

7th Bioinformatics Day on "Bioinformatics and Protein Science" April 20th 2006 in the Oxford Centre for Gene Function, Oxford.

Protein Science is experiencing continued major growth, both in terms of determination of structures (both quantitatively and qualitatively by an expanding repertoire of techniques), characterizing biological properties of proteins and the ability to model the dynamics of structures. Many other areas of the biosciences are undergoing a similar growth: It is interesting to consider how Bioinformatics can contribute to analysis of the expanding knowledge about proteins. Coming years will also see large scale modelling of biological systems and how will the increased knowledge of proteins fit into this process?

Andrew Dalby, Charlotte Deane and Jotun Hein have organised the Bioinformatics Day with focus on Protein Science. It is our intention to have such a half-day before every term. The day will be dedicated to interesting topics and will also be used to announce Genomics/Bioinformatics in the following term: seminars, study group and other meetings. If you intend to attend this event, please send an email to Cathy Went (went@stats.ox.ac.uk) by Tuesday April 18th 2006 with your name and affiliation. This is necessary, so we can arrange coffee and biscuits appropriately and since there is a limit of a 100 in the seminar room in OCGF. Last time 140 wanted to attend. Information on bioinformatics days can be found at <http://www.stats.ox.ac.uk/mathgen/bioinformatics/> under activities.

1.00pm Jotun Hein: Short Welcome

1.10pm Robert Esnouf: "Turning high-throughput into high-success in crystallography"

Developments in technology and automation give structural biology labs the potential for genuine high-throughput work. However, more capacity does not automatically lead to more results, especially for more challenging projects. At all stages of structure determination from target selection and construct design through to annotation of the final structure there is a large and growing need for the techniques of bioinformatics to guide both the overall choice of strategy and the details of individual experimental protocols. These needs will be discussed based on experiences at the Oxford Protein Production Facility covering work on bacterial enzymes through to components of human signaling systems and viral proteins where techniques such as disorder prediction, domain detection and homology modeling are all essential components of successful experimental structure determination by x-ray crystallography.

1.30pm Iain Campbell: 'Structural Biology, NMR and focal adhesions'.

I will give a very brief overview of current experimental structural biology and the role of NMR. Using the focal adhesion complex as an example (this is made up of at least 50 proteins at the cell surface), I will show that structural information is now available for most of the modular protein components. A major remaining problem is how inter-module linear motifs, used to mediate protein-protein interactions, recognise different proteins and how their affinity is modulated by phosphorylation. The utility of some databases such as SMART and ELM will be mentioned.

1.50pm Samantha Kaye: "Determining ligand efficacy in NMDA receptors: all in the hinge?"

NMDA receptors are ionotropic glutamate receptors that are thought to play a key role in learning and memory through the process of synaptic plasticity. They are heterotetrameric proteins which require both glutamate and glycine ligands for activation. Recent crystal structures of the ligand binding domain (LBD) of the glycine binding subunit, NR1, highlighted a subtle conformational change in the hinge region of the LBD, dependant on the efficacy of the bound ligand. We have used molecular dynamics simulations to look more closely at this region of the LBD in the presence of a variety of full and partial agonists, and antagonists. Our findings suggest a mechanism by which partial agonism might be conferred in this group of proteins.

2.10-2.30pm mini-BREAK

2.30pm David Jones: New tools for protein structure prediction

Protein fold recognition methods, particularly those based on the recognition of distant homology are becoming increasingly effective at predicting protein tertiary structure from sequence, as seen in the various CASP prediction experiments (e.g. Jones, 1999a). Despite this success, the better fold recognition methods often employ some degree of human expert intervention, which is clearly impractical if these methods are going to be applied to the annotation of uncharacterised genome sequences. In this talk I will be discussing several automatic methods which have been developed in my lab for predicting protein structure. These methods range from an extended version of a previously published fold recognition method for genome sequences, to a method which is capable of predicting entirely novel protein folds. In addition, I will briefly discuss some methods which have been developed to identify proteins which have novel folds and to identify possible domain boundaries from amino acid sequence alone.

2.50pm Andrew Dalby: Title: Networks within proteins

The discovery of small world networks and their application to biological systems has resulted in a flood of papers on the subject. Regulatory networks, protein interaction networks and metabolic networks have all been shown to have small world properties. This is hypothesised to result from the need for biological systems to be robust and it is supposed that small worlds will characterise living systems. Protein folding is another process that seems to be robust. Proteins fold to the same conformation whether in vivo or in vitro and protein folds can have very wide sequence diversity. So are there small world networks in protein structure?

3.10 pm Richard Copley: Alternative Splicing and Protein Repertoire

Alternative splicing of mRNA transcripts is often touted as a means by which metazoa have increased their proteomic complexity. I will discuss the significance of alternative splicing in animal evolution, focussing on its role as a modifier of protein-protein interactions.

3.30-4.00pm maxi-BREAK

4.00pm Charlotte Deane: Protein structure phylogenies and protein structure prediction

The gap between the number of known protein sequences and structures continues to widen, particularly as a result of sequencing projects for entire genomes. Recently there have been many attempts to generate structural assignments to all genes on sets of completed genomes using fold recognition methods. Our study shows how the outputs of such experiment can be fundamentally improved. We show that false positives generated by fold recognition methods can be identified by considering structural occurrence patterns on completed genomes; occurrences which are isolated within the phylogeny tend to be less reliable.

3.20 pm Andreas Heger: Decomposing protein sequences into domains with ADDA

ADDA is an algorithm to split protein sequences into domains. It applies Okham's razor to split the graph built from all-on-all alignments of protein sequences into components resembling protein domain families. In my talk I will present the ideas behind ADDA and discuss results.

4.40pm Discussion