4th ISMB Symposium
17-18 June 2010

Write-up of the event

This article was written by Dr. Clare Sansom (Department of Biological Sciences, Birkbeck, University of London).

The Institute of Structural and Molecular Biology holds a research symposium every two years. Its fourth such symposium was held at one of its constituent institutions, University College London, on June 17 and 18, 2010. As before, the symposium brought together all researchers and postgraduate students of the five departments from UCL and Birkbeck College making up the ISMB together for a two-day lecture programme. The 2010 symposium, attended by all researchers and postgraduate students of the five UCL and Birkbeck departments making up the Institute, as well as many external scientists, featured twelve excellent talks. Three each were taken from four of the Institute's six disciplinary programmes: structural biology, biophysics, chemical biology and bioinformatics. Each three-lecture slot featured one speaker from the ISMB, one from the UK and another from overseas: the external speakers were invariably established research leaders, whereas most ISMB speakers were promising principal investigators nearer the beginning of their research careers.

In introducing the symposium, **Mary Collins**, the Executive Dean of the Faculty of Life Sciences at University College London, congratulated ISMB's director, **Gabriel Waksman**, on putting together "one of the highest quality structural biology meetings" that she had been to in a long time. Waksman, whose many honours include Fellowship of the Academy of Medical Sciences and membership of the European Molecular Biology Organisation, is the head of both the Department of Biological Sciences at Birkbeck and the Research Department of Structural and Molecular Biology at UCL.

The first programme featured structural biology. Its speakers may have been chosen to complement each other, with one representing each of the three main structural biology techniques: X-ray crystallography, nuclear magnetic resonance, and electron microscopy. The first to speak was X-ray crystallographer **Laurence Pearl**, who moved very recently from London's Institute of Cancer Research to head the **School of Life Sciences at the University of Sussex, UK**. Pearl has a long association with both Birkbeck and UCL as student, lecturer and professor. He described structural studies of a protein known as heat shock protein 90 (Hsp90), which is a molecular chaperone. The concept of chaperones is a very familiar one at Birkbeck, thanks to Helen Saibil's ground-breaking work on the bacterial chaperone GroEL; they are proteins that help

other proteins to attain and maintain their folded conformation. Hsp90, which enables protein stability at a late stage of folding, is essential in eukaryotes.

Pearl has studied the structure of this protein and its constituent domains since his time at UCL in the 90s. The first structure of its N-terminal domain, which binds ATP, was published in 1998. Since then, he has been establishing the mechanism by which ATP binding drives a conformational change that leads to Hsp90 forming a dimer. When ATP binds, a "lid" of protein structure folds down over its phosphate groups, opening up a hydrophobic surface that can be stabilised by dimer formation. The chaperone's substrate proteins bind to the dimer, and ATP can be hydrolysed into ADP, after which the protein relaxes back into its monomeric state. Substrates that bind to Hsp90 include many protein kinases, which are key targets for drugs against cancer; Pearl has published structural studies of Hsp90-kinase complexes in collaboration with Birkbeck electron microscopist Cara Vaughan. The last part of his talk was devoted to unpublished studies of Hsp90 from plants in complex with co-chaperones that are necessary for its activity, and he explained how these structures are leading, in the light of the ATP "switch" described earlier, to new insights into the chaperone's mechanism of action.

The ISMB contribution to the structural biology programme was given by **John Christodoulou**, who took up his position as a lecturer in the **Research Department of Structural and Molecular Biology at UCL** in 2007. He uses NMR spectroscopy to study the mechanism of protein folding on the ribosome. These NMR studies of a large, complex molecular machine in motion are all the more remarkable as, traditionally, NMR is seen as a technique most suitable for studying relatively small molecules, and as the emerging protein chain is still rarely visible even in X-ray structures of ribosomes.

Each ribosome is formed from one large and one small subunit, which only come together to form the intact structure during protein synthesis. The newly synthesised protein chain emerges through a tunnel in the large subunit, beginning to fold into its stable structure as it leaves the tunnel. NMR can give useful information about the dynamics of these structures, and hence about the process of protein folding in vivo. Christodoulu described how "fingerprint" NMR spectra obtained at different points during the protein synthesis process showed significant differences between the folded and unfolded states of the emerging protein chain. He has used this technique to investigate the folding of proteins involved in diseases of protein aggregation, such as alpha-synuclein, which is implicated in Parkinson's disease, and alpha-1 antitrypsin. He has obtained a high resolution structure for the stalk protein on the intact ribosome and compared this to X-ray structures of this protein in isolation. He is now developing cell-free methods for obtaining NMR samples of ribosomes that can generate high-quality data faster and more cheaply than conventional methods, and is aiming to use these to probe the mechanism of protein folding – one of the "holy grails" of structural biology – even more closely.

