MS01 O1

Comparative and Structural Genomics to Explore the Evolution of Protein Function. Corin Yeats, Gabrielle Reeves, Oliver Redfern, Juan Ranea and Christine Orengo. Department of Biochemistry and Molecular Biology, University College London, UK.
E-mail: orengo@biochem.ucl.ac.uk

Key words: domain families, genome analysis

How can structural genomics initiatives target relatives from protein families in a manner that increases our understanding of the evolution of protein structures and functions within families? There are now nearly 100,000 domain structures in the CATH database which can be classified into approximately 2000 evolutionary superfamilies. Using HMM based methods to predict structural relatives in completed genomes we observe that more than half of the domain sequences can be assigned to known structural families in CATH [1]. This structural mapping allows us to probe more deeply into the evolutionary history of these families and their differential expansion in the genomes. Although, there are about 140 structural families that are common to all kingdoms of life, a small proportion of these (<20) are highly recurrent accounting for nearly 50% of domain structure annotations in the genomes. Furthermore, many of these very large families are observed to be highly structurally and functionally divergent, though functional divergence is generally limited to changes within a COG major functional class rather than a complete change of functional class. Structural analyses of the most divergent enzyme families reveals a mechanism whereby small accretions of secondary structural elements along the polypeptide change during evolution, are amplified in their impact on the structure through co-localisation in 3D. These secondary structure embellishments often modify the geometry of the active site or the structural characteristics on the surface of the protein promoting different protein-protein interactions [2]. Whilst local structure comparison methods and 3D-templates based on functional sites have difficulty in distinguishing functional subgroups within a structural superfamily, template methods based on global structural comparison show increased specificity and selectivity and reflect the ability of these approaches to capture a broader range of surface characteristics. Sequence based methods for predicting functional subgroups within superfamilies identifies functionally distint subfamilies with no close structural relatives available for homology modelling. These can be targetted by the structural genomics initiatives to improve our understanding of structure-function space.

[1]Comprehensive genome analysis of 203 genomes provides structural genomics with new insights into protein family space. Marsden RL, Lee D, Maibaum M, Yeats C, Orengo CA. (2006) Nucleic Acids Res 34, 1066-1080.
[2]Structural Diversity of Domain Superfamilies in the CATH Database. G.A. Reeves, T.J. Dallman, O.C. Redfern, A. Akpor & C.A. Orengo. (2006) Journal of Molecular Biology 360, 725-41.

MS01 O2

What is the Value of Automatic Protein Structure Prediction? Johannes Söding, *Max-Planck-Institute for Developmental Biology, Tübingen, Germany.* Present address: *Gene Center, University of Munich, Germany.* E-mail: soeding@lmb.uni-muenchen.de

Keywords: homology modeling, fold recognition, remote homology detection, structure prediction

This presentation will give a general introduction of automatic structure prediction methods and show up the possibilities and limitations of these powerful new methods. After explaining the basic procedure for manual structure prediction, I will give an overview of state-of-the-art automatic prediction methods that are easy to use via web servers. The bi-annual, blind structure prediction benchmark CASP (Critical Assessment of Structure Prediction) will be introduced and results for automatic methods in the 2006 competition will be summarized. An overview of some of the most successfull servers is given, including our own HHpred sever for protein structure and function prediction (http://hhpred.tuebingen.mpg.de). Several examples of applications will demonstrate their potential for structural biology.

MS01 O3

Structural Assignment of Spectra by Characterization of Conformational Substates in MbCO. M. Devereux, M. Meuwly, *Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056, Basel, Switzerland.* E-mail: Michael.Devereux@unibas.ch

Keywords: Molecular Dynamics; Ligand Binding; Protein Modeling

Residue motions of the distal heme pocket of the oxygen-storing protein Myoglobin have been shown to influence protein function, control ligand rebinding rates [1], and have been implicated in ligand recognition [2]. In Myoglobin systems binding NO (MbNO), experiment indicates that rebinding from different conformational substates follows distinct kinetics [1], which is likely to also hold true for carbonmonoxy Myoglobin (MbCO). In contrast to the former, for MbCO both the ligand bound (MbCO, A-state) and unbound (Mb···CO, B-state) have been characterized by x-ray crystallography. Because ligand binding and unbinding are transient processes in nature, it is difficult to experimentally characterize both structural and dynamic properties of the system. Atomistic simulations using validated force fields provide additional insight [3,4]. In further studies [2,5], Molecular Dynamics simulations have linked theoretical motions of residues within the heme pocket to changes in observed spectroscopic A-states. The distinct A-states of the bound CO ligand were attributed to different orientations and protonation states of the adjacent HIS64 moiety. Here we characterize the bound states A_0, A_1 and A_3 using a combination of molecular dynamics simulations and Density Functional Theory calculations. Calculated absorption bands were derived from different configurations for comparison with experimental results. The experimental data is then used to guide refinement of the original CHARMM model. Particular attention is paid to the charge model of the bound CO ligand, key to representing interaction with the local binding site.

[1] S. Kim, M. Lim, *J. Am. Chem. Soc.* 2005, *127*, 8909.
[2] A. Loccisano, O. Acevedo, et al., *J. Mol. Graph. Model.,* 2004, *22*, 369.
[3] D. Nutt, M. Meuwly, *PNAS ,* 2004, *101*, 5998.