

# ORAL PRESENTATIONS

---

## **Evolution of metal binding sites**

James Torrance, Malcolm McArthur, Janet Thornton

*European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD*

Bound metal ions are essential for the biological function of approximately a third of all proteins. The geometry of metal binding is well understood, but the evolution of metal binding sites has been less studied. We have looked at how metal binding site structure diverges between relatives, how often non-relatives converge on the same metal binding site structure, and how often metal binding sites are lost entirely in the course of evolution. Our dataset was dominated by zinc and calcium binding sites.

We found that the structural differences between metal-binding sites in related proteins were very small (typically less than 0.5 Å root-mean-square deviation). There was only a weak correlation between the size of this structural difference and the degree of overall sequence difference between two related proteins.

We have used structural templates representing metal binding sites to search the Protein Data Bank for cases where unrelated proteins have converged upon the same residue composition and geometry for metal binding. This has allowed us to identify nine “archetypal” metal binding site structures, each of which has arisen independently multiple times.

We found it was common for relatives of metal-binding proteins to lack metal-binding capacity, even where the original metal was essential for protein folding. Seventeen of the twenty-five metal-binding sites analysed displayed this type of metal binding loss among their homologous superfamily members. For most of the calcium-binding sites studied, the loss of metal binding in relatives was due to point mutation of the metal-binding residues (seven of the nine families where loss of metal binding occurred). For zinc-binding sites, loss of metal binding in relatives was rarer, and always involved loss of secondary structural elements or loops around the binding site. These differences between calcium and zinc are probably due to the fact that the zinc binding sites in our dataset play a structural role, whereas the calcium binding sites often have a regulatory role. Since the calcium ions are less important to the protein structure, they can be lost with less alteration to the structure.

---

## **Using multiobjective optimization to study the strengths of different interaction energies in protein-ligand complexes**

Sally Mardikian<sup>1</sup>, Valerie J. Gillet<sup>1</sup>, Richard M. Jackson<sup>2</sup>, David R. Westhead<sup>2</sup>

<sup>1</sup>*Department of Information Studies, University of Sheffield, Western Bank, Sheffield S10 2TN, UK*

<sup>2</sup>*Institute of Cellular and Molecular Biology, University Of Leeds, Leeds LS2 9JT, UK*

Despite the intense efforts that have been devoted to the development of scoring functions for docking, they are still limited in their ability both to identify the correct binding pose and to provide accurate estimates of binding affinities. Many scoring functions are based on summing contributions from different factors thought to influence binding. For example, molecular mechanics force field-based scoring functions sum the contributions from different interaction energy types (hydrogen bond, van der Waals, electrostatic) between ligand and protein atoms. The additive nature of these methods is such that the conformational search is directed by the final, total energy value, with no information on the importance of the different interaction energy types being used. Thus, the assumption is made that the influence of the different interaction energy types does not vary between different protein-ligand complexes.

We have developed a docking method based on a multiobjective genetic algorithm (MOGA) that is capable of simultaneously optimising individual interaction energy types of the GRID scoring function. The MOGA aims to lead the conformational search to what is known as the Pareto front. Solutions that fall on the front (the Pareto solutions) show different compromises in the different objectives, and they are all considered to be equivalent. By examining where the true solution falls on the Pareto surface we can gain insights into the relative importance of different contributions to binding for a given protein-ligand complex.

Twenty protein-ligand complexes have been tested using the MOGA optimisation method. For seventeen of these, the MOGA was able to obtain good solutions within the Pareto solution set (RMSDs of less than 2.0 Å from the crystal structure). An interesting feature of these solutions is that the different objectives do not show an equal influence. In fact, for many of the complexes, the van der Waals interactions alone appeared to be guiding the search. These results indicate that the strengths of different interaction energy types vary greatly from one complex to another, and that a scoring function should, in fact, be tailored to a given target. Our method could be used for scoring function development, by assessing the balance of the different interaction types for a single, or a group of related targets. It could also be used as a stand-alone docking tool that allows the user to select the most “realistic” of the Pareto solution set.

