

# Joint Departmental / ISMB Retreat 19-20 June 2007

## Write-up of the event

*This article was written by Dr. Clare Sansom and Dr. Halina Garavini.*

Following the success of the first Institute of Structural Molecular Biology retreat, held in Hinxton in June 2005, a second retreat was held almost exactly two years later. It took place at the Wellcome Trust Conference Centre in Hinxton, Cambridgeshire - which shares a campus with the Sanger Centre, where a third of the Human Genome was sequenced, and the European Bioinformatics Institute - on June 19 and 20, 2007. The packed, two-day-long programme included keynote lectures from three distinguished scientists, posters and oral presentations from students and postdoctoral researchers at Birkbeck College and University College, London, and an interesting session on careers in science.



*The Wellcome Trust Conference Centre in Hinxton*

The retreat opened with a Welcome address by **Professor Gabriel Waksman**, head of the Institute of Structural Molecular Biology and head of both the School of Crystallography at Birkbeck and the Department of Biochemistry and Molecular Biology at UCL. **Dr. Finn Werner (UCL)** then introduced the first keynote speaker, **Dr. Stephen Bell** from the **MRC Cancer Cell Unit in Cambridge**.

Stephen Bell spoke about “DNA replication in the third domain of life”. He has been investigating the DNA replication system in archaea, specifically in the species *Sulfolobus solfataricus*, a hyperthermophile that survives at temperatures of 80°C. Archaeal replication machinery is a simplified form of that found in eukaryotes, yet it provides a good model because archaea have a eukaryotic type of cell cycle. Like eukaryotes, some archaea have multiple origins of replication on their chromosome.

*Sulfolobus solfataricus* was found to have 2 origins of replication by 2-dimensional gel analysis. A third origin was then identified by marker frequency analysis, and the same technique, combined with computational modeling, suggested that all 3 origins fire synchronously in all cells, and that all three are used in each cell cycle. The origin binding proteins in archaea are homologues of the related eukaryotic Orc1 and Cdc6 proteins, and also contain an AAA+ domain.



*Dr. Stephen Bell from the MRC Cancer Cell Unit in Cambridge*

A large part of the afternoon was devoted to presentations by PhD students and post-docs. The session was begun by **Mr. Andrew Niewiarowski (UCL, Biochemistry and Molecular Biology)**. The title of his presentation was, "TIP48 and TIP49: united in mitosis or poles apart?" His project concerns the two homologous proteins, TIP48 and TIP49, which are ATPases that are individually essential in all eukaryotes. They are found together as components of multi-protein nuclear complexes and are implicated in a variety of cellular processes, including DNA repair, transcription regulation and chromatin remodeling.

Andrew has used the techniques of size exclusion chromatography and analytical centrifugation to study oligomerisation by the two proteins. His work so far suggests that TIP48 hexamerises in the presence of ATP, whereas TIP49 remains monomeric. It is also known that, although both proteins are involved in mitosis and cytokinesis, they have different localisations on the mitotic spindle and may operate independently during mitosis. Andrew's project aims to further investigate the mitotic roles of TIP48 and TIP49, and to search for an underlying mechanism that might connect the varied processes in which the two proteins are essential.

---

The second presentation was by **Dr. Stephanie Hunter (UCL, Biochemistry and Molecular Biology)**, and was entitled

"Analyzing the oral metagenome". The aim of her project is to investigate the microflora of the oral cavity, which contains more than 800 bacterial species. Some 50% of these oral species cannot be cultured, so non-culture techniques such as molecular approaches must be used instead. Bacteria adhere to various oral surfaces and also to each other to form biofilms, and this is a key process in the initiation of bacterial disease. An understanding of adhesion mechanisms would provide insight into how bacteria attach to human tissues and cause disease.

