilating Radial Distribution Functions to Build Water Models with Improved Structural Properties
Poster 12	Louis Evans, School of Science, University of Greenwich Probing the Reaction Mechanism of the Prins Cyclisation via a Combined Computational and Experimental Chemistry Study

Poster presenters will give their 2 minute “Lightning Talks” in the above order starting at 11:00.

Posters will be on display during the breaks and at lunchtime.

Poster 1

Optimization of the process parameters for PLGA microparticle formulation based on Taguchi design

Rosemond A. Mensah, Stewart B. Kirton, Michael T. Cook, Victoria Hutter, David Y.S. Chau

Department of Clinical and Pharmaceutical Sciences, School of Life and Medical Sciences, University of Hertfordshire, AL10 9AB, UK

Poly lactic-co-glycolic acid (PLGA) is a synthetic polymer that has been widely studied and used in the formulation of microparticles (MP) for the controlled and sustained release of drugs. The single emulsion oil-in-water (o/w) evaporation technique is often used for the preparation of MP, but with many processing variables present, it is difficult to predict the size of particles formed¹. Particle size is a significant characteristic as it determines the profile of the drug release and dictates the potential routes of administration (e.g. topical or injection). Moreover, particle size is important with respect to tissue irritation. Taguchi design is a statistical technique used to optimize parameters with complex interrelationship and its design protocol attempts to identify controllable factors that minimize the effect of “noise” (i.e. uncontrollable factors). Taguchi design exploits the use of orthogonal arrays (OA) to analyze many parameters with few experiments². The objective of this study was to employ Taguchi Level 12 and Level 18 OA designs to explore the effect of processing parameters on particle size using the single o/w evaporation technique. The parameters explored were: concentration of poly(vinyl alcohol) (PVA), molecular weight of PVA, concentration of PLGA, type of solvent (i.e. dichloromethane, ethyl acetate), concentration of PVA in the hardening bath, stirring/mixing speed, ratio of solvent/aqueous phases, emulsification rate, duration of speed and time for evaporation. The Taguchi OA allows for a model to be developed to identify conditions leading to 10-50 µm MP for topical drug delivery.

PLGA microparticles were successfully fabricated using an established solvent evaporation technique. The median particle diameters were measured using a SympaTec laser diffraction particle size analyser (SympaTec GmbH, Germany) and the morphology of the PLGA MPs were visualised using scanning electron microscopy. The first optimization step using the L12 design showed that all parameters significantly influenced the particle size of the fabricated MP with exception of concentration of PVA in hardening bath. In contrast, the L18 design results showed that molecular weight of PLGA does not significantly affect the particle size. The results from this study showed that application of Taguchi design can be used to understand the influence of parameters and predict the best combination of process parameters that can provide the optimal response condition- in this case, to formulate 10-50 µm MP for topical application.

References:

- [1] E. Bible, D.Y.S Chau, M.R. Alexander, J. Price, K.M. Shakesheff and M. Modo, *Nat Protoc.*, (2009) **4**, 1440-1453.
- [2] D.C. Montgomery, *J. Royal Stat. Soc.*, (1999) **48**, 159-177.

Poster 2

Functional annotation of ion channel structures: Predicting pore hydration state based on local radius and hydrophobicity

Shanlin Rao^a, Gianni Klesse^{a,b}, Phillip J. Stansfeld^a, *Stephen J. Tucker*^b & *Mark S. P. Sansom*^a

^a Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU;

^b Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, OX1 3PU

Ion channel proteins function as water-filled, ion-conducting pores that constitute crucial components of biological membranes. Ionic flow through these proteins in their resting or inactivated state may be interrupted at one or more positions ('gates') along the channel; conformational changes upon receipt of an activation signal typically lead to widening of the channel pore at its gate region(s). Driven by pharmacological interests in elucidating mechanisms of different ion channel families, and with advances in cryo-electron microscopy, an increasing number of ion channel structures are being determined, often capturing alternative conformations of the transmembrane pore.

