



MGMS and RSC MMG Young Modellers' Forum 2012 PROGRAMME & ABSTRACTS

Programme of Oral Presentations

9.00 – 9.30	Coffee and Registration
9.30 – 9.40	Welcome and Introduction
9.40 – 10.00	<i>In Silico</i> Prediction of Aqueous Solubility using Linear and Non-Linear Methods of Regression Analysis Jogoth Ali, <i>University of Hertfordshire</i>
10.00 – 10.20	Clustering and 'protein-based pharmacophore' modeling of protein binding-sites Lydia Siragusa, <i>University of Perugia, Italy</i>
10.20 – 10.40	Interplay of theory and spectroscopy: Study of an Fe^V-nitride complex and its photolytic formation Oliver Krahe, <i>Max-Planck Institute, Germany</i>
10.40 – 11.00	Modelling the Incorporation of Biomolecules into Calcite David Sparks, <i>University of Sheffield</i>
11.00 – 11.20	Predicting the morphology of organic solids Dinesh-Ramesh Mirpuri Vatvani, <i>University of Cambridge (CCDC)</i>
11.20 – 12.00	Lightning Talks All poster presenters, in poster number order
12.00 – 13.30	Lunch and Poster Session
13.30 – 13.50	<i>Ab initio</i> DFT-assisted structural characterisation of MOFs Jason Loader, <i>University of Sheffield</i>
13.50 – 14.10	The Link Between Macbeth's Banquo and Molecular Aromaticity Kate Horner, <i>University of York,</i>
14.10 – 14.30	The Statistical Limits of Molecular Scoring Gregory Ross, <i>University of Oxford</i>
14.30 – 15.00	Tea
15.00 – 15.20	Coarse-grained molecular dynamics simulations of the interactions of two cellpenetrating peptides with complex membrane models Jean Hélie, <i>University of Oxford</i>
15.20 – 15.40	Structural and thermodynamic insights into bacterial outer membrane lipid signaling by the innate immune system Teresa Paramo, <i>University of Cambridge</i>
15.40 – 16.00	Catalytic activity of fatty acid amide hydrolase investigated by EVB free energy calculations Ewa Chudyk, <i>University of Bristol</i>
16.00 – 16.20	Computer Simulations of the Molecular Role of EGFR Kinase Mutations Christopher Bull, <i>University of Southampton</i>
16.20 – 16.45	Judges Deliberations/Networking Session
16.45 – 17.00	Prize Presentations
17.00	End

Talk 1

***In Silico* Prediction of Aqueous Solubility using Linear and Non-Linear Methods of Regression Analysis**

Jogoth Ali^a, Patrick Camilleri^b, Marc B Brown^{a, c}, Andrew J. Hutt^a and Stewart B. Kirton^a

^a School of Pharmacy, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK

^b Bio-Chemical Solutions, 5 Morgan Close, Stevenage, Hertfordshire, SG1 4TG, UK

^c MedPharm Ltd, Unit 3, Chancellor Court, 50 Occam Road, Guildford, Surrey, GU2 7YN, UK

In silico models for predicting the aqueous solubility (logS) of compounds are widely used in fields including medicinal, physical, and environmental sciences. Here we present work where we have utilised numerous *in silico* techniques to develop accurate and predictive models of aqueous solubility.

Simple quantitative structure-property relationship (QSPR) models were created using partial least squares (PLS) regression based on the logP, melting point and topographical polar surface area (TPSA) of a chemical substance. A three descriptor model [1] accurately predicted (i.e. to within 1 log unit of the experimental logS) approximately 87% of compounds in a dataset of 1256 diverse chemical structures. These results are comparable to established models of aqueous solubility e.g. ESOL [2] and the GSE [3]. Hierarchical clustering according to chemical similarity highlighted molecules with phenolic and/or phenol-like moieties as being poorly predicted by all models. Incorporating a fourth descriptor pertaining to a simple count of aromatic hydroxyl groups improved the predictive ability to 89% for the dataset.

Following from this, artificial neural networks (ANN) were used to investigate non-linear relationships between chemical descriptors and aqueous solubility. Backpropagation networks (BPN) employing supervised error correction learning to train multilayer perceptron (MLP) networks for improved prediction of aqueous solubility. The orthogonal array testing strategy (a 'design of experiments' technique) was employed to guide and assess the factors involved in determining optimal ANN architectures. Numerous network architectures were trained and evaluated. The best performing four descriptor MLP network (4-4-10-1) showed superior predictive abilities in comparison to the best-performing linear PLS regression model with 92% of compounds being accurately predicted.

References:

- [1] Ali, J.; Camilleri, P.; Brown, M. B.; Hutt, A. J.; Kirton, S. B., Revisiting the General Solubility Equation: *In Silico* Prediction of Aqueous Solubility Incorporating the Effect of Topographical Polar Surface Area. *J. Chem. Inf. Model.* (2012), **52**(2), 420-428.
- [2] Delaney, J. S., ESOL: Estimating aqueous solubility directly from molecular structure. *J. Chem. Inf. Comput. Sci.* (2004), **44**(3), 1000-1005.
- [3] Jain, N. and Yalkowsky, S.H., Estimation of the aqueous solubility I: Application to organic nonelectrolytes. *J Pharm Sci.* (2001). **90**(2), 234-252.

Talk 2

Clustering and 'protein-based pharmacophore' modeling of protein binding-sites

Lydia Siragusa,^a Massimo Baroni,^b Valerio Ferrario,^c Cynthia Ebert,^c Lucia Gardossi,^c Simon Cross^b

^a Chemistry Department, University of Perugia, Via Elce di sotto 8, 06123 Perugia, Italy

^b *Molecular Discovery Limited, 215 Marsh Road, Pinner, Middlesex, London HA5 5NE, United Kingdom*

^c *Chemical and Pharmaceutical Sciences Department, University of Trieste, P.le Europa 1, Trieste, Italy*

Nowadays, the understanding of the key functional features of proteins is a fundamental requirement in protein molecular modelling. A new computational algorithm, for analysing similarities and differences across protein families has been developed. The algorithm was designed primarily for protein binding site comparison via three-dimensional superpositioning. This approach is based on the main concept that the function of a protein does not necessarily depend on its folding or sequence, but is related to essential interactions established by the three-dimensional spatial arrangement of protein residues in the binding sites.

The new procedure uses Molecular Interaction Fields (MIFs) generated with the GRID force-field in order to evaluate the type, strength and direction of the interactions that a binding site is capable of having. In order to compare protein cavities, it generates a Common Reference Framework which is applied for binding site comparison.¹

The method has been used in various scenarios: (i) multivariate statistical analysis to cluster functionally related proteins based on their GRID MIF similarities. We analysed a dataset of protein active sites belonging to MAP kinases in distinct subfamilies (p38 α , JNK, ERK2). Principal Component Analysis clustering highlighted the essential features for the MAP kinase inhibitors and helped to explain the ligand selectivity; (ii) generation of ‘protein-based pharmacophore’ models, derived from the MIFs that are common across several protein targets. Unlike ‘ligand-based pharmacophore’ models, which require the derivation of the bioactive conformation of the active ligands, the ‘protein-based pharmacophore’ models are independent from any ligand information.² We performed a superimposition of a set of two serine-hydrolases subfamilies (protease/esterase) and generated a ‘protein-based pharmacophore’ model for each. The common pharmacophoric fields and the electrostatic and physico-chemical differences enabled us to rationalise the catalytic differentiation for each subfamily.