---

## **Simulation of conformational transitions and Free Energy Calculations of PcrA DNA Helicase by Targeted Molecular Dynamics and WHAM**

Hao Wang and Charles Laughton

*Centre for Biomolecular Sciences, School of Pharmacy, University of Nottingham, Nottingham, UK, NG7 2RD*

DNA helicases are important enzymes involved in all aspects of nucleic acid metabolism, ranging from DNA replication and repair to recombination, rescue of

stalled replication, and translation. DNA helicases are molecular motors: through conformational changes caused by ATP hydrolysis they move along the template double helix, breaking the hydrogen bonds holding the two strands together, and separating the template chains so that the DNA can be replicated and the genetic information can be accessed. Our project is focused on using molecular modelling to study the molecular mechanism of PcrA DNA helicase. Crystal structures are available for this helicase bound to DNA in both substrate (ATP bound) and product (ADP bound) states. We have used Targeted Molecular Dynamics (TMD) in AMBER8 to simulate transitions between known end structures. By TMD, the system is slowly forced to find a path from its initial state to the final conformation, by enforcing a restraint on the RMS coordinate deviation from this structure. In this case TMD has been used to identify the reaction pathway through which PcrA helicase moves along DNA as it ‘burns’ ATP, and calculate to the potential energy along the pathway.

The WHAM (Weighted Histogram Analysis Method) is an Umbrella Sampling method for free-energy and potential of mean force calculation. By combining TMD, searching the conformation space along the reaction, and WHAM together, the free energy changes along the reaction pathway are calculated. This approach is very general, permitting, in theory, a very ‘natural’ path to be found between the initial and final states. The result gives us a clear view of the mechanism of PcrA DNA helicase motion along single-stranded DNA, and the conformational changes in the protein and DNA that are required.

---

## **Drug binding to the hERG K<sup>+</sup> channel: a molecular modelling study**

Phillip J. Stansfeld<sup>1</sup>, John S. Mitcheson<sup>1</sup>, Peter Gedeck<sup>2</sup>, Michael J. Sutcliffe<sup>3</sup>.

<sup>1</sup>*Department of Cell Physiology and Pharmacology, University of Leicester, University Road, Leicester, LE1 7RH, UK*

<sup>2</sup>*Novartis Horsham Research Centre, Wimbleshurst Road, Horsham, West Sussex, RH12 5AB, UK*

<sup>3</sup>*School of Chemical Engineering and Analytical Science (CEAS), University of Manchester, The Mill, PO Box 88, Manchester, M60 1QD, UK*

The characterisation of a novel compound’s safety profile is an essential aspect when developing medicinal drugs. Recently, the risk of unintentional drug block of the human ether-a-go-go related gene (hERG) potassium (K<sup>+</sup>) ion channel has become a major safety concern, due to the chance of lethal cardiac arrhythmias. It is the central cavity that drugs bind to, which lies in the conduction pathway of hERG. This binding may prevent normal K<sup>+</sup> ion flow and consequently reduce the rate of cardiac myocyte repolarisation, during the ventricular action potential. A number of residues in the central cavity region of hERG have been identified, by site-directed mutagenesis, to be critical for high affinity drug binding. These include two aromatic amino acids, Tyr652 and Phe656, which are deemed to be critical and are specific to the EAG-family of K<sup>+</sup> channels, and a further four residues, Thr623, Ser624, Val625 and Gly648, which highly conserved throughout the K<sup>+</sup> channel family. This study aims to further characterise the binding site within hERG by using homology modelling, ligand docking and molecular dynamics techniques, to facilitate the analysis and

visualisation of interactions between the drugs and central cavity. The ultimate goal of the study is to produce an accurate high throughput computational method that permits the recognition of compounds liable to bind to hERG, at an early stage of development. Homology models of hERG were created, describing the predicted open (based on KvAP and MthK) and closed (KcsA & KirBac1.1) states of the channel pore domain (S5-S6). Ligand docking techniques were then used to predict binding modes and scores of a total of 25 ligands within the models, with the output compared to the known mutagenesis and IC<sub>50</sub> data. Based on the results of the ligand docking, the homology models were further refined, primarily to improve correlation with the mutagenesis. To investigate the homology models and drug binding in greater detail, molecular dynamics simulations were performed to reproduce the intrinsic motions of the channel and as an alternative, potentially more accurate, method for predicting drug affinities. Furthermore, homology models of other hERG domains have also been developed, including the voltage sensor domain (based on the isolated KvAP voltage sensor and Kv1.2 crystal structures) and C-terminal domain (based on HCN2).

