3354-Pos Board B401
Development of Cardiac Integrated Database Management System (CIDMS)
Pradeep Kumar Sreenivasaiah, Bharathi Mohanraju, Do Han Kim.
GIST, Gwangju, Republic of Korea.
A major emphasis in systems biology is to unfold the dynamic network properties of complex biological processes through modeling and simulation utilizing various experimental data. CIDMS was developed to provide an integrated bioinformatics platform for systems biology approach studying mammalian cardiac systems. CIDMS has following information which was manually curated from published literatures and public databases: 1) list of genes, proteins, pathways critical for development and functioning of heart; 2) list of human cardiac diseases; and 3) mathematical / computational models employed to study cardiac biological processes and diseases. These lists are supplemented with key annotations like function, regulatory mechanisms, interactions and references to the source of information in context of cardiology. CIDMS also provides a large amount of additional automated data integrated within the framework of manually curated data. Data such as structure, function, domains, location, phenotype, molecular interactions, and pathways from numerous life-science databases are made available. In addition to these qualitative data, CIDMS integrates quantitative data such as reaction kinetics, stoichiometric and model information. This broad data integration has become possible by the state of the art technologies like data-warehouse and through webservices (SOAP interfaces) of well established public database servers ensuring up to the minute information. The comprehensive information can be easily searched and browsed by categories and ontologies with an advanced web interface hosted at http://cidms.org.

3355-Pos Board B402
Towards Unifying Systems Biology - Using Pathway Data in Biopax Format for SBML Simulators
Oliver Ruebenacker, Ion I. Moraru, **Michael L. Blinov**.
University of Connecticut Health Center, Farmington, CT, USA.
Thousands of biochemical interactions are available from public sources in the Biological Pathways Exchange (BioPAX) format. However, the current standard for exchange of simulation-ready biological models is System Biology Markup Language (SBML). This markup language is structurally and semantically different from BioPAX. Some conversion schemes exist, using annotations and based on simple one-to-one mappings between SBML and BioPAX objects, which ignores semantic differences and therefore often leads to significant loss of information or meaning. A comprehensive modeling framework capable of representing the complex relationships between SBML and BioPAX data is needed to take full advantage of existing pathway data in kinetic modeling, thus integrating these two formats by gluing them together.
Here we describe such a framework that we are developing as a part of the Virtual Cell (http://vcell.org) modeling and simulation environment. Systems Biology Linker (SyBiL, http://vcell.org/biopax) is a tool for analyzing and visualizing BioPAX data, and converting them to SBML. Based on the Jena Semantic Web framework for Java, SyBiL supports handling of generic RDF/OWL data (such as visualization and reasoning) as well as functions specific to handling SBML and BioPAX data. SyBiL uses Systems Biology Pathway eXchange, called SBPAX, as a generic approach to integrate model-centric formats similar to SBML with pathway-centric formats similar to BioPAX. SBPAX is an OWL-based schema that serves as a glue to integrate different data formats, despite semantic differences. Effectively, SBPAX provides a bridge between SBML and the Semantic Web world. SyBiL offers various visualization modes showing reaction networks to varying degrees of details, including displaying nodes for reactions only as well as displaying Petri nets consisting of reaction nodes and reaction participants and catalysts.