Annotation of the conductive state represented by a newly resolved ion channel structure (i.e. whether its conformation would allow permeation of water and ions) forms a key aspect towards understanding its functional relevance and requires consideration of physicochemical properties of the pore lining. A hydrophobic region may disfavour liquid-phase water and undergo spontaneous capillary evaporation, thereby presenting an energetic barrier to the flow of water and ions without steric occlusion of the pore. The critical radii above which a channel pore may become stably hydrated is modulated by the degree of hydrophobicity of the pore lining, as initially demonstrated by simulation studies on model nanopores¹.

Here we apply molecular dynamics simulation-based calculations (www.channotation.org)² to 190 ion channel structures. Through water density and free energy estimations, we examine the influence of local radius and hydrophobicity on channel pore hydration. The results of our analysis provide a heuristic model for predicting the solvation state of new channel structures as they emerge.

References:

- [1] O. Beckstein and M. S. P. Sansom, *Phys. Biol.*, 2004, **1**, 42–52.
- [2] J. L. Trick, S. Chelvaniththilan, G. Klesse, P. Aryal, E. J. Wallace, S. J. Tucker and M. S. P. Sansom, *Structure*, 2016, **24**, 2207–2216.

Poster 3

Computational Modelling of Two-Dimensional Atomic Networks

David Ormrod Morley, Mark Wilson

Department of Chemistry, University of Oxford, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford, OX1 3QZ

From armadillo shells to bubble baths to the stones on Giant's Causeway, two-dimensional networks are commonplace throughout nature. It is no surprise then that they are also found in Chemistry, and of particular interest owing to the recent experimental synthesis and visualisation of materials such as amorphous graphene and silica bilayers. These materials have the potential to form a range of technologically useful materials and prove an interesting problem for theoretical and computational study owing to their low dimensionality.

Modelling two-dimensional networks and their ring structure presents a unique set of challenges. Euler's law tells us that the mean ring size must be six, and the full ring distribution may be known from experiment, however this is not sufficient to fully describe the network topology. The ring-ring correlations must also be accounted for through the Aboav-Weaire law¹. We have developed a method to generate samples in which both the ring distribution and ring-ring correlations are imposed on the network, by exploiting relationships between the atomic network and its *dual graph*. This method takes inspiration from "bond-switching" algorithms, where crystalline samples are amorphised to a target structure, but with Monte Carlo moves instead performed on the dual graph - before the results are translated into real space. This has allowed us to explore the why certain naturally occurring structures are energetically favourable, and provide physical insight into the meaning of the Aboav-Weaire parameter.

Alternative Monte Carlo methods have been used to grow networks from seed in order to mirror experimental data. This allows generation of configurations which are larger than those available from experiment and which are artefact free. In addition, network properties can be tuned to make structures which are as yet experimentally unrealisable. These samples, and those from experiment, are necessarily aperiodic and so are unsuitable for conventional molecular dynamics techniques. However, we are currently working on an implementation of novel "sliding boundary conditions", which can allow aperiodic two-dimensional networks to be investigated using molecular dynamics simulations².

References:

[1] S. N. Chiu, *Mater. Charact.*, (1995), **34**, 149-165.

[2] L. Theran, A. Nixon, E. Ross *et. al.*, *Phys. Rev. E*, (2015), **92**, 053306.

Poster 4

Multiscale molecular dynamics to explore voltage-gated sodium channel oligomerisation

William Glass¹ and *Philip C. Biggin*¹

¹Department of Biochemistry, University of Oxford, Oxford, UK

william.glass@chem.ox.ac.uk

Electrical signalling is key to a vast number of biological processes within the nervous system and for this to occur specific types of transmembrane proteins, named voltage-gated ion channels, are utilised¹. Voltage-gated sodium ion (Na_v) channels are vital to the correct function of nerve cells for the effective transmission of an action potential.

Typically, Na_v channels are not present in isolation and are known to be associated with one or more β-subunits². These transmembrane proteins are multifunctional and interact with the main α-subunit of the Na_v channel in addition to having a cell adhesion function through their extracellular immunoglobulin domain. In Na_{v1.5} the β3-subunit is known to noncovalently interact with the central α-subunit³, however the exact mode of interaction is unclear. Additionally, β3 has been found to form trimers within the membrane through interactions between the extracellular Ig domains as well as cross-linking of α-subunits^{4,5}.