---

## **Modelling drug metabolism in cytochrome P450 enzymes**

Jolanta Zurek, Jeremy Harvey, *Adrian Mulholland*

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

Cytochrome P450 enzymes are a group of haem-monoxygenases, which are involved in the biosynthesis of various regulators and oxidation of xenobiotics. These enzymes are responsible for the metabolism of the majority of drugs, catalysing a range of stereospecific activations of C-H bond. In addition to a broad specificity, P450 enzymes show genetic polymorphism, which can lead to altered enzyme activity or specificity, resulting in different effects of drug administration. It is therefore of great interest to be able to predict interactions of various therapeutics with these enzymes. Better understanding of drug-P450 interactions and reactions should be useful in the development of new drugs, for example in predicting drug metabolism and its dependence on polymorphisms.

We have used state-of-the-art methods to study the reaction of hydroxylation of ibuprofen, catalysed by human P450 CYP 2C9. High level combined quantum mechanics/molecular mechanics (QM/MM) calculations have been used to model the reaction in the enzyme, using hybrid density functional techniques (B3LYP/CHARMM). The reaction pathways have been refined with the nudged elastic band technique. The electronic state of the reactants, products and transition state has been characterised, as well as their geometries.

A reactive and unreactive mode of ligand approach has been identified. These geometric descriptors may be useful in docking.

These results give insight into atomic-level details of P450 mechanism.

---

## **Online Drug Design**

Alasdair T. R. Laurie, Richard M. Jackson

*Institute of Molecular Cell Biology, University of Leeds, Leeds, LS2 9JT*

The Online Drug Design system is a computational method for generating potential lead compounds. It comprises of three interlinked tools: (1) binding site prediction (Q-SiteFinder) (2) virtual screening (Q-fit and FlexLigDock) and (3) lead optimisation (Q-LeadFinder).

Q-SiteFinder uses the interaction energy between the protein and a van der Waals probe to locate energetically favourable binding sites (Laurie & Jackson, 2005). The success rate of Q-SiteFinder (90%) is higher than a pocket detection algorithm, Pocket-Finder, tested on the same data set. Additionally, Q-SiteFinder is twice as effective as Pocket-Finder in generating predicted sites that map accurately onto ligand coordinates. The predicted binding sites can be exported into Q-fit and Q-LeadFinder.

Q-fit is a rigid docking tool that employs a geometric hashing method to generate different orientations of a ligand in a binding site (Jackson, 2002). GRID forcefield parameters are used to estimate binding affinities and score these orientations. FlexLigDock is a flexible docking algorithm based on Q-fit (Oledzki *et al.*, unpublished data). The web-based interface allows energetic analysis and spatial comparison of docked ligands. Ligands can be exported to Q-LeadFinder for optimisation.

Q-LeadFinder is a lead optimisation tool that enables both fragment-based and atomic mutations to improve the binding affinity of a bound ligand. The ligand can be derived from a crystal structure, or it can be docked into the protein. Binding energies are calculated using GRID forcefield parameters. Lipinski's Rule of Five parameters are calculated to assess drug-like properties (Lipinski *et al.*, 1997). The process is human-directed and has not yet been comprehensively tested or benchmarked. An automated version is currently in development.

Laurie, A. T. R. and Jackson, R. M. (2005). Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, **21**, 1908-1916.

Jackson, R. M. (2002). Q-fit: a probabilistic method for docking molecular fragments by sampling low energy conformational space. *J Comput Aided Mol Des*, **16**, 43-57.

Oledzki, P., Lyon, P. and Jackson, R. M. FlexLigDock, a flexible ligand docking tool. *Unpublished data*.

Lipinski, C. A., Lombardo, F., Dominy, B. W. and Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*, **23**, 3-25.