In all cases the results validated the effectiveness of our method and illustrated its potential application, both for molecular drug-design and bio-catalysis. Furthermore, the new method can be exploited for several other applications, including off-targeting, protein engineering, inferring a function of a novel protein, multi-target therapy, and de novo drug design.

Part of this work (L.G.) was financially supported by the EU-FP7 “IRENE” project under the FP7-KBBE-2008-2B grant agreement n° 227279

References:

[1] Baroni, M., Cruciani, G., Sciabola, S., Perruccio, F. and Mason, J., A Common Reference Framework for Analyzing/Comparing Proteins and Ligands. Fingerprints for Ligands And Proteins (FLAP): Theory and Application, *J. Chem. Inf. Model.*, (2007) **47**, 279-294

[2] Cross, S.; Baroni, M.; Goracci, L.; Cruciani, G. GRID-based 3D pharmacophores I: FLAPpharm, a novel approach for pharmacophore elucidation, *J. Chem. Inf. Model.*, (2012) In Press. (DOI: 10.1021/ci300153d)

Talk 3

Interplay of theory and spectroscopy: Study of an Fe^V-nitride complex and its photolytic formation

Oliver Krahe, Eckhard Bill, Taras Petrenko, Frank Neese

Max-Planck Institute for Chemical Energy Conversion, Mülheim an der Ruhr, Germany

In bioinorganic chemistry a wide range of highly developed spectroscopic methods have been used in the study of bio-relevant complexes and over the last two decades, computational chemistry has gained prominence as a powerful tool to provide insight into e.g. bio-relevant reactions from an additional perspective. Both fields, spectroscopy and theory, are individually very powerful, but a combination of both can facilitate the interpretation of data and permits an even more detailed understanding of the studied systems.

The nitridoiron model system I focus on in my PhD work is a six coordinated Fe^V-nitride supported by a cyclam derived ligand which is formed by photolysis of its Fe^{III}-azide counterpart[1]. Combining spectroscopy and theory we have been investigating the formation process of Fe^V by N₂ elimination and the electronic structure of the resulting high-valent iron complex. Using DFT we are able to accurately model the measured Mößbauer parameters, but a more sophisticated insight in the electronic structure was obtained by multi-reference calculations (CASSCF/NEVPT2), which also made it possible to reproduce the measured g-values accurately. To map the processes that accompany photo excitation, resonance Raman spectroscopy is a very applicable method, with DFT once again very useful in band assignment.

References:

- [1] Grapperhaus, G., Mienert, B., Bill, E., Weyhermüller, T., Wieghardt, K., *Inorg. Chem.*, (2000), **39**, 5306-5317.
[2] Aliaga-Alicante, N. DeBeer, S., Mienert, B., Wieghardt, K., Neese, F., *Angew. Chem. Int.Ed.* (2005), **44**, 2968-2972.

Talk 4

Modelling the Incorporation of Biomolecules into Calcite

David J. Sparks, Colin L. Freeman, John H. Harding

Department of Materials Science and Engineering, Sir Robert Hadfield Building, University of Sheffield, S1 3DJ.

Biomaterial composite materials exhibit mechanical properties that are superior than their fully inorganic counterpart. Recent synthetic work has found that the incorporation of biomolecules such as polymers or amino acids within calcite crystals improves the mechanical properties, including the fracture toughness and hardness, of these composite structures without considerable disruption of the crystal lattice [1,2]. Recent work on the screening of the incorporation of all 20 common amino acids using high-resolution powder diffraction has shown a varying degree of incorporation of these amino acids [3].

Molecular Dynamics simulations have been performed to model the incorporation of these small biomolecules within the calcite structure. The amino acids Aspartic acid (Asp) and Glycine (Gly) have been selected as their behaviour within the calcite structure is seen to be distinctly different. A range of conformations with an increasing weight percent amino acid has been studied. From the insertion energies and radial distribution data calculated it is apparent that Aspartic Acid is more readily included into the calcite crystal, with a minimum disruption to the crystal lattice. The inclusion of Glycine however, is seen to be energetically less favourable and a larger disruption of the crystal lattice is observed upon the insertion of Glycine. The results obtained with these Molecular Dynamics simulations are in good agreement with experimental results.

References:

[1] Kim, Y. Y. *et al. Adv. Mater.*, (2010), **22**, 2082-2086

[2] Kim, Y. Y. *et al. Nature Mater.*, (2011), **10**, 890-896

[3] Borukhin, S. *et al. Adv. Funct. Mater.*, (2012), doi: 10.1002/adfm.201201079

Talk 5

Predicting the morphology of organic solids

Dinesh-Ramesh Mirpuri Vatvani^{a,b}, *William Jones*^a, *Colin Groom*^b

^a *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW.*

^b *Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ.*

In the field of crystallography, morphology is a term used to describe the macroscopic shape of a crystal. Morphology influences the physical properties of crystalline materials (e.g. pharmaceutical products), it would therefore be invaluable for industry to predict morphology reliably.

Several commonly used methods exist to predict crystal morphology, but these methods have only been validated using a very small number of crystal structures, so their domain of applicability is still a matter of debate.

Our work utilises the Crystal Structure Database (CSD), which contains over a three hundred thousand structures with descriptions of experimental morphologies, to validate two of the most widely used morphology prediction models; the BFDH model [1] and the attachment energy model [2]. We have quantified the relative performance of these models and well as developed a new morphology prediction model based on the 2-D nucleation crystal growth mechanism, which rivals the BFDH and attachment energy models.

References:

[1] Donnay, J. D. H., Harker, D., A new law of crystal morphology extending the law of Bravais, *Amer. Min.*, (1937), **22**, 446-467.

[2] Hartman, P., Perdok, W. G., On the relations between structure and morphology of crystals. *Acta Crystallographica*, (1955), **8**, 49–52.

Talk 6

Ab initio DFT-assisted structural characterisation of MOFs

Jason R. Loader, Charles A. Mason, Paul Smart, Benjamin J. Irving, *Anthony J. H. M. Meijer and Lee Brammer*

Department of Chemistry, University of Sheffield, Brook Hill, Sheffield, S3 7HF.

Metal-organic frameworks (MOFs) are crystalline materials of metal ions individually linked through polytopic organic bridging ligands; yielding large, open and typically porous framework structures. This structural property enables their potential in highly desired applications such as gas storage and heterogeneous catalysis. Furthermore, their simple two-component nature readily allows the framework topology to be predesigned and tailored, through conscious choice of both metal and organic elements, for a specific application. However, the incorporation of additional functional groups, for enhanced host-guest interactions within the framework, is required to drive this field forward. Unfortunately, additional functional groups within the traditional one-pot self-assembly synthesis of MOFs often interfere with the formation of the targeted framework topology; necessitating an alternative method to introduce framework functionality. Current techniques around this issue

include post-synthetic modifications (PSM) and exchanges (PSE) that introduce functionalization after the framework is self-assembled. Such transformations however, seldom occur as a single-crystal-to-single-crystal mechanism resulting in difficulties for full structural characterisation of the new material.