Stephanie is making a metagenomic study of the oral microflora, which is a study of all the genomes in this bacterial community. Her method is to make phage display libraries. The process involves finding an appropriate DNA extraction method, and also a method for removing human DNA from bacterial samples, either by selective lysis of human cells or by differential centrifugation. Whole genome amplification of the bacterial genome is then performed, and phage display libraries are constructed by inserting bacterial DNA fragments into a phagemid, which allows the bacterial protein to be expressed on the surface of the phage host. The library can be screened for peptides and protein domains that bind to known ligands which can be linked to the DNA sequence. Fosmid libraries can be used to complement the phage display. So far, conditions have been optimised for the production of representative phage display libraries, and samples from 60 patients are being collected and investigated.

---

The next presentation was given by **Miss Asvi Francois (UCL, Biochemistry and Molecular Biology)**, and was entitled, "Flavin-containing monooxygenases: genetic polymorphisms and their role in drug metabolism". Flavin-containing monooxygenase (FMO) proteins are membrane-bound enzymes in the endoplasmic reticulum. They are used in detoxification reactions and also metabolise

drugs. Rarely, they can take part in bioactivation reactions which produce harmful compounds such as thioicarbamides and thioureas.

Five members of the FMO family are known to metabolise drugs: FMO1,2,3,4 and 5. FMO2 is found in the lung in several mammalian species, but in humans, specifically in Asians and Caucasians, there is a mutation in the FMO2 gene which causes premature truncation. However, 25% of the population of sub-Saharan Africa have a wild type FMO2 gene and produce a catalytically active protein. Asvi's project provides evidence that thiacetazone, a cheap and commonly prescribed anti-tuberculosis drug in Africa, is a substrate for FMO2. Thiacetazone can have severe side effects and produce liver toxicity. Individuals who express FMO2 are at greater risk, and this finding has implications for the prescribing of anti-tuberculosis drugs to susceptible individuals in Africa.



*Student and Postdoc talks in the Francis Crick Auditorium*

Next to speak was **Mr. Osman Salih (Birkbeck, Crystallography)**, on the subject of "Structural analysis of the Saf pilus by electron microscopy and image processing". Bacterial pili are fibrous protein organelles, which are important in pathogenesis because they are involved in initiating attachment to host cells. Osman's study has focused on the *Salmonella* atypical fimbriae (or *Saf* pili) in *Salmonella typhimurium*, using negative stain electron microscopy and

single-particle image analysis to determine their 3-dimensional structure.

The aim of Osman's project was to study the organisation of the subunits in an atypical fibrillum. The major pilus subunit is known as SafA. Visualisation of pili by electron microscopy has shown that the subunits are globular structures arranged in a row, linked to each other through short thin linker regions. The linker regions are highly flexible, and allow a large range of movement between consecutive subunits in the fibre. The results provide an explanation for the flexibility of atypical fibrillae, and therefore explain why the fibrillae appear unstructured on bacterial surfaces.

The final presentation was that of **Mr. Tjelvar Olson (UCL, Biochemistry and Molecular Biology)**, and was entitled, "Structure, free energy, enthalpy and entropy in protein-ligand interactions: a database approach". Tjelvar has been investigating protein-ligand interactions experimentally by isothermal calorimetry, and has created a database, Scorpio, of the experimental data on 85 ligand-protein complexes. The structures have been analysed in terms of buried apolar and polar surface area, and the structural parameters have been correlated to the Gibb's free energy, the enthalpy and the entropy of binding.

The commonly-held belief is that there is a relationship between buried apolar surface area in a protein-ligand complex and the entropy. However, Tjelvar's analysis controversially shows no such general relationship, but instead a significant correlation between buried apolar surface area and the Gibb's free energy. He used multiple linear regression to arrive at a minimal adequate model of the binding interactions, and apolar surface area has been found to contribute favourably to  $\Delta G$ . It has also been found that the buried polar surface area remains relatively constant for a wide range of ligand sizes.

The afternoon finished with a session on Careers in Science. Four speakers presented their individual perspectives and experience. **Dr. Cara Vaughan (ISMB Lecturer in Crystallography)** spoke of her career path to getting a lectureship. **Dr. Snezana Djordjevic (UCL, Senior Lecturer in Biochemistry)** spoke of what is involved in being a university lecturer. **Dr. Renos Savva (Birkbeck, Crystallography)**, who is a co-founder of and scientific advisor to the successful company Domainex Ltd., gave a different perspective of working in industry, and finally **Ms Kate Travis (Contributing Editor, ScienceCareers.org)** spoke of her experience and career in science communication. The presentations were followed by a lively session of panel discussion and questions from the audience, with a lot of interest shown by students and post-docs in alternatives to the "traditional" career path of a researcher in a university, the public sector or industry.