To gain an insight into the α – β subunit relationship both atomistic and coarse grained molecular dynamics methodologies are adopted. Longer time scales granted through the latter technique allow the reported trimerisation and higher order oligomers to be studied and the likely interactions and dependence on domains assessed. Increasingly large-scale simulations are beginning to elucidate the relationship between the β3-subunit and the central α-subunit and increase our understanding of ion channel modulation.

References

1. A. L. Hodgkin and A. F. Huxley, *J Physiol*, **117**, 500-544, **1952**
2. Das et al., *eLife*, **5**, e10960, **2016**
3. Brackenbury & Isom, *Front Pharmacol*, **2**, **2011**
4. Namadurai, S., et al., *Open Biol*, **5**, **2015**
5. Namadurai, S. et al. *J. Biol. Chem.* **289**, 10797–10811 **2014**.

Poster 5

De Novo Prediction of Pyrazinamide Resistance in *Mycobacterium tuberculosis* from Protein Structure

Joshua J Carter¹, Timothy M Walker¹, A Sarah Walker^{1,2}, Timothy EA Peto^{1,2}, Derrick W Crook^{1,2,3}, Philip W Fowler^{1,2}

¹ Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK

² National Institute of Health Research Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK

³ Public Health England

In 2016, tuberculosis (TB) infected 10.4 million and killed 1.7 million people, making it the leading cause of death by infectious disease. Pyrazinamide is a critical component of both first-line and multiple drug resistant (MDR) TB treatment regimens; it acts on non-growing “persister” bacteria, shortening treatment times and increasing their sterilizing effects. However, the evolution of pyrazinamide resistance has been implicated in poor treatment outcomes. Pyrazinamide is a prodrug that is converted to its active form by an enzyme called PncA, a nonessential protein that contains mutations in the majority (72-99%) of pyrazinamide-resistant isolates. Advances in next-generation sequencing and PCR-based line probe assays have enabled the detection of emerging resistance from genotype more quickly than from standard drug susceptibility testing. However, resistance-causing mutations have been found dispersed across the promoter and open reading frame of the *pncA* gene, complicating the development of a complete catalogue of resistant mutations. A method to predict resistance for all possible nonsynonymous mutations in PncA would increase our ability to detect and respond to emerging pyrazinamide resistance.

This study exploits the availability of (i) a high-resolution crystal structure of PncA from *M. tuberculosis* and (ii) a large and well-characterized mutagenesis dataset (322 unique amino acid mutations) [1] in order to estimate sequence-based structural properties that machine-learning models can exploit to predict the effect of each individual PncA mutation on pyrazinamide sensitivity. Applying the model to a novel set of *pncA* mutations collected from a recent study of over 10,000 clinical isolates (163 unique amino acid mutations) [2] demonstrates that we can achieve 86% sensitivity and 99% specificity in identifying PncA mutations associated with phenotypic pyrazinamide sensitivity. Alongside recent advances in genetics-based diagnostic technologies, this technique provides proof-of-concept for allowing *de novo* identification of emerging pyrazinamide resistance, enabling more rapid clinical responses. Importantly, this model is able to make a prediction for every nonsynonymous mutation possible in PncA, enabling its extension to novel SNPs detected in emerging outbreaks. Future work could extend this analysis to predict resistance in other prodrug antibiotics such as ethionamide or bedaquiline.

Poster 6

In silico free energy calculations to guide bromodomain inhibitor optimisation

Joseph P. Bluck^{1,2}, Larissa See², Amy R. Scolah², Jessica R. Reynolds², Angelina Sekirinik², Francesco Rianjongdee³, Garrett M. Morris⁴, Stuart J. Conway², Philip C. Biggin¹

¹Department of Biochemistry; ²Department of Chemistry;

³Epigenetics Discovery Performance Unit, GlaxoSmithKline R&D, Stevenage;

⁴Department of Statistics

Development of ligands for bromodomains (BRDs) has expanded our understanding of gene regulation and epigenetic links to disease. Though many potent probes have been developed for some of the BRDs, many BRDs are yet to have a selective probe. This is partially due to a lack of polar residues in proximity to the binding site of some BRDs, along with difficulties in gaining selectivity over other BRDs.