We are interested in how computational modelling of MOFs can assist in characterisation when experimental data is limited and/or poor. We have focused on one MOF system that undergoes a two-step PSE to introduce functionality; building a family of functionalised MOFs from the same parent MOF, which are often unobtainable in a single step. Product characterisation of the parent MOF was fulfilled by single-crystal XRD but that for both the intermediate and final stages was limited to powder XRD. For the intermediate phase, the data were insufficient to build any form of structural model, and for the latter, only a very rough model based on the original parent structure could be constructed. Fortunately, the PSE of this MOF system can be followed via infra-red (IR) spectroscopy revealing additional structural clues for the intermediate and final phases. Our computational studies involve both simple isolated molecular fragment and fully periodic models to investigate this MOF system. We explore the plausibility of using molecular fragments to model the entire MOF framework with verification of their accuracy made by comparison of experimental and computational IR spectra. We also show how fully periodic models aid analysis of diffraction data that, along with calculated IR data, enabled full structural determination of both the intermediate and final phases; the latter of which also required the exceptionally atypical use of ligand-protein docking software (GOLD) to great success.

Talk 7

The Link Between Macbeth's Banquo and Molecular Aromaticity

Kate E. Horner, *Peter B. Karadakov*

Department of Chemistry, University of York, Heslington, York, YO10 3UD

Aromaticity is a concept that is introduced to chemists at an early stage in their education. However, the complexities of characterising this property, as well as antithetical antiaromaticity, only become apparent on later investigation. With the proliferation of computational techniques throughout all branches of science, the nucleus-independent chemical shift (NICS) [1] has become increasingly popular as a magnetic criterion for describing aromaticity. This process involves the calculation of the isotropic chemical shielding value at a point at the centre of an aromatic ring and again 1 Å above, the latter of which avoids the σ -contribution, focussing on the π -density responsible for aromaticity. The method uses ghost atoms (named Bq atoms from Banquo), as opposed to molecular probes, which prevents any wavefunction perturbation. The shielding tensors are integrals of the electron density which take into account both the density and electronic movement hence provide more information than electron density plots alone.

This work takes the technique further and calculates these NICS values at many points across a fine, two-dimensional grid using *Gaussian 09*. This plane can be oriented through the molecular plane or in a vertical direction perpendicular to the former. These shielding values are then illustrated with contour plots which give remarkable insight into, not only the nature of aromaticity and antiaromaticity, but also into that of chemical bonding.

The archetypal systems exhibiting aromaticity and antiaromaticity, benzene and cyclobutadiene, were probed in this manner and a clear distinction was seen. Benzene displays moderately shielded bonding regions, a small area of low shielding at the ring centre and intriguing deshielded regions around the carbon nuclei. Cyclobutadiene also presents

deshieldings around the carbon nuclei but in this case, also demonstrates a dumbbell-shaped deshielding feature at the ring centre that extends above and below the molecular plane. Moreover, there is a clear distinction between adjacent carbon-carbon bonds with evidence of bent-bonding caused by the degree of ring strain. Further revelations were made on investigation of aromatic heterocycles regarding the nature of atoms that exhibit these halos of deshielding. This technique shows significant potential for a variety of applications and would benefit many fields.

References:

[1] Schleyer, P. v. R., Maerker, C., Dransfeld, A., Jiao, H. and van Eikema Hommes, N. J. R., Nucleus-Independent Chemical Shifts: A Simple and Efficient Aromaticity Probe, *J. Am. Chem. Soc.*, (1996), **118**, 6317-6318

Talk 8

The Statistical Limits of Molecular Scoring

Gregory A. Ross^a, Garrett M. Morris^b and Philip C. Biggin^a.

^a*Structural Bioinformatics and Computational Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom.*

^b*InhibOx, Pembroke House, 36-37 Pembroke Street, Oxford, OX1 1BP, United Kingdom.*

The three dimensional structure of a protein can give a great deal of insight into the mechanism of its function. Perhaps surprisingly, it is very difficult to infer the binding affinity of a complex even when its structure is known. Addressing this problem is of great concern to virtual screening, a process in which up to many millions of compounds are assayed *in silico* against a protein target. A fast and trustworthy affinity estimator could potentially streamline the drug discovery process, reducing reliance on expensive wet lab experiments, speeding up the discovery of new hits and aiding lead optimisation. Typically applied using a single 'snapshot' of a molecular complex, structure-based scoring functions are fast and simple enough to meet the computational demands of testing vast chemical libraries. However, despite three decades of development, their performance remains highly variable and dependent on the protein system studied [1]. As a result, it is not yet certain whether it is even possible to create a scoring function that is universally applicable to all protein-ligand complexes.

Using statistical learning and information theory, we have investigated the maximum theoretical accuracy and fundamental sources of error in structure-based scoring functions. Our analysis has indicated that, using current methods, a single scoring function cannot achieve the minimum theoretical error on all protein systems. This result is independent of the regression method and the training set used. To test the theoretical framework, we have developed scoring functions with a variety of descriptors, which include explicit water information using our own water placement method [2]. We also have utilised linear and non-linear machine learning techniques. Our results are consistent with our predictions. It is hoped that highlighting the fundamental limitations and successes of scoring functions in a rigorous fashion will facilitate the creation of new methods for fast free energy calculations.

References:

[1] Cheng, T., Li, X., Li, Y., Liu, Z., and Wang, R., Comparative assessment of scoring functions on a diverse test set. *J. Chem. Inf. Model.*, (2009), **49**,1079-1093.

[2] Ross, G. A., Morris, G. M., and Biggin, P. C., Rapid and accurate prediction and scoring of water molecules in protein binding sites. *PLoS ONE*, (2012), **7**, e32036.

Talk 9

Coarse-grained molecular dynamics simulations of the interactions of two cell penetrating peptides with complex membrane models

Jean Hélie^a, Francesca Milletti^b, Mickael Lemousin^a, Charlotte M. Deane^c and *Mark S. P. Sansom*^a

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

^b*Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG*

^c*Hoffmann-La Roche, Nutley, NJ, USA*

Cell-penetrating peptides (CPP) can permeate cellular membranes and are therefore attractive vectors for gene therapy and drug delivery. However their uptake mechanisms are still poorly understood [1]. Peptides classified as CPP are usually enriched in basic residues and thus positively charged but except from this shared characteristic display a wide range of physico-chemical properties. A unifying interpretation of experimental results is rendered all the more difficult by the great diversity in experimental setups; nevertheless there is evidence that some CPP can be internalized by both endocytic and direct translocation pathways. In particular, peptide concentration and amphipathicity have been shown to be important for membrane disruption and passive permeation. Membrane lipid composition and electrostatic properties also appear to play a crucial role in the peptides activity.

Molecular dynamics simulations are useful to gain molecular level insights into CPPmembrane interactions but full atomistic descriptions remain very computationally expensive thus making the investigation of various membrane properties difficult. Here we present coarse-grained (CG) simulations over several microseconds of two wellstudied

CPP (penetratin and transportan) in large, asymmetric bilayers with complex, biologically relevant lipid compositions. The influence of peptide secondary structure was also explored. We discuss the development of a modelling approach focusing on membrane properties as well as peptides characteristics to investigate peptide-lipid interactions.