\*  
\* \*

Wednesday's sessions started with the second keynote speaker, **Trevor Smart** from the **Department of Pharmacology at University College, London**. Professor Smart, who is a Fellow of both the Royal Society of Pharmacology and the Academy of Medical Sciences, described his work to elucidate the function and mechanism of the GABA<sub>A</sub> receptor as "trying to solve a structural issue, in the absence of structural data or crystals". GABA receptors are ion channels, secreted by neurons and present in 30-35% of synapses in the brain. Their function is to control the excitability of nerve cells, by opening to allow the passage of chloride ions into the cell in response to the neurotransmitter gamma-amino butyric acid (GABA). Increased chloride conductance inhibits nervous system activity, so agonists at this receptor - molecules that replicate the action of GABA - may be good drugs for central nervous system (CNS) conditions including epilepsy, simple anxiety and alcohol and drug abuse.



*Professor Trevor Smart  
from the Department of  
Pharmacology at  
University College,  
London*

GABA receptors and related ion channels are membrane-bound structures consisting of pentamers of similar subunits, each subunit forming a characteristic bundle of four alpha helices. The channel forms in the centre of the structure, bounded by the second helix of each subunit. Structural studies of membrane-bound proteins are notoriously difficult, and there are still no atomic-level structures even of individual subunits. Smart and his colleagues have therefore turned to molecular biology to probe the mechanism of these receptors further and identify amino acid residues involved in binding GABA and neurosteroid agonists. His group first examined structures of known steroid binding sites and identified a "canonical" binding pocket that was likely to be replicated in these structures. They also identified a similar receptor in *Drosophila* that responded to GABA but was not affected by neurosteroids, and constructed chimeric receptors from parts of the *Drosophila* and human sequences to identify residues involved in steroid binding and, separately, receptor activation. These experiments identified one residue, thought to be in the ion channel lining, that is involved in hydrogen bonding to steroids, and revealed that the human GABA<sub>A</sub> receptor has two binding sites, one "potentiating" the receptor and the other activating it, so steroid binding to both sites is necessary for full receptor activity. These results are an important step towards a full understanding of the mechanism of these receptors and their response to drugs for some important diseases.



The next session featured three student and post-doc presentations concerning applications of bioinformatics. The first of these was given by **Eleni Rapsomaniki**, a third-year Ph.D. student working in Dr. Adrian Shepherd's bioinformatics group in the **School of Crystallography at Birkbeck**. She described a comparative analysis of the methods used to study protein-protein interactions and the databases in which those interactions are stored. These high throughput methods, which include yeast two-hybrid and co-immunoprecipitation, are notoriously inaccurate, with a high false-positive rate. Eleni mined a number of widely used databases of these interactions, compared the full set of reported interactions between proteins in seven proteomes with literature and bioinformatics evidence and used the statistical technique of imputation to "fill in" some missing data. Not unexpectedly, her results showed that the databases with the highest coverage had the lowest specificity, i.e. they had the highest rate of false positives. More surprising was the finding that interactions stored in several databases were no less likely to be false-positives than those stored in a single one.

---

**Stathis Sideris**, who works with Paul Kellam and Christine Orengo in the **Department of Biochemistry and Molecular Biology, University College London**, gave the second bioinformatics talk. Clustering is an important statistical technique for the identification of genes that have similar expression profiles and that may, for example, be co-expressed, from microarray experiments. However, there are many different clustering methods, which may give inconsistent results, and it is not clear which is the most appropriate to use. Gene Ontology (GO) is a well-known pseudo-hierarchical database of terms describing the function of gene products. Stathis validated clustering algorithms by calculating the "semantic distance" between the annotated functions of clustered genes, taken from the Gene Ontology database. The most useful results were obtained by using the more specific

terms within GO; generic terms linking a gene to a process such as "sensory perception" or "immune response" gave poor results.

