understanding the principles of protein interactions. Experimental studies like alanine scanning mutagenesis require significant effort; therefore, there is a need for computational methods to predict hot spots in protein interfaces. We present a new efficient method to determine computational hot spots based on sequence conservation and solvent accessibility of the interface residues (Tuncbag et al.; Guney et al.). The predicted hot spots are observed to correlate with the experimental hot spots with an accuracy of 71% and a positive predictive value of 79%. Several machine learning methods (SVM, Decision Trees and Decision Lists) are also applied to predict hot spots and compared to our method. The results reveal that our empirical approach performs better. We observe that both the change in accessible surface area upon complexation and residue accessibility in the complex forms improve detection of hot spots. Predicted computational hot spots for all protein interfaces (49512 interfaces as of 2006) are available at HotSprint database. HotSprint (a database of computational hot spots in protein interfaces) can be accessed at http://prism.ccbb.ku.edu.tr/hotsprint.

Guney E, Tuncbag N, Keskin O, Gursoy A: HotSprint: database of computational hot spots in protein interfaces Nucleic Acids Res 2008, 36(Database issue):D662-666.

Tuncbag N, Gursoy A, Guney E, Nussinov R, Keskin O: Architectures and Functional Coverage of Protein-Protein Interfaces J Mol Biol 2008.

3360-Pos Board B407

Ancestral Sequence Reconstruction and Homology Modeling Link Temperature Adaptation and Conservation of Function With Sequence Evolution in Parvalbumin

A. Carl Whittington[1], Timothy S. Moerland[2].

[1]Florida State University, Tallahassee, FL, USA, [2]Kent State University, Kent, OH, USA.

Temperature is a key factor influencing protein structure and function in poikilotherms. Previous studies of enzymes have shown that orthologs from species acclimated to different thermal niches can maintain a relatively similar level of function at those species' respective physiological temperatures. In some well-characterized enzymes, this conservation of function is correlated with differences in primary structure that lie outside active sites. Information gained from thermal adaptation studies of enzymes can be extended to non-catalytic proteins, which are less thoroughly examined, through the study of parvalbumin structure and function. Parvalbumins are intracellular calcium-binding proteins of the EF-hand type that are thought to function in muscle cells as calcium sinks permitting more rapid unloading of troponin-C, leading to more rapid contraction/relaxation cycles. Parvalbumins contain two functional, highly conserved binding sites and one non-functional site, the AB domain, thought to be an important area of modulation for cation binding. Parvalbumins from teleosts of the sub-order Notothenioid and the unrelated Arctic cod, Boreogadus saida, have converged on a common phenotype. That is, they show similar thermal sensitivity patterns of calcium binding. To explore the underlying structural basis of this similarity in phenotype, we have used ancestral sequence reconstruction combined with homology modeling to identify potential changes in primary structure that have allowed parvalbumins from these disparately related groups of fish to function similarly in their polar habitats. For instance, an Asn to Cys change at position 26 (located in the AB domain) in the evolution of B. saida parvalbumin may lead to the loss of a hydrogen bond in the B-helix. This may provide the change in tertiary structure needed for this parvalbumin to function at polar temperatures.

3361-Pos Board B408

Evolutionary Analyses Of KCNQ1 And HERG Voltage-gated Potassium Channel Sequences Reveal Location-specific Susceptibility And Augmented Chemical Severities Of Arrhythmogenic Mutations

Heather Jackson, **Eric Accili**.

UBC, Vancouver, BC, Canada.

Mutations in HERG and KCNQ1 potassium channels have been associated with Long QT syndrome and atrial fibrillation, and more recently with sudden infant death syndrome and sudden unexplained death. In other proteins, disease-associated amino acid mutations have been analyzed according to the chemical severity of the changes and the locations of the altered amino acids according to their conservation over metazoan evolution. Here, we present the first such analysis of arrhythmia-associated mutations (AAMs) in the HERG and KCNQ1 potassium channels. Using evolutionary analyses, AAMs in HERG and KCNQ1 were preferentially found at evolutionarily conserved sites and unevenly distributed among functionally conserved domains. Non-synonymous single nucleotide polymorphisms (nsSNPs) are under-represented at evolutionarily conserved sites in HERG, but distribute randomly in KCNQ1. AAMs are chemically more severe, according to Grantham's Scale, than changes observed in evolution and their severity correlates with the expected chemical severity of the involved codon. Expected chemical severity of a given amino acid also correlates with its relative contribution to arrhythmias. At evolutionarily variable sites, the chemical severity of the changes is also correlated with the expected chemical severity of the involved codon. Unlike nsSNPs, AAMs preferentially locate to evolutionarily conserved, and functionally important, sites and regions within HERG and KCNQ1, and are chemically more severe than changes which occur in evolution. Expected chemical severity may contribute to the overrepresentation of certain residues in AAMs, as well as to evolutionary change.