References:

[1] A. Ziegler, Thermodynamic studies and binding mechanisms of cell-penetrating peptides with lipids and glycosaminoglycans, *Adv. Drug Del. Rev.* (2008) 60, 580-597.

Talk 10

Structural and thermodynamic insights into bacterial outer membrane lipid signaling by the innate immune system

Teresa Paramo, Susana Tomasio, *Peter J. Bond*

Unilever Centre for Molecular Informatics, University of Cambridge, Department of Chemistry, Lensfield Road CB21EW (Cambridge).

Lipopolysaccharide (LPS) from bacterial outer membranes is a potent early indicator of microbial infection and the primary inducer of fatal septic shock syndrome. Its recognition by the immune system is carried out by Toll-like receptor 4 (TLR4) when associated with its co-receptor MD-2, an immunoglobulin-like protein. MD-2 adopts a characteristic "beta-cup" fold with a large hydrophobic cavity, and is able to bind a variety of lipid species. Subtle alterations in the structure of LPS derivatives can profoundly alter the resultant immunological response[1], hampering the rational design of TLR4 immunomodulators. To unravel the associated structure-activity relationships, we have performed long-timescale, all-atom molecular dynamics simulations and free-energy calculations of the isolated MD-2 co-

receptor and the entire signaling-active receptor complex in the presence of a variety of LPS species, as well as an LPS membrane. Unbiased simulations revealed that the MD-2 cavity is highly conformationally flexible, identifying spontaneous switching between active signalling-competent and inactivated states dependent upon the presence of different ligands, leading us to propose a conserved receptor activation mechanism. Novel code designed for efficient characterization of protein cavities in simulation trajectories revealed a strong correlation between the size of the bound ligand and cavity volume, consistent with experimental data. To gain insights into the thermodynamic determinants of endotoxin recognition, extensive umbrella sampling has been applied to estimate the potential of mean force (PMF) for the binding of LPS molecules to MD-2 co-receptor. Strikingly, stronger binding to signalling-inactivate MD-2 were observed for antagonists, and conversely, stronger binding to active MD-2 for agonists. Comparison of this data to the first ever PMF calculated for extraction of LPS from a model of the bacterial outer membrane has revealed how MD-2 creates a "membrane-like" environment within its protein cavity, providing a mechanism for sensitive LPS recognition by the innate immune system.

References:

[1] Teghanemt A, Zhang D, Levis EN, Weiss JP and Gioannini TL., Molecular basis of reduced potency of underacylated endotoxins, *J. Immunol.*, (2005), **175**, 4669-4676.

Talk 11

Catalytic activity of fatty acid amide hydrolase investigated by EVB free energy calculations

Ewa Chudyk¹, Maria Frushicheva², Arieh Warshel², Jitnapa Sirirak¹, Alessio Lodola³, Marco Mor³, Adrian J. Mulholland¹

¹Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, UK, Ewa.Chudyk@bristol.ac.uk

²Department of Chemistry, University of Southern California, Los Angeles, CA 90089

³Dipartimento Farmaceutico, Università degli Studi di Parma, 43100 Parma, Italy

Fatty acid amide hydrolase (FAAH) is an enzyme that deactivates neurotransmitters involved in inducing sleep, anxiety, inflammatory and pain states. Such a crucial role in living organisms requires more detailed knowledge of the molecular basis of FAAH-catalyzed oleamide hydrolysis in order of its rational control which might find applications in inhibitor design process. It is a target for drug design.

The rate-limiting step in this reaction occurs during the acylation stage and is thought to include two proton transfers followed by tetrahedral intermediate formation. Geometrical descriptors for this reaction were investigated by QM/MM calculations, first by B3LYP/6-31Gd(+)//PM3-CHARMM potential energy surface calculations [1], and then analyzed using the differential transition state stabilization method. This provided a general insight into the reaction mechanism, as well as indicating a crucial role of conformational fluctuations within the FAAH active site.

To investigate catalysis in FAAH we have used an empirical valence bond (EVB) approach, with a combined free energy perturbation umbrella sampling method [2]. Structural changes within FAAH active site during reaction strongly influence its energetic conditions, and may be important in specificity. We have also investigated mutant enzymes. These results along with the enzyme's stabilization effects compared to the equivalent 'reference' reaction in water are discussed.

References:

- [1] A. Lodola et al., Conformational Effects in Enzyme Catalysis: Reaction via a High Energy Conformation in Fatty Acid Amide Hydrolase (2007), *Biophys. J.*, 92: L20-L22
[2] A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (1991), Wiley Interscience, New York

Talk 12

Computer Simulations of the Molecular Role of EGFR Kinase Mutations

Christopher Bull^a, Richard Ward^b, Jonathan Essex^a

^a*School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ.*

^b*Oncology iMed, AstraZeneca, 35S155, Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TG*

Epidermal Growth Factor Receptor (EGFR) is a signalling protein whose mutation and upregulation is implicated in various cancers. Dimerisation promotes the large conformational changes that are required for activation of the kinase domain of wild-type (WT) EGFR, however mutants are able to adopt the active conformation far more readily. Additionally, mutations modulate the efficacy of EGFR kinase inhibitors. Despite the potent effects of a number of activating and resistance mutations, the mechanisms behind those effects are not well understood. While much progress has been made using computational techniques, the deep energy minima within which the active and inactive conformations of EGFR lie make it difficult to examine the “big picture” of WT and mutant EGFR's conformational dynamics.

The current work employs microsecond time-scale molecular dynamics (MD) on WT EGFR and three clinically relevant mutations, in addition to dimensionality reduction techniques aimed at differentiating minima conformations from the initial noisy dataset. We show that these techniques are capable of identifying possible transition states involved in long time-scale conformational change, as well as the differential sampling of trajectories due to mutation. By highlighting regions in which motions differ between mutants we identify possible sites of interest for accelerated sampling techniques such as targeted MD and reversible digitally filtered MD that may help us to understand better the origin of the different properties of the EGFR mutants, as well as having wider application to systems with similarly deep energy minima.

References:

- [1] Shan, Y., Eastwood, M. P., Zhang, X., Kim, E. T., Arkhipov, A., Dror, R. O., Jumper, J., Kuriyan, J., Shaw, D. E., Oncogenic Mutations Counteract Intrinsic Disorder in the EGFR Kinase and Promote Receptor Dimerization, *Cell*, (2012), **149**, 860-870.