The name of **Wolfgang Baumeister** of the **Max-Planck Institute of Biochemistry, Martinsried, Germany**, is synonymous with technical developments in electron microscopy and, particularly, the development of electron tomography – imaging in sections - as a tool for understanding cell biology. This can produce relatively high resolution images of non-repetitive structures the size of cells in their natural environment.

Baumeister started his talk, however, with an overview of more recent work in his laboratory using single particle analysis. He used as examples two proteases involved in the protein degradation pathway, the proteasome, which chops unwanted proteins into peptides about 8-12 amino acids long, and the TPII protease, which breaks down those short peptides further. He can now obtain high quality structures of the proteasome, a cylindrical structure with a "cap" at each end, quickly and almost routinely, and is exploring its mechanism further with structural studies of

the complex bound to proteins that proteomics experiments have suggested as interaction partners.

Electron tomography explores the structures of proteins in the context of intact cells or organelles. Baumeister presented his studies of the nuclear pore complex through which proteins are transported into and out of the cell nucleus. Higher resolution structures of protein components of the pore complex have been docked into a lower resolution structure of a section across the intact complex to produce a full model. He is now turning his attention to the human ribosome, which is larger than those of lower organisms and less tractable to other structural techniques.

The second programme, on biophysics, was opened by **Sheena Radford** of the **Astbury Centre for Structural Molecular Biology and Institute of Molecular and Cellular Biology, University of Leeds, UK.** Radford's research, like Christodoulou's, focuses on protein folding, misfolding and disease, and she is no stranger to the ISMB, having established productive collaborations with Gabriel Waksman and Helen Saibil. Early protein folding theories proposed a single mechanism for folding each sequence into a single low-energy structure with a single function. Folding is now known to be much more complex, with folded and unfolded structures in some kind of equilibrium. Many if not most proteins, can form an amyloid-type structure consisting of ladders of beta sheets as well as their functional folds, and amyloid aggregation is a cause of many diseases. Understanding the process that drives proteins into either the folding or the aggregation pathway is crucial to understanding these diseases. Radford presented two pieces of work focusing on different parts of the aggregation pathway. One, which is not yet published, uses trapped intermediate structures of mutant proteins to identify and characterise the folding precursors of amyloid aggregates in beta-2 microglobulin.

Radford also presented her group's investigation of the structure and function of intact, fully assembled amyloid fibrils using infra-red spectroscopy and electron microscopy. She discovered that structurally identical fibrils of different lengths form from the assembly of globular subunits. The lengths and physical properties of the fibrils depend on the environment in which they form, with slower growth producing longer fibrils. Interestingly, she suggested that intermediates and short fibrils are more toxic to cells than the inert long fibrils; this insight may suggest new targets for intervention in amlyoid disease.

Mark Williams, a lecturer in the **Department of Biological Sciences at Birkbeck** and director of the **ISMB Biophysics Centre**, represented the "home team" in the biophysics section. His talk, intriguingly entitled "A Wet Walk in the Thicket of Thermodynamics", described some uses of thermodynamics - the study of energy conversion and heat changes - to aid the understanding of protein complex formation. Whenever molecules interact, by either covalent or non-covalent binding, heat is gained or lost. These gains and losses can be measured and monitored over time using a technique known as isothermal titration calorimetry (ITC).

Williams illustrated the use of ITC using examples from the Hsp90 chaperone system described by Pearl. The switch mechanism of Hsp90 is driven by a conformational change that closes a lid when ATP binds. Williams used the measurement of heat capacity when ATP and its analogues bind to answer the question of whether this alone is sufficient to close the lid, concluding that lid closure is at least partially driven by ATP. In collaboration with other ISMB researchers, he has set up SCORPIO, a database of the thermodynamics of about 250 different protein-ligand complexes. An analysis of these data shows that that the balance between enthalpic and entropic contributions to energy change is often very close, and that whereas the free energy of binding is strongly correlated with hydrophobic surface burial, this free energy is not clearly derived from the change in entropy to which it is traditionally attributed.