3362-Pos Board B409

Evolution of Mammalian HCN Channels - Positive Darwinian Selection Identified at the Molecular Level

Andrew Roth, Heather Jackson, Eric Accili.

UBC, Vancouver, BC, Canada.

HCN channels are important for regulating spontaneous electrical activity and membrane potential in excitable cells. We hypothesize that the four mammalian HCN genes were established by duplication after the divergence of urochordates and before the divergence of fish from the tetrapod lineage. A question in which we are interested is how did the differences in structure and function among the four mammalian channels arise? These differences are due to changes in primary sequence that have occurred since duplication. At the molecular level, changes in DNA sequence over the course of evolution may be identified based upon the Neutral Theory of Molecular Evolution, which states that the majority of changes in DNA are neutral. Neutrality can be estimated directly by comparing the number of non-synonymous changes (dN) and synonymous changes (dS) at the DNA level over a given period of evolutionary time. Neutrality implies that dN and dS are equal (dN/dS=1). A ratio less than one implies purifying selection, whereas a ratio of more than one implies positive selection. Here, we use phylogenetic and statistical analyses of mammalian HCN sequences to identify positive selection among the four mammalian HCN isoforms. We find that the HCN2 isoform yields a very high value for dN/dS ($>>1$), with strong statistical support. Further analysis of HCN2 uncovered a number of specific sites that have undergone positive selection, including an unusual triplet of residues in the outer pore. Our results suggest that positive selection has contributed to differences in structure and function between HCN2 and the other three isoforms.

3363-Pos Board B410

Genomic Identification of Transmembrane β-Barrels (TMBBs)

Thomas Freeman, William C. Wimley.

Tulane University, New Orleans, LA, USA.

Transmembrane beta-barrels (TMBBs) are a special structural class of proteins predominately found in the outer membranes of Gram-negative bacteria, mitochondria, and chloroplasts. It is estimated that 2-3% of a bacterial genome encodes TMBBs, yet less than 40 non-redundant structures have been solved. It would be highly advantageous to have methods to rapidly identify TMBBs from increasingly available genomic databases. A prediction algorithm proposed by Wimley in 2002 was based on the physicochemical properties of TMBBs of known structure. This method used relative amino acid abundances to predict the position of beta-strands and beta-hairpins, which are the major structural subunits of TMBBs, and a mathematical simplification of the topology prediction data called a beta-barrel score. To test the accuracy of this algorithm we scored proteins from a non-redundant database of protein sequences from the Protein Data Bank (NRPDB). The results revealed that the algorithm's ability to discriminate true TMBBs from other proteins, while strong, could be significantly improved. First, we updated the relative amino acid abundances to include the latest structural information. Second, we altered the beta-strand prediction method to account for the fact that certain amino acids have a higher propensity to situate near the lipid/water interface than in the hydrophobic core of the bilayer. Third, we adjusted the calculation of the beta-barrel score to address the lowered beta-hairpin density of larger TMBBs such as BtuB. We reanalyzed the NRPDB and the modifications resulted in a 5-fold decrease in the number of false positives, many of which are either non-bacterial proteins or from Gram-positive organisms. We will use this method to analyze the available genomes of Gram-negative bacteria and the results, along with the signal peptide predictions of SignalP (Bendtsen, et al. 2004) will be deposited into a publicly available database.

3364-Pos Board B411

Predicting Binding Sites of EH1-like Motifs from Their Amino Acid Sequences

D.S. Dalafave.

The College of New Jersey, Ewing, NJ, USA.