Programme of Poster Presentations

Poster 1	Design and validation of a novel method of MD trajectory analysis for polymer and polymer-drug characterisation <i>Glib Meleshko, Cardiff University</i>
Poster 2	Metal-Support Interactions in Catalysis: A Theoretical Investigation <i>Soon Hoh, Cardiff University</i>
Poster 3	Homology Modelling of PepT1, a Human Peptide Transporter <i>Firdaus Samsudin, University of Oxford</i>
Poster 4	Predicting the regioselectivity and reactivity of 1,3-dipolar cycloaddition reactions of organic azides with quinone derivatives using QM calculations <i>Maysaa Saleh, University of Nottingham</i>
Poster 5	The radical reactions of adenosylcobalamin dependent enzymes <i>Patrick von Glehn, University of Bristol</i>
Poster 6	Conformational flexibility of peptides with chemopreventive activity <i>Sara Kramar, UCL School of Pharmacy</i>
Poster 7	Plane of Best Fit: A Novel Method to Characterize the Three-Dimensionality of Molecules <i>Nicholas Firth, Institute of Cancer Research</i>
Poster 8	Modelling and Dynamics of the SV2A Synaptic Vesicle Protein <i>Joanna Lee, University of Oxford</i>
Poster 9	Towards Low Cost Virtual Biological Laboratories: Molecular Modelling Simulation on Commodity Hardware <i>Ardita Shkurti, University of Southampton</i>
Poster 10	Annotating Targets With Pathways: Extending Approaches to Mode of Action Analysis <i>Sonia Liggi, University of Cambridge</i>
Poster 11	Computational modelling and the reduction of animal testing in toxicology studies <i>Grace Edmund, University of Surrey</i>
Poster 12	Predictive modelling of P-glycoprotein binding compounds using IC₅₀ values <i>Mohsen Sharifi, University of Kent</i>

Poster 1

Design and validation of a novel method of MD trajectory analysis for polymer and polymer-drug characterisation

Glib Meleshko, David J Willock, James A Platts, Alison Paul

School of Chemistry, Cardiff University, Park Place, Cardiff

The conformation that polymer-drug conjugates adopt in solution has a significant effect on properties important for their performance as drug delivery systems. For N-(2-hydroxypropyl)methacrylamide (HPMA) copolymer conjugates it is known that aggregation number, size and shape affect the rate at which the drug is enzymatically cleaved from the polymer backbone.¹ Therefore structural properties such as size, shape and density distribution of a range of HPMA copolymers have been investigated. The suitability of atomistic force fields has been assessed against rotational barriers and relative conformational energies obtained from ab initio and DFT data for a monomer and dimer of HPMA.

Following this, the AMBER99 parameter set was chosen for all molecular dynamics simulation. Radius of gyration (Rg), radial distribution function (RDF), shape, and density profiles of particular atom types were calculated for a range of HPMA homopolymers sizes from 4 to 200 repeat units (2 to 35 kDa), and results interpreted in the context of Flory's mean field approach, and compared with data obtained from small angle neutron scattering (SANS) experiments. Results of this study were used for complex investigation of HPMA conjugates with drug mimics. A range of linear amines (aminohexane(C6), aminooctane(C8), aminododecane (C12)), hydroxyl and fluoro terminated linear amines as well as aromatic aminoanthracene (ANC), aminocrysene(AC) and aminoanthraquinone (ANQ), bound to the polymeric carrier via a tetrapeptide linker (glycine-phenylalanine-leucine-glycine) (GFLG) (Mw ~ 30 kDa) were selected as model objects for study of the effect of drug type and loading on HPMA copolymer conformation. SANS experimental scattering curves were compared with theoretical curves, predicted from MD simulations. Parameters such as size and shape fitted to SANS data were compared with relevant simulated structures. Results for aromatic drug conjugates showed better agreement with experiment than linear amines, and suggest evidence of aggregation for some structures. Computational simulation of polymer-drug conjugates has a practical application in obtaining details on morphological information of polymer structure, and can be combined with experimental techniques to provide a view of solution behaviour. This allows design of new materials for drug delivery, characterization of known biological agents and investigation of structure dependence on copolymer types and drug loading in conjugate systems.

References:

[1] Bader, H., H. Ringsdorf and B. Schmidt, *Water-Soluble Polymers in Medicine, Macromol.Chem. and Phys.*, (1987), **188**, 1261-1272

Poster 2

Metal-Support Interactions in Catalysis: A Theoretical Investigation

Soon W. Hoha, David J. Willock^a, Glenn Jones^b

^a*Cardiff Catalysis Institute, School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff, CF10 3AT.*

^b*Johnson Matthey Plc, Technology Centre, Blounts Court, Sonning Common, Reading RG4 9NH*

The use of catalysts which contains both metals and metal oxides is of great interest with widespread applications in heterogeneous catalysis. A few published literatures of carbon monoxide adsorbed on the metal reacting with oxygen species associated with the oxide has implied the importance of interactions across the metal-oxide interface.[1-2] Our project involves the Au/oxide catalysts recently developed at Cardiff and the of study metal-support interactions and their role in oxidation catalysis. To date, the interface of Au/Fe₂O₃ is found to strongly stabilise atomic oxygen. This means that on an Au₁₀ cluster in isolation, the barrier for O₂ desorption is lower than that for dissociation, whereas at the oxide interface the reverse is true. So that for an isolated nanoparticle only molecular oxygen will be present, but at the transition metal oxide interface atomic oxygen will form and Au atoms will be oxidised. Studies are extended to M/CeO₂ (with M = Au, Pd, Pt, Ni) and the creation/role of surface oxide defects. For example, is it easier to remove oxygen from the lattice at the particle/oxide interface than from the clean oxide? This will lead to the investigation of oxidation mechanisms involving a Mars-van Krevelen mechanism. A number of studies show that metal particles on CeO₂ can be in various oxidation states depending on the catalyst environments.

The initial steps in this process can be addressed from our small particle (10 metal atoms) on defective oxide surfaces. Periodic DFT calculations are applied to the study of the influence of metal oxide supports on the band structure of metal overlayers. A simple slab methodology will be used to construct periodic models of metal/oxide interfaces. These will be used to calculate the band structure of the metal overlayers for comparison with the pure metal case. Besides, the adsorption of probe molecules such as CO and CH₂O; and the activation of primary and secondary C-H bonds in propane will also be considered.

References:

[1] A. Holmgren, F. Azarnoush and E. Fridell., Influence of pre-treatment on the low-temperature activity of Pt/ceria, *Appl. Catal., B*, (1999), **22**, 49-61

[2] S. Golunski, R. Rajaram, N. Hodgeb, G.J. Hutchings, and C.J. Kiely, Low-temperature redox activity in co-precipitated catalysts: a comparison between gold and platinum-group metals, *Catal. Today* (2002), **72**, 107-113

Poster 3

Homology Modelling of PepT1, a Human Peptide Transporter

Firdaus Samsudin, *Philip W. Fowler, Mark S. P. Sansom*

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU.

The mammalian solute carrier 15 transporter family consists of four proton-coupled oligopeptide transporters PepT1, PepT2, PHT1 and PHT2. PepT1 is expressed in the small intestine and facilitates absorption of di- and tripeptides from protein digestion. Due to the broad substrate specificity, this transporter also mediates the uptake of some pharmacologically relevant compounds including the β -lactam antibiotics and some antiviral drugs. Structural studies of PepT1 is therefore of major interest for rational drug design to increase oral bioavailability. Recent publication of the crystal structure of bacterial peptide transporters, PepTSo [1] and PepTSt [2], provides an excellent model system to study mammalian peptide transporters due to the high sequence similarity of the transmembrane region and the conservation of key residues in peptide transport.