---

The next talk was from **Mark Halling-Brown**, a post-doc in David Moss' group at **Birkbeck**, who gave a short overview of the EU Framework 6 project, Immunogrid, which funds his work. This ambitious project aims to develop a "clinically relevant systems biology model of the human immune system" that can simulate immune system processes at the levels of organs, cells, and molecules. The immune system is diverse and extremely complex, and operates on many levels. Many groups within Europe, including the Birkbeck group, have used models to simulate parts of this process, including interactions between molecules and cells. The task of the Immunogrid consortium is to introduce molecular level detail into the developing cellular and organ-scale models, by reading both experimental and simulated data from a database repository. These calculations are computationally intensive, and will become more so, and so a Grid implementation will be necessary.

---

**Satpal Virdee**, a PhD student working with Gabriel Waksman, also at **Birkbeck**, described a detailed structural and biophysical analysis of SH2 (Src homology 2) domains. These are fairly short, conserved protein domains found in many different eukaryotic proteins involved in signal transduction; their function is to recognise, and bind to, phosphorylated tyrosine residues. They have been very well studied by structural and biophysical methods; Waksman was one of the first to elucidate their structure. Satpal and co-workers used a semi-synthetic method, involving expressed protein ligation and native chemical ligation, to build an SH2 domain from three constituent fragments. Mutations could be introduced in any fragment, and the central fragment could be synthesised containing non-standard amino acids. The group is currently having problems with

yield, but is trying to solve this problem by introducing cysteine-serine mutants.

---

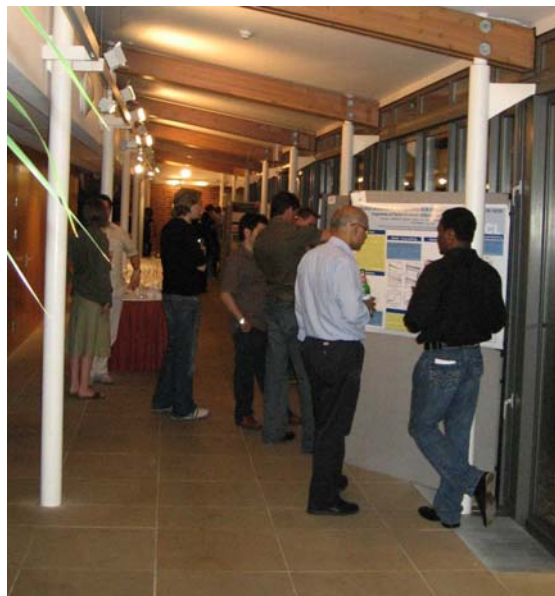
The final two talks from students and post-docs came from the same lab: Derek Macmillan's group in the **Department of Chemistry at University College**. **Silvia Marchesan**, who has recently joined the group, presented work carried out in her previous position at the University of Edinburgh. Mannosyltransferases are enzymes that transfer the monosaccharide, mannose, onto glycopeptides or glycolipids, and they are essential for the synthesis of the common core of N-linked glycoproteins. Silvia and her co-workers have developed a method of "engineering" mannosyltransferases to recognise unnatural substrates, including residues that incorporate an azido ( $-N_3$ ) group. This modified monosaccharide can be recognised very quickly and easily using a simple colorimetric screen, and therefore should aid the development of a simple non-radioactive probe for the activity of the enzyme that catalyses its addition to a peptide or lipid group.

---

Finally, **Jonathan Richardson**, also from Derek Macmillan's group at UCL, described the semi-synthetic synthesis of the human glycoprotein, erythropoietin (EPO). This is a small protein with four attached carbohydrate groups, secreted from the kidney, and its function is to stimulate the production of red blood cells. It can therefore be used as a drug, to treat anaemia. It is currently produced from cultured human CHO cells, but this method is very expensive and it is difficult to keep up with demand. Jonathan and his colleagues have developed a method of synthesising human EPO from two fragments: a chemically synthesised N-terminal one containing acetylated cysteine residues, and a bacterially expressed C-terminal one. This produced a protein that, unlike EPO produced from culture, can be decorated with homogeneous glycan residues. This semi-synthetic protein has not

yet been fully tested, but if it proves to be both active and non-immunogenic, the procedure will provide a reliable method of producing this vital biological product.