Several compounds that bind to the BRD of TRIM33 were identified through a high throughput screen, one of which has been selected for optimisation. Without a ligand co-crystal structure, docking studies were employed to enable structure-based drug optimisation. However, synthesis of compound analogues gave surprising structure-activity relationships (SAR) that did not fit the proposed binding model. The binding model was subsequently scrutinised through molecular dynamics (MD), which highlighted alternative binding modes.

The MD based binding model was further improved through rigorously calculating the absolute binding free energies of the ligands, to increase confidence in ranking binding modes. Following our previous work, the energies of binding-site water molecules were probed with Grand Canonical Monte Carlo (GCMC) calculations.¹ Combining both aspects has led to an improved binding model that fits the experimental SAR and is now being used to yield rational designs for future compounds.

1. Aldeghi, M. *et al.* Large-scale analysis of water stability in bromodomain binding pockets with grand canonical Monte Carlo. *Commun. Chem.* **1**, 19 (2018).

Poster 7

Multi scale simulations of de novo designed self-assembling peptides for antimicrobial and gene-delivery strategies

Irene Marzuoli^a, Dr. Chris D. Lorenz^a, Dr. Max Ryadnov^b, Prof. Franca Fraternali^a

^a Randall Division & Physics Department, King's College London, New Hunt's House, Guy's Campus, SE1 1UL

^b Biotechnology National Physical Laboratory, Teddington

Antimicrobial resistance and drug delivery have been two main focuses of medical research in the last decades. Recently engineered virus-like peptidic capsids¹ have been shown to promote bacterial membrane poration and, at the same time, to be suitable for gene delivery, as they co-assemble with siRNA and promote its transfection to human cells. However, the atomistic details of both their antimicrobial activity and their assembly, which is necessary for their functions, are still unknown.

The project aims at studying through Molecular Dynamics simulations the two processes, comparing them with experimental data available through the partnership with the National Physical Laboratories. The *in-silico* investigation proceeded from the simulation at atomistic level of small assemblies in solution to the modelling of increasingly complex ones, with both atomistic and coarse grain descriptions. The structures found to be stable were selected for further simulations of their interaction with a bacterial model membrane. The protein-membrane interaction has been enhanced via electroporation simulations to observe the full poration process within the accessible simulation time. An accurate validation of the membrane force field parameters has been undertaken to assess the performance and accuracy of the newly available ones.

Integrated results from the different sets of simulations highlighted the role of the amphiphilic structure of the molecule as driven promoter of the structure stability, while its interaction with the membrane has been investigated under applied electric field showing that the protein has a stiffening effect to the membrane which ultimately leads to poration.

This investigation points to some major determinants in the antimicrobial activity effectiveness of the molecule, showing that computational techniques can contribute to the molecular understanding of its functional activity. Future developments will include an exploration of amino acids mutations to test and select sequences that can more effectively display such antimicrobial and delivery double task.

1. V. Castelletto et al., *Chem. Sci.*, (2016) **7**(3), 1707–1711.

Poster 8

The *in silico* Medicinal Chemist: Neural Networks for Kinase Inhibitor Design

Silvia Amabilino ^{a, b}, David R. Glowacki ^a, Michael P. Mazanetz ^b

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

^bNovaData Solutions Ltd, PO Box 639, Abingdon-on-Thames, Oxfordshire, OX14 9JD

There are currently over 40 FDA approved Kinase inhibitors [1]. However, not all of these drugs are single target kinase inhibitors and many treatments have undesirable side effects. It is therefore essential to design new chemical entities with cleaner profiles. Standard techniques for this purpose include high throughput screening, structure-based design and analogue synthesis. We propose an alternate approach, an *in silico* Medicinal Chemist, that considers the wealth of publicly available medicinal chemistry knowledge in the design process.

In this work, we applied generative recurrent neural networks (GRNN) in the design of new kinase inhibitors [2]. GRNNs are able to learn the syntax of SMILES strings and they can be fine-tuned to predict specific molecular targets. This makes them a powerful technique to design completely new molecules that bear similarities to existing ones.