Computational methods, namely homology modelling and atomistic molecular dynamics (MD) simulations, have been employed to gain additional insights into this membrane protein. Using bacterial peptide transporters as the template structure, a full-length homology model of human PepT1 was built. Essential residues for peptide transport as suggested by various mutational studies align well with the hydrophilic binding cavity of the model transporter. Also, most deleterious SNPs from the 1000 Genomes Project are mapped on the transmembrane region as expected. To assess the structural integrity of the PepT1 model, atomistic MD simulations were conducted in a model membrane. This model shows relatively low RMSD progression and very good secondary structure preservation during the simulations.

Our PepT1 homology model provides an excellent framework for future studies such as pharmacophore modelling, molecular docking of potential drugs and modelling of conformational changes involved in transport mechanism. Collectively, these computational methods will enhance our understanding of mammalian peptide transport system and help improve drug delivery.

References:

[1] Newstead, S., Drew, D., Cameron, A. D., Postis, V. L. G., Xia, X., Fowler, P. W., Ingram, J. C., Carpenter, E. P., Sansom, M. S. P., McPhearson, M. J., Baldwin, S. A., and Iwata, S.,

Crystal structure of prokaryotic homologue of the mammalian oligopeptide-proton symporter, PepT1 and PepT2. *EMBO J.*, (2011) **30(2)**, 417-426.

[2] Solcan, N., Kwok, J., Fowler, P. W., Cameron, A. D., Drew, D., Iwata, S., and Newstead, S., Alternating access mechanism of the POT family of oligopeptide transporters, *EMBO J.*, (2012) **31(16)**, 3411-3421.

Poster 4

Predicting the regioselectivity and reactivity of 1,3-dipolar cycloaddition reactions of organic azides with quinone derivatives using QM calculations

Maysaa Saleh^{1,2}, Christopher Moody², and Charlie Laughton¹

¹School of Pharmacy and ²School of Chemistry, University of Nottingham, Nottingham NG7 2RD.

The unlimited proliferation of cancer cells depends on a telomere maintenance mechanism which is provided by telomerase enzyme¹. The telomeric ends form structures called G-quadruplexs. Stabilization of these structures by small binding molecules called G4 ligands inhibits telomere elongation, and targets telomere maintenance mechanism, resulting in ultimately delayed cell death and abrogation of tumourigenicity *in vivo*². In this project, we are aiming to develop a new series of triazoles which are designed to bind to and stabilize G-quadruplex structures selectively, and which may therefore have potential as anti-cancer drugs. These triazoles are synthesized via 1,3-dipolar cycloaddition reactions of organic azides with quinone derivatives which afford two possible bis-triazole products: centrosymmetric and noncentrosymmetric regioisomers. The different regioisomers are expected to have different DNA-binding characteristics. Quantum mechanics calculations have been used to understand the regioselectivity of 1,3-dipolar cycloaddition reactions between the organic azides (1,3-dipoles) and quinones (dipolarophiles), and to predict the identity of the preferred regioisomer. The calculations are helpful in the study of the effect of electron-withdrawing and/or electron-releasing groups and conjugation on the molecular orbitals' energies and coefficients of azides and quinones, and thus on the rates of azides (dipoles) – quinones (dipolarophile) 1,3-dipolar cycloaddition reactions. Models for the azide and quinone molecules were generated and optimized using the Sybylx program [RFT], while energies and coefficients of molecular orbitals for the optimized molecules were calculated using MOPAC [RFT]. Quantum mechanics calculations predicted that in all 1,3-dipolar cycloaddition reactions, which are carried out in this project, between organic azides and *p*-benzoquinone, the second cycloaddition reaction is controlled by the dipole (azide) HOMO – dipolarophile (mono-triazole) LUMO interaction, which results in the formation of the centrosymmetric regioisomer as the most stable geometry and favoured product. Furthermore, QM calculations predict that the regioselectivity and reactivity of 1,3-dipolar cycloaddition reactions between azides (dipole) and quinone derivatives (dipolarophile) are affected by the presence of phenyl and benzyl substituents next to the interacting centre (N=N=N) of the dipole, two carbonyl groups next to the interacting centre (C=C) and the triazole ring fused to the quinone in the dipolarophile. Subsequent synthetic studies have confirmed the modelling predictions.

References:

[1] Moorhouse, A. D.; Haider, S.; Gunaratnam, M.; Munnur, D.; Neidle, S.; Moses, J. E., Targeting telomerase and telomeres: a click chemistry approach towards highly selective G-quadruplex ligands. *Mol. BioSyst.*, (2008), **4**, 629–642.

[2] Campbell, N. H.; Patel, M.; Tofa, A. B.; Ghosh, R.; Parkinson, G. N.; Neidle, S., Selectivity in Ligand Recognition of G-Quadruplex Loops. *Biochemistry*, (2009), **48**, 1675–1680.

Poster 5

The radical reactions of adenosylcobalamin dependent enzymes

Patrick von Glehn, Jeremy Harvey, Adrian Mulholland

School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 ITS

Patrick.vonGlehn.07@bristol.ac.uk

Adenosylcobalamin (AdoCbl) dependent enzymes catalyse difficult carbon skeleton rearrangement reactions via radical mechanisms. The initial and radical generating step in all of these enzymatic reactions is the homolytic cleavage of the AdoCbl Cobalt-Carbon bond which is estimated to be accelerated by up to 12 orders of magnitude in enzyme compared to in solution. There are several competing hypothesis as to the origin of this large catalytic effect.[1]

Radical formation is followed by hydrogen abstraction from the substrate. Unusual temperature dependencies of experimentally observed kinetic isotope effects suggest that there is significant hydrogen tunnelling taking place but it is hotly debated whether this is a catalytic effect or not.[2]

In collaboration with the experimental group of Neil Marsh (Michigan) we have already provided new insight into these systems using force field molecular dynamics and hybrid quantum mechanics/molecular mechanics (QM/MM) potential energy surface scans. We would like to investigate dynamic and nuclear quantum effects in this reaction, but reactive and quantum molecular dynamics simulations would be computationally intractable using quantum mechanics. Our current efforts are focused on applying the much cheaper empirical valence bond method combined with path integral simulations to investigate the free energy surfaces and quantum dynamics of adenosylcobalamin systems in enzyme and solution.

References:

[1] Sharma, Pankaz K., Zhen T. Chu, Mats H. M. Olsson, and Arieh Warshel. "A New Paradigm for Electrostatic Catalysis of Radical Reactions in Vitamin B-12 Enzymes" *PNAS* (2007), **104** (23), 9661–9666.

[2] Doll, KM, BR Bender, and RG Finke. "The First Experimental Test of the Hypothesis That Enzymes Have Evolved to Enhance Hydrogen Tunneling." *JACS* (2003), **125** (36), 10877–10884.

Poster 6

Conformational flexibility of peptides with chemopreventive activity

Sara Kramar, Geoff Wells, Mire Zloh

Department of Pharmaceutical and Biological Chemistry, UCL School of Pharmacy, 29-39 Brunswick Square, London WC1N 1AX, UK

Protein-protein interactions (PPIs) play an important role in cellular signalling pathways and are often used as targets for the development of new drug candidates. Inhibition of Nrf2-Keap1 interaction leads to the release of transcription factor Nrf2, which plays an important

role in xenobiotic and oxidative stresses, [1]. Direct inhibitors of the Nrf2-Keap1 PPI could therefore act as new chemopreventive drug candidates.