---

## **Molecular Surface Property Graphs**

Vishwesh Venkatraman, Dr. David Whitley, Prof. Martyn Ford

*Centre for Molecular Design, University of Portsmouth, King Henry Building, King Henry I Street, Portsmouth, Hampshire, PO1 2DY*

The presentation focuses on the use of topological methods to describe molecular surface properties. The basic idea is to transform geometrical concepts of similarity into a flexible notion of equivalence.

Starting from a level set description of the molecular surface, the aim is to characterise the behaviour of a scalar molecular surface property (molecular electrostatic potential, molecular lipophilicity potential etc.) in terms of a directed graph derived from its gradient vector field. This graph is formed by the trajectories of the vector field, linking the critical points (sources, saddles and sinks) and is called a *molecular surface property graph* (MSPG). The issue of property comparison can then be reduced to a comparison of these graphs, which can be addressed as a common subgraph isomorphism problem.

Typical MSPGs are locally stable, unaltered by small conformational changes. Larger variations may alter the graphs, corresponding to gradient field bifurcations that change the number of critical points and rearrange the edges.

These graphs provide an effective representation of both shape and property with potential applications in

- i. Database searching for molecules with similar properties.
- ii. Classification of property-based conformational shapes.
- iii. QSAR analysis using topological indices derived from the MSPGs.
- iv. The investigation of ligand-receptor complementarity.

An algorithm to determine the graphs will be outlined and examples of how MSPGs can be used in similarity analysis will be presented.

References:

Whitley, D. (2002). *Molecular Surface Property Graphs*, *Euro QSAR*. Bournemouth, UK, 428-430

---

### **Where do the protons go: Protonation states in protein-ligand complexes**

Paul Czodrowski<sup>1</sup>, Ingo Dramburg<sup>2</sup>, Christoph A. Sotriffer<sup>1</sup>, Jens Erik Nielsen<sup>3</sup>,  
Gerhard Klebe<sup>1</sup>

<sup>1</sup>*Department of Pharmaceutical Chemistry, Philipps University Marburg, Marbacher Weg 6, 35032 Marburg, Germany*

<sup>2</sup>*BioSolve IT GmbH, An der Ziegelei 75, 53757 Sankt Augustin, Germany*

<sup>3</sup>*School of Biomolecular and Biomedical Science, University College Dublin, Dublin 4, Ireland*

Protonation states of amino acid residues in proteins can be significantly modulated by ligand binding. The major contribution to changes in protonation states comes

from electrostatic interactions which can be modelled by continuum electrostatics methods based on solving the Poisson Boltzmann (PB) equation. A crucial initial step for PB calculations is the assignment of partial atomic charges. Currently, no generic charge model applicable to both, ligand and protein, is available for the use in pKa calculations.

We therefore developed an optimized form of the established Gasteiger-Marsili charges [1]: They were re-parameterized to describe solvation free energies of polar amino acids as well as small organic molecules. Applied to proteins for which experimental pKa values have been recorded a convincing agreement between computer simulation and experiment could be achieved. This stimulated us to apply our charge model to protein-ligand complexes.

Examples will be presented for protein-ligand complexes with experimentally observed changes in protonation states. These examples cover the serine proteases trypsin and thrombin, the aspartic protease HIV protease, and the oxidoreductase dihydrofolate reductase. As indicated by isothermal titration calorimetry these complexes either remain with unchanged protonation states or they pick-up or release protons upon binding. Our computer simulations succeed to correctly reproduce the experimental observations.

Finally, a new tool for the computation of protonation states in protein-ligand complexes is developed as an open source project. It detects automatically titratable groups for ligands and initiates straight-forward pKa calculations of protein-ligand complexes based on the newly optimized Gasteiger-Marsili charge model.

References:

[1] Gasteiger, J., Marsili, M. Iterative Partial Equalization of Orbital Electronegativity - A rapid Access to Atomic Charges, *Tetrahedron* (1980) **36**, 3219-28.