Bioinformatics and computer modeling were used to predict binding sites of engrailed homology-1 (eh1) -like motifs from their amino acid sequences. According to previous studies, an eh1 motif provides its transcriptional function

by binding to the pore of WD domain, a highly conserved region of Gro/TLE protein family. Reliable methods for predicting binding sites would allow a better understanding of protein selective recognition mechanisms. This, in turn, would help developments of new organism-specific medications. To the best of the author's knowledge, this is the first study that uses amino acid sequences to predict binding sites of eh1-like motifs to the WD domain. Three-dimensional models of known eh1-like motifs were generated. Their interactions with the WD domain were studied and optimized using Deep View program and the Swiss-Model server. Spatial distributions of binding sites, residue properties, and bonds' stabilities were used to devise a scoring function. The scoring function was employed to predict and evaluate putative binding sites of randomly generated eh1-like sequences. Bioinformatics database searches were used to check whether the scoring function consistently discriminated between viable and non-viable eh1 candidates and corresponding binding sites. The scoring function expressed well a general relationship between putative motifs' residue sequences and their binding sites with the WD domain. Although the function gave a few false positive findings, it reliably identified sequences that did not form stable bonds with the WD domain. The results of this study should lead to a better understanding of mechanisms of transcriptional regulation and selective protein recognition. The method presented may be used to predict binding sites of other regulatory motifs.

3365-Pos Board B412

Genome-wide modeling and analysis of BAR Domains in *Arabidopsis thaliana.*

Deoranie Nikita Brasse, Ramy Abdel-Naby, **Shaneen M. Singh**.
Brooklyn College, Brooklyn, NY, USA.

The BAR (Bin-Amphiphysin-Rvs) Domain is a conserved dimerization protein domain which senses membrane curvature and binds to the lipid plasma membrane. The Amphiphysin protein family, which contain the BAR domain, are thought to be the modulators of the early phases of endocytosis and intracellular transport. It is theorized that this crescent shaped domain dimer show a preference for binding to highly curved, negatively charged membranes. It has been noted that in mammals and higher organisms, the BAR domain with an N-terminal amphipathic helix and BAR domain (N-BAR) can drive the membrane curvature both *in vivo* and *in vitro*. We have studied the BAR domain in the small flowering plant which is widely used as the model for plant biology, *Arabidopsis thaliana*, where its function remains largely unexplored. We have modeled all the BAR domains of *Arabidopsis thaliana* using an automated modeling pipeline with manual refinement methods and investigated their mechanism relative to other higher organisms. We have identified eight non-redundant domain sequences in *Arabidopsis thaliana*, which can be grouped into three different classes, based on their electrostatic profiles and domain architecture. We provide new insight into the features of plant BAR domains including distinct electrostatic profiles for domain sequences categorized within the same class and atypical electrostatic profiles showing a concentration of positively charged residues at both extremities of the structural fold. Our results are important in understanding the differences in signaling through BAR domains in plants and its implication in plant signaling and membrane trafficking.

3366-Pos Board B413

Interactive Visualization of Protein Dynamics in Ribbon Mode

Manuel Wahle, Stefan Birmanns.
University of Texas Health Science Center at Houston, Houston, TX, USA.

In recent years, experimental methods have uncovered more and more dynamic properties of molecular systems. Novel techniques for post-processing data from cryo-electron microscopy or small angle X-ray scattering nowadays routinely reveal conformational differences linked to functional states of a biological system. As opposed to molecular dynamics, these methods typically yield information about larger systems, which in effect makes an interactive visualization of the results challenging.

This poster presents an approach that computes a depiction in the more abstract ribbon or cartoon mode, which highlights the secondary structure information of a protein. To accomplish interactive frame rates, the algorithm offloads computational work from a PC's CPU to its graphical processing unit (GPU). This is achieved by separating two phases in the calculation of the geometry representing the protein. The first one involves creating a smooth curve along the protein's backbone, which requires global information and thus has to be computed by the CPU. In the second phase, vertices along that curve are moved to build the geometry for the molecule's depiction. Because local information is sufficient for that, this is handled exclusively by the GPU.