A series of peptides that interact with Keap1 and their binding activities have been reported [2] and in this work we are trying to gain insights into the relationship between ligand structure and activity by exploring the conformational flexibility of peptides and possible influence of flexibility on binding to Keap1. The peptide conformations were explored using different approaches, such as conformational search (Macromodel) and simulated annealing (Desmond), followed by docking (Hex, GLUE).

The well defined but large binding site of Keap1 is not fully occupied by its ligand in the X-ray structure of the Keap1-Nrf2 peptide complex (PDB entry: 1X2R) and water molecules may play an important role in formation of this complex. Our preliminary molecular dynamics (MD) studies with explicit water solvation indicated that the initial conformation of the peptide is crucial for the success of reproducing the complex from the crystal structure. To obtain initial conformations of all peptides with activity a conformational search of each peptide was carried out and the conformation with the lowest energy was used as an input for molecular docking. Resulting structures were subjected to simulated annealing protocol established to relax the resulting docking solutions. We have gained some understanding of intermolecular interactions between Keap1 and ligand peptides, and role of water molecules. These interactions can be improved or waters could be displaced by modifying the peptide sequences or designing new peptidomimetic agents with an improved binding affinity and potentially better inhibitory activity.

References:

[1] Motohashi, H., Yamamoto, M., Nrf2-Keap1 defines a physiologically important stress response mechanism, *Trends Mol Med.*, (2004), **10**, 549-557.

[2] Hancock, R., Bertrand, H.C., Tsujita, T., Naz, S., El-Bakry, A., Laoruchupong, J., Hayes, J.D., Wells, G., Peptide inhibitors of the Keap1-Nrf2 protein-protein interaction, *Free Radic Biol Med.*, (2012) **52**, 444-451.

Poster 7

Plane of Best Fit: A Novel Method to Characterize the Three-Dimensionality of Molecules

Nicholas Firth, Nathan Brown, Julian Blagg

Cancer Research UK Cancer Therapeutics Unit, Division of Cancer Therapeutics, The Institute of Cancer Research, 15 Cotswold Road, Belmont, Surrey, SM2 5NG, UK.

Recently there has been a trend in drug design to tackle molecular targets classed on the edge of druggable space, for example protein-protein interactions. The move towards these targets is associated with the design of drugs with significant three-dimensional (3D) character [1]. Molecules with 3D character also increase the chance of higher solubility due to the nature of their solid-state crystal packing.

In this work we have sought a method that will rapidly assess the 3D character of large compound libraries, providing an easily interpretable score. The method that we have developed generates a plane of best fit for a conformation of a given molecule and calculates the distance of each heavy atom from the plane in ångströms and outputs the average of these distances.

We have applied this method to a number of diverse datasets and compared with methods of shape characterization from the literature. We have investigated these datasets to understand

the populated areas of chemical space using Plane of Best Fit (PBF), generating plots that describe the dynamic range of the datasets (Figure 1). When compared with other methods for characterizing the three-dimensionality of molecules, the PBF was found to have greater granularity and description of 3D character [2].

References:

- [1] D.C. Fry. Drug-Like Inhibitors of Protein-Protein Interactions: A Structural Examination of Effective Protein Mimicry, *Curr. Protein and Pept. Sci.*, 2008, **9**(3):240-247.
- [2] N.C. Firth, N. Brown, J. Blagg, Plane of Best Fit: A Novel Method to Characterize the Three-Dimensionality of Molecules. *J. Chem. Inf. Model.*, Article ASAP.

Poster 8

Modelling and Dynamics of the SV2A Synaptic Vesicle Protein

Joanna Lee^{a*} and Philip C Biggin^a

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

**e-mail: joanna.lee@bioch.ox.ac.uk*

SV2A is a transmembrane protein which is found on the synaptic vesicle. It has been found to be the target for Levetiracetam, an anti-epileptic drug (AED). This is a novel target for AEDs and so there is a clear desire to understand the mode of action for continued drug design. The SV2A protein has been suggested to be homologous to the MFS transporter family [1], but with an additional amino terminal domain that is thought to have a calcium-regulated interaction with synaptotagmin [2]. Previous mutational studies [3], performed on the basis of an homology model derived from the LacY structure, demonstrated a number of key residues involved in racetam binding. Due to its similarity to MFS transporter proteins, SV2A is thought to undergo transitions between outward and inward-facing conformations [4], therefore it is desirable to understand to what extent these residues are compatible with drug-binding in alternative conformations. To that end the recent crystal structures of EmrD, FucP and PepT provide additional start points for alternative models.

Homology modelling has been used to determine the inward and outward-facing structures of SV2A. We have also used molecular dynamics simulations to investigate the dynamic behaviour of the SV2A protein within a lipid bilayer environment in these different conformational states in order to gain insight into which residues are most likely to play a role in drug-binding. Our results should lead to hypotheses which we will test via site-directed mutagenesis and drug-binding assays.

References:

- [1] Bajjalieh, S.M., Peterson, K., Shinghal, R., Scheller, R.H., SV2, a Brain Synaptic Vesicle Protein Homologous to Bacterial Transporters, *Science*, (1992), **257**, 1271-1273.
- [2] Schivell, AE., Mochida, S., Kensel-Hammes, P., Custer, KL., Bajjalieh, S.M., SV2A and SV2C contain a unique synaptotagmin-binding site, *Mol. Cell. Neurosci.*, (2005), **29**, 56 – 64 .
- [3] Shi, J., Anderson, D., Lynch, B.A., Castaigne, J.G., Foerch, P., et al., Combining modelling and mutagenesis studies of synaptic vesicle protein 2A to identify a series of residues involved in racetam binding, *Biochem. Soc. Trans.*, (2011), **39**, 1341-1347.
- [4] Lynch, B.A., et al., Visualization of SV2A conformations in situ by the use of Protein Tomography, *Biochemical and Biophysical Research Communications*, (2008) **375**, 391–395.

Poster 9

Towards Low Cost Virtual Biological Laboratories: Molecular Modelling Simulation on Commodity Hardware

Ardita Shkurti^{ab}, Andrea Acquaviva^a, Elisa Ficarra^a, Mario Orsi^b, Enrico Macii^a, Jonathan W. Essex^b

^a *Department of Control and Computer Engineering, Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129, Torino, Italy*

^b *School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom*

Many essential cell processes, such as the conformation of embedded proteins, membrane permeability, interaction with drugs and signalling, are directly connected to the molecular dynamics of cell membranes. The importance of this biology has led to an intensifying demand for hardware and software optimized models and tools, implemented on commodity high performance low-cost hardware, in order to provide the scientific community with virtual low cost laboratories.