---

## **The Role of Water in Protein-Ligand Interactions: Implications for Rational Drug Design**

Caterina Barillari, Russell Viner, Jonathan W. Essex

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

Water molecules play a crucial role in mediating the interaction between a ligand and a macromolecular receptor. Some water molecules located in the binding pocket of the protein can be displaced upon ligand binding and this process is favoured by a gain in entropy of the system; other water molecules can stabilize the complex by acting as a bridge between the two components.<sup>1</sup> If we knew *a priori* which waters are likely to be displaced by the ligand and which ones are instead tightly bound to the protein, molecules could be designed specifically to displace only targeted waters and interact with the others, thus increasing the binding affinity for the protein.

The aims of this work were to study water molecules mediating the interaction between proteins and their inhibitors from a thermodynamic point of view and to seek a relationship between the binding free energy of the waters and their microenvironment. For this purpose, six proteins, each in complex with different

ligands, were selected using Relibase+<sup>2</sup> and a total of 51 water molecules were studied. For each water molecule, the absolute binding free energy was calculated using the double decoupling method.<sup>3</sup> The free energy methodology used was Replica Exchange Thermodynamic Integration (RETI)<sup>4</sup> in Monte Carlo simulations.

It was found that conserved water molecules are generally more tightly bound than those which can be displaced by ligands.

Simple molecular descriptors characterizing the water environments were then calculated and they were used to build statistical models for the prediction of the binding free energy of waters.

References:

1. C.Poornima, P.Dean, J. Comp. Aid. Mol. Des. (1999), **9**, 500.
2. J. Gunther, A. Bergner, M. Hendlich, G. Klebe, J. Mol. Bio. (2003), **326**, 621
3. M.K. Gilson, J. A. Given, B. L. Bush, J. A. McCammon, Biophysical J. (1997), **72**, 1047
4. C.J.Woods, J.W.Essex, M.A.King, J. Phys. Chem. B (2003), **107**, 13703

---

### **Investigation of the charge transfer mechanism in 9-mesityl-10-methylacridinium : Rationalizing electron transfer in perpendicular systems.**

Ben Allen, James Rostron, Peiyi Li, *Anthony Harriman*

*Molecular Photonics Laboratory, Bedson Building, University Of Newcastle Upon Tyne, Newcastle Upon Tyne, NE1 7RU*

According to Marcus Theory, for efficient photoinduced charge transfer (CT) to occur, a system requires strong electronic coupling between the initial and final states. Recently however, it has been reported<sup>1</sup> that a small molecular dyad with a perpendicular geometry, 9-mesityl-10-methylacridinium (mes-ac), exhibits not only rapid formation of a CT state upon photo-excitation, but also that this state has an exceptionally long lifetime, especially at low temperatures. While the interpretation of some of the results in the original report are still being discussed<sup>2</sup>, it seems clear that the initial CT state is formed with reasonably high efficiency, certainly far quicker than would be expected for a molecule of this type. CT is believed to occur by excitation of the acridinium moiety into a locally excited singlet, then subsequent transfer of an electron from the mesityl group to the acridinium, resulting in the formation of a singlet bi-radical species.

The present investigation is focussed on using theoretical techniques to explain the apparently anomalous high rate of CT. To this end, several computational techniques have been employed, of which molecular dynamics and semi-empirical C.I. have proved to be most illuminating. The optimised mes-ac geometries of the various electronic states show that the initial photo-accessible state  $S_1$  is a local  $\pi \rightarrow \pi^*$  transition located on the acridinium moiety. The  $S_2$  state shows marked charge transfer character, but lies between 0.2 eV and 0.6 eV above  $S_1$  (depending on the

state geometry chosen). Therefore, on purely energetic grounds, it seems unlikely that CT will occur in this molecule. However, molecular dynamics simulations of mes-ac at room temperature show that there are certain geometries at which the local  $\pi \rightarrow \pi^*$  and CT states 'flip' in energy, enabling electron transfer to theoretically take place.

While for some geometries the CT reaction is energetically feasible, the coupling elements at these configurations are prohibitively small. However, an analysis of the vibrational frequencies shows that for certain frequencies, proceeding along the normal coordinate raises the energy of the HOMO of the donor, until it is degenerate with the lower SOMO of the acceptor moiety, with both the HOMO and SOMO extending over the donor and acceptor moieties, resulting in large orbital overlap, and therefore strong coupling. This massive increase in the coupling between the  $S_1$  and  $S_2$  states and provides a possible mechanism for the CT process observed in mes-ac.