The speed-up factor achieved moves a range of large time-varying proteins into the category of those which can be depicted with interactive frame rates. For intermediate sized molecules, the speed-up results in an even smoother animation and an overall increase of the reactivity of the whole program. Especially simultaneously running processes, e.g. calculations for multi-scale modeling, can benefit from the additionally available CPU resources, too.

This work was supported by startup funds of the University of Texas Health Science Center at Houston, by NIH grant R01GM62968, and by a grant from the Gillson Longenbaugh Foundation.

3367-Pos Board B414

Dynamics-based Alignment: A Novel Tool For Comparing Large-scale Movements In Proteins With Same Or Different Fold

Andrea Zen[1], Vincenzo Carnevale[2], Arthur M. Lesk[3], Cristian Micheletti[1].
[1]International School for Advanced Studies, Trieste, Italy, [2]Department of Chemistry, University of Pennsylvania, Philadelphia, PA, USA, [3]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA.

The biological function of several proteins and enzymes is assisted by large scale conformational changes that are excited in thermal equilibrium. In terms of the traditional logical cascade, sequence -> structure -> function, it is expected that the functional movements of a protein are influenced by the structural architecture. Proteins with similar structures are known to sustain similar large-scale movements; yet it has recently emerged [Carnevale et al. JACS 2006, Capozzi et al. J_Proteome_Res 2007, Zen et al. Prot_Sci 2008] that similar functional movements are shared by proteins with different architecture or topology.

This observation parallels the known paradigm that (i) proteins with similar primary sequences usually attain a similar fold but also that (ii) the same fold is adopted by non-homologous proteins. The sophisticated interplay between sequence and structure has now been extensively characterized thanks to the availability of sequence and structural alignment methods. By analogy, the availability of quantitative methods for comparing the functional-oriented dynamics in proteins would allow to take to a new level the investigation of the structure/function relationship.

We report on a first attempt in this direction by discussing a pairwise alignment scheme that identifies groups of amino acids that undergo similar concerted movements in proteins. The alignment method is based on a coarse-grained elastic network model and requires as input the sole proteins' native structures. No prior detection of structure and sequence correspondence is used. The scheme is first used to perform a dynamics-based alignment (and grouping) of a data set of >70 representative enzymes covering the main functional and structural classes. Finally we discuss an application where the method is used to identify the putative nucleic-acid-binding regions of proteins having AXH-domains [de Chiara et al. Structure 2005].

3368-Pos Board B415

Template-Based Modeling of Protein-Protein Interfaces

Petras Kundrotas, Ilya A. Vakser.
University of Kansas, LAWRENCE, KS, USA.

Modeling of protein binding sites is important for 3D prediction of protein complexes. Statistical analysis of target-template PSI-BLAST sequence alignments was performed for 329 two-chain target protein complexes selected from DOCKGROUND database. For 214 complexes (~65 %) the alignments contained all interface residues (full interface coverage or FIC alignments) for both complex monomers and 101 (~30 %) complexes had FIC alignments for one of the monomers. The FIC alignments were observed even in the case of poor alignments where only a small portion of the target sequence (as low as 40%) was aligned to template sequence with low alignment identity (40%) alignments, whereas for the low-identity alignments the situation is opposite. Homology models were built based on the FIC alignments with target sequence coverage < 60 %. The results showed that one third of the target sequences with such short FIC alignments produced models with interface RMSD (i-RMSD) < 5 Å, suitable for low-resolution ab initio docking. The proteins with i-RMSD < 5 Å had domain structure, whereas models with 5 Å $<$ i-RMSD < 8 Å (accuracy suitable for structure-alignment methods) were generated for single-domain proteins as well. The results provide guidelines for building 3D protein models for docking studies.

3369-Pos Board B416

Homology Modeling and Molecular Dynamics Simulation Studies of the human resistin protein

Chilamakuri C. Sekhar Reddy[1], Ramanathan Sowdhamini[2], Bernard Offmann[1].
[1]University of Reunion, Saint-Denis, Reunion, [2]National Centre for Biological Sciences, Bangalore, India.

Resistin is a member of a secretory protein family, known as resistin-like molecules (RELMs) which were exclusively found in mammalian genomes. Though human resistin molecule has high sequence similarity with mouse, its structural and physiological roles differ considerably from mouse and the