In the light of these considerations, we implemented an accelerated version of a molecular dynamics coarse-grain lipid bilayers simulator on commodity Graphic Processing Units (GPU) architectures. The characteristics of this molecular dynamics model, such as new force fields for pair potentials that include an unconventional representation for water and charges, were particularly challenging. We introduced new algorithms and data structures required by coarse-grain models compared to atomistic ones, for the modelling of the integration timestep, neighbour list generation, and nonbonded force interactions. We characterized the impact on performance of biological systems of differing complexity in terms of size, particle type and timestep. We also compared the simulations of many particle-type systems against single particle-type systems, to evaluate the overhead of additional structures needed to model more complex molecules. Moreover, we performed a detailed analysis on the profiling of the simulation code and its execution flows due to the computation of the non-bonded forces. Finally, we characterized the acceleration and accuracy of the simulations on three GPUs having different computation capabilities and parallelism, achieving one order of magnitude faster simulation execution times.

The lessons learned from this study have implications for the porting of molecular modelling software to these new, emerging, computational architectures in terms of the effort required to achieve an efficient and effective software implementation, and the improvements in performance that can be expected.

References:

[1] Shkurti, A., Acquaviva, A., Ficarra, E. et al., Characterization of coarse-grain molecular dynamic simulation performance on graphic processing unit architectures, *Proceedings of the International Conference on Bioinformatics Models, Methods and Algorithms*, (2012), 339-342.

[2] Orsi M., Michel J. and Essex J. W., Coarse-grain modelling of dmpe and dopc lipid bilayers, *J. Phys.: Condens. Matter*, (2010) **22**, 155106.

Poster 10

Annotating Targets With Pathways: Extending Approaches to Mode of Action Analysis

Sonia Liggi, Alexios Koutsoukas, Robert C Glen, Andreas Bender

Unilever Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK

When attempting to treat pathological conditions, it is often necessary to act on multiple targets in order to modify the implicated biological network(s). A systems biology approach would help the drug discovery field by providing a link between chemistry and biology [1]. In particular, annotation of targets with pathways would allow a better understanding of a drug mechanism of action, side effects and promiscuity, as well as complex network modulation. A deep understanding of these factors is fundamental to design a drug with the desired pharmacological profile.

In this context, a cheminformatics target prediction tool [2] was used on several small molecule phenotypic datasets, and the predicted targets were annotated with the pathways they belong to. Predicted targets and annotated pathways were subjected to an enrichment calculation in order to eliminate the prediction noise and highlight only those likely to be implicated in the phenotype studied. Several enrichment methods were tested and compared.

Comparison of several enrichment methods demonstrates that there is not a single method that stands out, but instead a combination of several methods is needed to increase the significance of the readouts.

The results obtained show that pathways annotation can give new insight into understanding drugs' mechanism of action, underlining pathways that could be implicated in the studied phenotype, which would not be brought out by the consideration of only predicted targets.

References:

[1] Hood, L. and Perlmutter, R.M., The impact of systems approaches on biological problems in drug discovery, *Nat Biotechnol* (2004), **22**, 1215-1217.

[2] Koutsoukas, A., Lowe, R., Kalantar Motamedi, Y., Mussa, H.Y., Klaffke, W., Mitchell, J.B.O., Glen, R.C., Bender, A., In silico target predictions: comparing multiclass Naïve Bayes and Parzen-Rosenblatt Window and the definition of a benchmarking dataset for target prediction, *J Chem Inf Model* (2012) **Submitted**.

Poster 11

Computational modelling and the reduction of animal testing in toxicology studies

Grace. H. C. Edmund, *Brendan J. Howlin, Dave Lewis*

Department of Chemistry, University of Surrey, Stag Hill, Guildford, GU2 7XH

Cytochrome P450 enzymes (CYP450) are central to drug metabolism, carrying out the metabolism of 75% of known drugs in current clinical use. Important effects such as adverse drug reactions and genetically determined differences in drug toxicity and efficacy depend on CYP450 activity, yet this cannot be easily predicted from protein structures alone and is often found through the use of animal models [1-3].

Computational methods allow for a better understanding of the molecular determinants of reactivity and specificity, potentially contributing significantly to both drug development, assisting in the prevention of the growing and expensive problem of late stage drug failures (which often occur due to CYP450 mediated ADME-Tox properties), and in the reduction (and eventual replacement) of animal testing at this stage of drug development.

Our aim is to provide validated computer models of the CYP450 enzymes that can be used in *in silico* screening of new drug molecules thereby greatly reducing the need for animal testing

in toxicity studies. By using 3D homology modelling and molecular docking, we are investigating the species difference observed for rat and human on the binding of the two stereoisomer's of well known inhibitors of CYP450 2D6, quinidine and quinine. By investigating the structural and electronic features that determine selectivity in these reactions, we hope to properly evaluate the reliability of animal models in predicting human ADME-Tox properties while also creating a series of simple descriptors for the accurate prediction of species selectivity and human ADME-Tox properties reducing animal testing in the pharmaceuticals industry.

References:

- [1] de Groot, M. J., Vrije Universiteit, 1996.
- [2] Mulholland, A. J. *Drug Discovery Today* **2005**, *10*, 1393.
- [3] Rowland, P.; Blaney, F. E.; Smyth, M. G.; Jones, J. J.; Leydon, V. R.; Oxbrow, A. K.; Lewis, C. J.; Tennant, M. G.; Modi, S.; Egglestone, D. S.; Chenery, R. J.; Bridges, A. M. *Journal of Biological Chemistry* **2006**, *281*, 7614.

Poster 12

Predictive modelling of P-glycoprotein binding compounds using IC₅₀ values

Mohsen Sharifi, *Taravat Ghafourian*

Medway School of Pharmacy, University of Kent, Central Avenue, Chatham Maritime, Kent ME4 4TB

P-glycoprotein 1 (also known as MDR1/ ABCB1 or ATP-binding cassette sub-family member 1) is a pump within the cells that efflux chemicals out of the cells and detoxifies cells. In the cancer treatment as we increase chemotherapeutics within the body, P-gp can be over-reactive in pumping chemotherapeutics out of the body and make the treatment ineffective. The prediction of inhibitory concentration (IC₅₀) in P-gp is therefore of paramount importance in the pharmacokinetics characterisation of drugs. This can be achieved by Quantitative Structure – Activity Relationships (QSAR) studies using several QSAR modelling methods, where IC₅₀ in P-gp is related to structural attributes of the drugs using statistical methods. The validated QSARs can then be used to estimate IC₅₀ of new drugs with the help of their chemical structures.

Inhibitory concentration values for 149 compounds were collated from University of California web page UCSF-FDA TransPortal [1] with the addition of some new data from the literature. Dataset was partitioned into a training set of 115, validation set of 29 compounds and 5 compounds in external set.

Molecular modelling followed by molecular descriptor calculations enabled development of linear regression and nonlinear SVM regression tree models and also more complex models of C&RT and MARSplines. The selected descriptors by stepwise regression analysis in Minitab and classification and regression trees in STATISTICA resulted in some insight into major factors that can affect inhibitory of P-gp. The best models were able to predict the logIC₅₀ values for the validation and external set with R-sq of 0.862. The most accurate models were the C&RT and MAR Spline model II respectively with the lowest average error of predicted IC₅₀ values and highest R-square values for the validation set.

References:

- [1]. <http://bts.ucsf.edu/fdatransportal>