The final biophysics talk was given by **Steve Block** of **Stanford University in California, USA**. Block is a pioneer in the emerging field of single-molecule biophysics: studying proteins or nucleic acids, literally, one molecule at a time. Block's group has developed a novel technique known as the optical trap or optical tweezers, which can exert exceptionally small forces via focused laser light, and has deployed this to measure the tiny force that is caused by the actions of a single molecular machine such as RNA polymerase. RNA transcription is the key nexus of gene control; it is a highly regulated and extremely complex process, particularly in eukaryotes. In the double optical trap, the force arising from the motion of a single RNA polymerase molecule along the DNA strand draws two beads of the optical trap together, and the passage of the polymerase can be tracked by measuring this change. Reaching the precision necessary to measure the almost infinitesimal forces represented by the passage of a single DNA base has been an enormous technical challenge. The latest versions of the assay, however, can, indeed, measure the movement of 3.4 Å that occurs during the synthesis of a single base of RNA. Block's most recent work, however, has focused on the motions involved in the folding and unfolding of RNA secondary structures including simple hairpins and the natural RNA aptamers that control gene expression by selectively binding regulatory regions of genes. The double optical track has allowed him to track the folding of an RNA aptamer in real time, showing that folding and binding occur on the same timescale and that one RNA loop, the P2 loop, always forms first.

This section also included the final of a video competition for postgraduate students in ISMB. This was funded through the Roberts Fund for innovation in training for early-stage researchers. The four finalists all produced entertaining videos explaining their research to non-scientists. The winning video, chosen unanimously by a panel made up from the non-ISMB speakers, was made by Luke Goodsell, Maria Karyadi, Hélène Launay and Xiaolin Wang from John Christodoulou's lab, and showed each of them taking on the persona of one component of the ribosome and interviewing each other about their function. The second-placed video, made by Lamya Al-Haj and Linda Dekker from Saul Purton's group at UCL, explained the use of cyanobacteria in biofuels using a rap.

First to speak in the chemical biology section was **Ben Davies** from the **Department of Chemistry at Oxford University, UK**. Davies has been interested in carbohydrate chemistry and carbohydrate-protein interactions for many years; his work in this field has generated many honours, including the Wolfson award for Research Merit, and he is a founder of the biotech company Glycoform. He linked his presentation to many earlier ones through the idea that chemistry can be used as a tool to understand biology. By analogy with synthetic chemistry, the characterization, synthesis and artificial variation of biological molecules in "synthetic biology" can be used effectively in drug discovery.

At the beginning of the "genome era", the fact that the complex biology of higher eukaryotes can be generated from relatively few genes was considered a paradox. We now know that protein complexity can be generated post transcription and post translation. Glycosylation is one very common post-translational modification, and one protein may exist in many "glycoforms". Davies' group has developed methods to generate synthetic mimics of common glycosylations, for example by replacing an oxygen bridge by a disulphide. This "tap and modify" strategy allows for a wide range of molecules to be bound to proteins with exquisite specificity. He is exploring synthetic models for therapeutic proteins in collaboration with Glycoform, and has developed synthetic proteins that mimic specific protein-protein interactions, giving an antibody-like response.

Derek MacMillan from the **Department of Chemistry at UCL** continued the themes of glycosylation and complex biomolecule synthesis with a presentation describing a novel method for synthesizing peptide and protein thioesters. This reaction will cleave proteins at sites adjacent to cysteine residues within peptides and proteins, with the amount of thioester produced dependent on the amino acid preceding the cysteine in the sequence: cysteines immediately following a histidine residue are the most readily targeted; those preceded by another Cys or by Gly also give fairly good yields. MacMillan ended his talk by describing some unpublished work, in which he demonstrated that it is possible to synthesise cyclic analogues of anti-microbial peptides using this new chemical reaction.

Gregory Verdine, Irving Professor of Chemistry at **Harvard University and the Dana-Farber Cancer Institute, Cambridge, Massachusetts**, shares with MacMillan an interest in designing molecules to interact with intractable protein targets; he is the

founder and chair of several biotech companies. In an entertaining talk concluding the chemical biology section, he outlined the theory that a new approach to drug design is needed to access the full range of potential targets. Currently, for a protein to be "druggable" it needs either to be extracellular, and so accessible to biological drugs, or to contain a binding site suitable for small molecules; most proteins fit neither category. Verdine's work involves creating constrained and stabilized peptides that can mimic a partner protein's binding to a protein-protein interaction site. Many of these protein recognition elements are alpha helices, which can be constrained using covalent bonds or "staples" between adjacent turns of the helix; molecular dynamics simulations have shown them to be conformationally stable.