The model, our *in silico* Medicinal Chemist, was first trained on a large subset of ChEMBL to learn the syntax of SMILES strings. Once the model was able to predict valid SMILES strings, it was then fine-tuned on a much smaller data set of kinase inhibitors.

The molecules predicted by the fine-tuned model were highly similar to the existing inhibitors, without being duplicates. In order to compare the predicted molecules and the original data set, over 50 medicinal chemistry-favoured descriptors were evaluated. A principal component analysis showed a good overlap of the predicted kinases and the training set, suggesting that the predicted molecules have the potential of being good candidates as kinase inhibitors. Compounds are currently being prioritised for chemical synthesis and biological evaluation.

References:

[1] P Wu, T. E. Nielsen, M. H. Clausen, *Trends in Pharmacological Sciences*, (2015), **36** (7), 422-439

[2] A. Gupta, A. T. Muller, B. J. H. Huisman, J. A. Fuchs, P. Schneider and G. Schneider, *Mol. Inf.*, (2018), **37**, 1700111.

Poster 9

Training a Boolean model of the Mitotic Exit Network

Rowan Howell, Peter Thorpe, Attila Csikasz-Nagy

The Francis Crick Institute, London; and Randall Division of cell and molecular biophysics, King's College London

The Mitotic Exit Network (MEN) is a collection of proteins that coordinate late mitotic events in *S. cerevisiae* [1]. MEN proteins localize together at the Spindle Pole Body (SPB) and there is a complex interplay between function and localization. The complexity is increased through cross-talk with the cdcFourteen Early Anaphase Release (FEAR) pathway. The study of mitotic exit is of interest as errors in control of mitotic exit can lead to aneuploidy and cancer.

We collected roughly 200 genetic phenotypes resulting from knock-down or overexpression of MEN and FEAR proteins from the literature. We used this dataset to train a Boolean model of the MEN using the CellNOptR tool [2]. The CellNOptR tool uses a genetic algorithm to find optimal wirings of the network that achieve a good match between the model's behaviour and the data.

This model provides a formal representation of our understanding of the MEN by identifying a minimum set of interactions required to explain current observations on the MEN. Additionally, it can be used to make testable predictions about complex genetic interactions. In the future, we aim to expand the model to account for the effects of localization using a systematic dataset on forced relocalization of MEN components.

References:

- [1] Scarfone, I., & Piatti, S., *Small GTPases* (2015), **6(4)**, 196–201.
- [2] Terfve, C., et al., *BMC Systems Biology*, (2012). **6(1)**, 133-147.

Poster 10

Simulation-led redesign and evaluation of a natural Diels-Alderase

Sam Johns, Dr Marc Van der Kamp, Dr Paul Race

School of Biochemistry, University of Bristol

Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, United Kingdom

The Diels-Alder reaction, which involves the [4+2] cycloaddition of a conjugated diene to a dienophile, is a key reaction in organic synthesis and as such there is a strong demand for enzymes that can catalyse this reaction. Recently, the enzyme AbyU was identified as a true natural Diels-Alderase¹ that is attractive as a biocatalyst for industry. Due to its simplicity and stand-alone nature, it is also a practical target for engineering. We thus aim to explore the potential for broadening the substrate scope of this enzyme by evaluating variants that are predicted by computer simulation to produce the desired change in substrate scope. An *in silico* simulation workflow of protocols including molecular docking, molecular dynamics simulation and combined quantum mechanics / molecular mechanics (QM/MM) reaction simulation² is being developed to enable high-throughput screening of proposed enzyme variants with alternative substrates in order to identify promising candidates. The elements of the workflow are explained and illustrated with initial results.