3356-Pos Board B403
Prediction of Functional Sites in 50,000 Protein Domains Using Dynamics Perturbation Analysis
Judith D. Cohn, Dengming Ming, **Michael E. Wall**.
Los Alamos National Laboratory, Los Alamos, NM, USA.
Dynamics perturbation analysis (DPA) [1-3] finds regions in a protein structure where proteins are ''ticklish,'' *i.e.*, where interactions cause a large change in protein dynamics. Such regions corresponded to native binding sites in a test set of hundreds of protein-ligand complexes [3]. We have further tested DPA for prediction of protein functional sites by applying an accelerated algorithm, Fast DPA [4], to over 50,000 structures of protein domains from the SCOP database [5]. We compared the predicted sites with known functional sites obtained from two sources: catalytic residues in the Catalytic Site Atlas; and putative binding sites identified by finding protein residues near small molecules in crystal structures. Combining predictions with information from sequence conservation and multiple sequence alignments reduced false positive rates. Overall, prediction of functional sites using DPA recapitulates much of the known information about functional sites in SCOP domains, and validates the use of DPA to predict functional sites in proteins. These results further suggest that functional sites in proteins tend to evolve at control points where interactions cause a large change in protein dynamics [1, 3].
1. Ming D. and Wall M.E. Quantifying allosteric effects in proteins. 2005. Proteins 59:697-707.
2. Ming D. and Wall M.E. Allostery in a coarse-grained model of protein dynamics. 2005. Phys Rev Lett 95:198301.
3. Ming D. and Wall M.E. Interactions in native binding sites cause a large change in protein dynamics. 2006. J Mol Biol 358:213-223.
4. Ming D., Cohn J.D. and Wall M.E. Fast dynamics perturbation analysis for prediction of functional sites. 2008. BMC Struct Biol 8:5.
5. Murzin A.G., Brenner S.E., Hubbard T. and Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. 1995. J Mol Biol 247:536-540.

3357-Pos Board B404
From Abstract Graphs To (biophysical) Reality - Graphs In Mutual Information Space Help Identify Highly Correlated Positions In G Protein-coupled Receptors
Sarosh N. Fatakia, Stefano Costanzi, Carson C. Chow, Vipul Periwal.
NIDDK, NIH, Bethesda,, MD, USA.
G protein-coupled receptors (GPCRs) are a superfamily of seven transmembrane-spanning proteins involved in a wide array of physiological functions and are the most common targets of pharmaceuticals. Are there idiomatic relations between positions on these protein sequences that are evident in their co-evolution? We investigate the semantic graph of the constituent amino acid (AA) positions using an information theoretic approach. Using a multiple sequence alignment of the seven transmembrane (7-TM) domains, we calculated the mutual information (MI) between all pairs of aligned positions. Representing TM positions as vertices and pairing them by their MI we compute the planar acyclic graph that maximizes the total MI. The total MI of this graph is much greater than the total MI of a random planar acyclic graph. From this graph, we identify few positions which have a significantly high degree (edges incident to the vertices) when contrasted with the others. The positions, from class A and class C GPCRs, with the leading degree values are found to be associated with the experimentally determined binding pocket, confirming our previous studies involving MI graphs.

3358-Pos Board B405
Local surface shape-based protein function prediction using Zernike descriptors
Daisuke Kihara[1], Lee Sael[1], Rayan Chikhi[2].
[1]Purdue University, West Lafayette, IN, USA, [2]Ecole Normale Superieure de Cachan, Britanny, France.
Structural genomics projects have been solving an increasing number of protein structures with unknown function. As the number of uncharacterized protein structures continues to grow, there is increasing need for effective computational strategies for structure-based protein function prediction. As of now, there are more than 2400 protein structures in the PDB database, which are categorized as unknown function and thus are awaiting functional characterization. Previously, we have reported a fast global protein surface shape comparison method based on 3D Zernike descriptors. 3D Zernike descriptors are series expansions of a given three-dimensional function that compactly represent a protein surface and its corresponding physicochemical properties. Consequently, it takes approximately a minute to compare a query structure against thousands of protein structures in a database.
Here, we apply the method for local protein surface shape comparison, focusing on surface pockets, which are potential ligand molecules binding sites. Applications of the methods presented in this work include classification of ligand binding sites of TIM barrel proteins according to their surface electrostatic potential and classification of protein local surface regions in terms of the similarity of shape and physicochemical properties. We also show that prediction of binding ligand molecule by pocket similarity search. Taken together, we demonstrate the high throughput applicability of Zernike descriptors in recognizing local physicochemical of protein surfaces. This work was supported in part by NIH (R01 GM075004).

3359-Pos Board B406
Large Scale Prediction of Computational Hot Spots in Protein Interfaces
Nurcan Tuncbag[1], **Ozlem Keskin**[1], Ruth Nussinov[2], Attila Gursoy[1].
[1]Koc University, Istanbul, Turkey, [2]SAIC-NCI, Frederick, MD, USA.
Hot spots are residues comprising only a small fraction of interfaces yet accounting for the majority of the binding energy. These residues are critical in