---



*Poster session in the Cloisters*

Fifty-eight posters were presented at the meeting, all by students and post-docs from the departments involved in the ISMB: Crystallography and Biology/Chemistry at Birkbeck, and Biochemistry and Molecular Biology, Chemistry, and Biology, at University College London. They covered a wide range of topics within the disciplines of chemistry, biochemistry, molecular and structural biology, and bioinformatics. Work presented in the poster session included, from Birkbeck, biophysical studies of the KlenTaq DNA polymerase (PhD student **William Allen**) and the assembly of a complex between Braf35, a protein associated with the breast cancer susceptibility protein BRCA2, and the kinesin KIF4 (PhD student **Matthew Webster**). Highlights from UCL included a novel method of protein function prediction (PhD student **Anna Loble** from the Department of Computer Science) and structural studies of peroxidases from the important protozoan parasite *Trypanosoma cruzi* (post-doc **Syeed Hussain**) and of the regulation of human pyrophosphatase 2 by protein methylation (PhD student **Meng-Lin Tsai**).

---



*Professor Carol Robinson,  
from the School of Chemistry  
at the University of  
Cambridge*

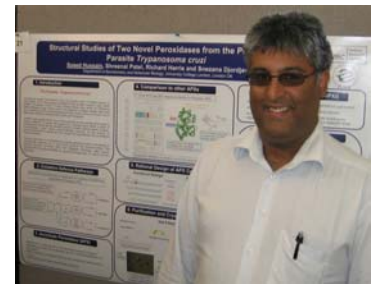
The final keynote lecture was given by **Carol Robinson**, a professor in the **School of Chemistry at the University of Cambridge**. Her expertise is in the area of mass spectrometry, and she has developed techniques for maintaining the structural integrity of protein complexes within a mass spectrometer, enabling the elucidation of complex features of these large (at the molecular scale) structures. She first described a proof of the technique in principle, using the recombinant carrier protein, transthyretin, which functions as a 56 kDa dimer of dimers. Although it was important to prove that this could travel through the mass spectrometer as an intact tetramer, the technique did not reveal anything new about its structure. The TRAP (TNF-related activation protein) complex forms a ring-shaped 11-mer which, in turn, complexes with tryptophan and mRNA. In conventional electrospray mass spectroscopy, this complex dissociates into its constituent monomers. Robinson, however, was able to show the complex remaining intact. Without RNA bound, and in highly charged states, the ring-shaped structure could collapse, and Robinson's MS technique was able to distinguish between stable and collapsed rings. RNA binding prevented the collapse of the rings. Finally, she described the use of this technique to study the intact exosome - a complex "molecular machine" that degrades RNA - to identify each of the ten proteins making up this complex, and, by identifying the subunits present in sub-complexes, to determine which are in contact with which.

The last task of the retreat organisers was a very pleasant one: the award of ISMB Young Investigator Awards to the presentation of the best talk and the best poster. The judges admitted that the high quality of work presented made their decisions very difficult, but finally the prize for the best talk was awarded to **Jonathan Richardson**, for his elegant, semi-synthetic synthesis of erythropoietin, and the poster prize was awarded to **Syed Hussain**, for structural studies of *T. cruzi* peroxidases. It is particularly encouraging that both prizes were awarded for high quality basic research with important potential clinical applications.



*On left - ISMB Young  
Investigator Award for the  
best talk:  
Dr Jonathan Richardson*

*On right - ISMB  
Young Investigator  
Award for the best  
poster:  
Dr Syeed Hussain*



The retreat could not have been the success it undoubtedly was without the dedication of its scientific organising committee, and the invaluable behind-the-scenes work of the ISMB's able administrator, Anne-Cécile Maffat. On behalf of all ISMB students and researchers, we would like to thank them for their efforts, and we hope that the next retreat will offer a similar rich feast of science.