Verdine illustrated his theory by showing, firstly, that stapled peptides do, unlike the macromolecules from which they are derived, enter cells, and that they can be used to target particular organelles. He also demonstrated their potential in drug development using, as an example, the transcription factor NOTCH, which is mutated in a form of acute leukaemia known as T-ALL. The crystal structure of NOTCH with its interaction partners shows its binding site to be occupied by a long alpha helix. Taking this helical sequence, and mutating two residues not involved in interactions with NOTCH to form a staple, Verdine has been able to generate a stable, functional peptide that can repress NOTCH-driven DNA transcription in vivo and that may well have value as an anti-leukaemia agent.

The final symposium programme covered bioinformatics, and its first speaker was **Mark Sansom** of the **Department of Biochemistry, University of Oxford**. Sansom's long and distinguished research career has focused on modeling the structure and function of membrane-bound proteins, which are far less easily studied experimentally than soluble ones, using computational techniques including molecular dynamics simulation. This is an important topic in drug discovery as the 20% of protein products of the human genome that are membrane-bound make up 40% of all known drug targets.

Long molecular dynamics simulations are the most instructive, but these are very expensive in memory and CPU time. Sansom's group has developed a form of "coarse grain" modeling that speeds up calculations by grouping atoms together, switching back to fully atomic simulations when the system is stable. He has used this to model the self-assembly of proteins in a membrane, starting with a simple model system: the helix dimer of glycophorin. This was shown to assemble into its experimentally determined structure during the simulation, and the technique then applied to more complex systems, including the dimer assembly of epidermal growth factor receptor (EGFR) which has extracellular and intracellular domains joined by a single helix. The simulated structures of the glycophorin dimer and other systems including the M2 proton channel from the influenza virus have been shown to correlate well with experimental evidence.

The ISMB's contribution to the bioinformatics programme was given by **Andrew Martin** from the **Research Department of Structural and Molecular Biology at UCL**. Martin's work on the analysis of antibody sequences and structures has attracted the attention of pharmaceutical companies and he holds several patents. He began his talk, like some others, by illustrating the importance of his topic with a statistic: the total market in therapeutic antibodies will be worth some $30Bn in 2010. The large number of antibody structures known still represents only a fraction of antibody sequences, so modeling is extremely important, and standard automatic alignment and modeling methods may be inappropriate for these molecules. Several groups have published classifications and residue numbering schemes for human and mouse antibodies, and used them to generate alignments and profiles. With funding from pharmaceutical giant UCB Celltech, Martin's group has assembled all antibody sequence and structural information into a single database, Abysis, which now contains over 40,000 sequences and over 1,000 non redundant structures. This comprehensive database is being used in homology modeling and to test the extent to which any proposed therapeutic antibody resembles a human sequence, and so, whether it can be expected to be free of toxic side effects.

Andrej Sali of the **University of California, San Francisco** wound up the bioinformatics session and thus the symposium with a lecture on the integration of methods for characterizing macromolecular assemblies. He, like Pearl, started his scientific career in Bloomsbury; during his PhD studies with (now Sir) Tom Blundell at Birkbeck he developed the program Modeller, which is still one of the most widely used homology modeling methods.

Proteins function in vivo by assembling to form complexes or molecular machines. The proteome of any organism, however complex, will produce only a few thousand such assemblies, involved in a few hundred biological processes. Conversely, much structural information is available only for proteins in isolation. Sali described his method for combining information from many sources to yield assembly structures that are as detailed and as functionally correct as possible. This involves a four-stage scheme not dissimilar to that used for solving NMR structures. Firstly, data is collected from a variety of sources, and then the structure is modeled using both theoretical and experimentally derived restraints. This generates an ensemble of structures that fits the restraints as far as possible, and an analysis of the ensemble generates more restraints so the process can be iterated. The method is a general one that allows data from a wide range of structural studies to be incorporated if it is available. Sali cited, as an example, his partly unpublished work on the structure of the nuclear pore complex. This analysis combines low resolution data obtained using electron microscopy, protein localization, affinity purification and other techniques with atomic resolution data from such X-ray structures of the consistent nucleoporin sub-complexes as are available to derive models that are "as close as possible" to the complete structure of the whole pore, or, at somewhat higher resolution, of groups of proteins assembled into sub-complexes.

In summing up, **Gabriel Waksman** thanked all speakers and stressed the theme of molecular complexes and machines running through the four programmes. Almost invariably, biological macromolecules function in complexes rather than singly; this is increasingly becoming the focus of biological research and reflects the direction he would like the ISMB to take. Finally, **Nicholas Keep, Executive Dean of the Faculty of Science at Birkbeck**, thanked Waksman, stressing his irreplaceable contribution to the leadership of the ISMB and his indefatigable work as head of two departments and as a research leader and mentor.