References:

[1] M.J. Byrne, N.R. Lees, L.C. Han, M.W. Van der Kamp, A.J. Mulholland, J.E. Stach, C.L. Willis and P.R. Race, *JACS*, (2016), **138**, 6095

[2] M.W. Van der Kamp and A.J. Mulholland, *Biochemistry*, (2013), **52**, 2708

Poster 11

Assimilating Radial Distribution Functions to Build Water Models with Improved Structural Properties

Alexander D. Wade^a, David J. Huggins^a, Lee-Ping Wang^b

^aUniversity of Cambridge, TCM Group, Cavendish Laboratory, 19 JJ Thomson Avenue, Cambridge CB3 0HE, United Kingdom

^bUC Davis, Department of Chemistry, One Shields Ave, Davis, CA 95616, United States

Understanding the physical behaviour of water is of critical importance to an accurate description of biological and engineered systems. Despite the simplicity of water as a individual molecule its structure and dynamics in bulk liquid remains a major scientific challenge to be solved. A central goal of computational bio-chemistry is to develop classical force fields which are more accurate than the standard of chemical accuracy 1 kcal/mol. Within this limit reasonable chemical predictions can be made and computational methods could be used to cheaply design or screen new drugs. Providing an valuable tool within the pipeline of rational drug design.

The reward for developing accurate force fields is high, however the difficulty of parametrizing classical force fields which are accurate over a range conditions is equally high, often described as a 'black art'¹. As such many parametrizations of water exist and all these parametrisations fit to a different set of target properties. During fitting, these parametrisations rarely, if ever², include the entropy. Which is a crucial component of a system's free energy.

This work presents a method to target the two body translation entropy of a material during the systematic parametrization of its force field. This is achieved by allowing experimental radial distribution functions (RDFs) to be included in the set of target properties. RDFs are a measure of the structure and correlation of the material. Therefore by tuning the parameters of the force field, until the computational RDF is best agreed with experiment, we improve the accuracy of the models structural properties and as a result the entropy.

Applying this methodology to parametrise water yields models with improved calculations of entropy. The error in the calculation of the entropy is reduced from 11% in the most popular water models (TIP3P, TIP4P/2005) to 2% in the new models presented here (TIP3P-Buckingham). Additionally these new models have improved predictive power for many non-fitted properties.

[1] L-P. Wang, T, Martinez, V.S. Pande, *J. Phys. Chem. Lett.* (2014), **11**, 1885-1891

[2] A.V. Onufriev and I. Saeed, *Wiley Interdiscip Rev Comput Mol Sci*, (2018) **8**, 1347-1387

Poster 12

Probing the Reaction Mechanism of the Prins Cyclisation via a Combined Computational and Experimental Chemistry Study

Louis Evans, *Adrian Dobbs*, *Jiayun Pang*

School of Science, University of Greenwich, Central Avenue, Chatham Maritime, Kent.

The Prins cyclisation reaction is a highly powerful reaction used in the preparation of a wide range of heterocyclic structures.¹ Despite finding many applications in natural product and pharmaceutical synthesis², the exact reaction mechanism of the Prins reaction is still disputed, with three main approaches proposed in the literature – direct cyclisation (Dobbs), allyl-transfer *via* [3,3]-sigmatropic rearrangement (Willis) and aldehyde incorporation (Maseras). A detailed understanding of the exact reaction mechanism is highly sought after as it may allow for further developments of the Prins cyclisation and unlock further synthetic potential, especially with regard to currently difficult-to-synthesise heterocyclic structures.

Through the use of computational methods (density functional theory calculations performed with Gaussian) and detailed experimental observations, all three key routes offered in the literature have been studied and compared for the first time. It is shown that the reaction proceeds *via* oxocarbenium ion intermediates, whereby direct cyclisation is most favoured, albeit with allyl-transfer also taking place as a high-energy competing route. The role of the Lewis acid has also been investigated, showing an intimate relationship throughout the reaction that differs to the existing proposed mechanisms.

This work will be discussed in the presentation, showing how the structural geometry, calculated energies and experimental results have led to a much more detailed reaction mechanism for the Prins cyclisation. In addition, preliminary mechanistic results of our investigation in to the silyl-Prins reaction, a useful variation leading to dihydropyran products will also be shown.

References:

- (1) For a review of the Prins cyclisation reactions see: C. Olier, M. Kaafarani, S. Gastaldi and M. P. Bertrand, *Tetrahedron*, 2010, **66**, 413-445
- (2) M. D. Delost, D. T. Smith, B. J., Anderson and J. Y. Njardardon, *J. Med. Chem.*, 2018, Article ASAP