#### References:

**1:** Fukuzumi, S.; Kotani, H.; Ogo, S.; Tkachenko, N.V.; Lemmetiyen, H; J. Am. Chem. Soc. 2004, 126(6), 1600

**2:** Benniston, A. C.; Harriman, A.H.; Li, P.; Rostron, J. P.; Verhoven, J. W.; Chem Comm; 2005, 21, 2701

---

### **$\pi$ -stacking in aromatic structures: a DFT and Atoms in Molecules study**

Arturo Robertazzi, Mark P. Waller, *James A. Platts*

*Department of Chemistry, Cardiff University, Cardiff, Cf10 3TB, UK*

Intermolecular forces such as Hydrogen Bonding and  $\pi$ -stacking play a fundamental role in assembling the structure of nucleic acids. While an ample variety of computational tools are available for describing Hydrogen Bonding,  $\pi$ -stack interactions present major difficulties. In particular, post-HF methods such as MP2 and CCSD are needed to characterise such forces. However, the computational resources required for these methods increase rapidly with molecular size and hence are practically limited to small model systems. Therefore, there is considerable demand for a computationally efficient electronic structure method capable of describing  $\pi \dots \pi$  stacking. Here we propose a new method based on Becke's half-and-half functional, along with Atoms in Molecules (AIM) theory, in order to elude such obstacles. The hybrid functional has been tested for more than 50 molecules, including substituted benzenes, pyridines and DNA bases. These data are compared to MP2 and CCSD calculations wherever available, confirming that BHandH qualitatively reproduces high-level results. AIM methods were developed to decompose total binding energy into contributions from Hydrogen Bonding and  $\pi$ -stacking, allowing us to recognize the source stabilisation in several DNA adducts of interest. These methods will be applied to intriguing cases like realistic DNA models and metal-DNA adducts.

---

### **Modelling Aromatic Interactions via Molecular Mechanics Simulation**

Christopher M. Baker<sup>1,2</sup> and Guy H. Grant<sup>2\*</sup>

<sup>1</sup> *Department of Chemistry, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford, United Kingdom*

<sup>2</sup> *Unilever Centre for Molecular Informatics, University Chemical Laboratory, Cambridge, United Kingdom*

The role of weak intermolecular interactions in molecular, and particularly biomolecular, systems is of growing interest. Perhaps the most versatile of the groups involved in such interactions are aromatic rings, which can participate in favourable interactions with both cations and anions, and act as hydrogen bond donors or acceptors. These interactions are known to be important in a number of biological situations, including neuroreceptors, GPCRs, transporters and enzymes.

The rise in the appreciation of the importance of aromatic interactions has not, however, been matched by an increase in the sophistication of computational methods used to model these interactions. Within the framework of molecular mechanics simulation, molecules are typically represented as a collection of point charges centred on the nuclear sites. Whilst this representation is adequate for many molecules, reproducing for example correct dipole moments, it fails in the case of aromatic molecules. Within the point charge representation of benzene, partial charges are placed on the carbon and hydrogen positions. The individual bond dipoles created by this charge distribution will sum to give the required zero dipole moment but this ignores the fact that benzene has an anisotropic charge distribution that results in a molecular quadrupole moment. And it is this quadrupole that determines the correct interaction geometry for the aromatic ring.

In an attempt to resolve this problem, Hunter and Sanders<sup>1</sup> proposed a simple model that accounted for the non-atom centred charge distribution within aromatic molecules by separating the charges into separate contributions for the sigma and pi systems. In the present work, this model is adapted for the treatment of liquid benzene via Monte Carlo simulation and, after careful parameterisation, is shown to reproduce both thermodynamic and structural data more accurately than comparable all atom force fields.

The newly developed model is then applied to the specific case of protein-ligand binding in acetylcholinesterase. Preliminary results are encouraging but suggest that there is room for further improvement of the model.

#### References:

<sup>1</sup> Hunter, C. A., Sanders, J. K. M., *J. Am. Chem. Soc.*, (1990), **112**, 5525