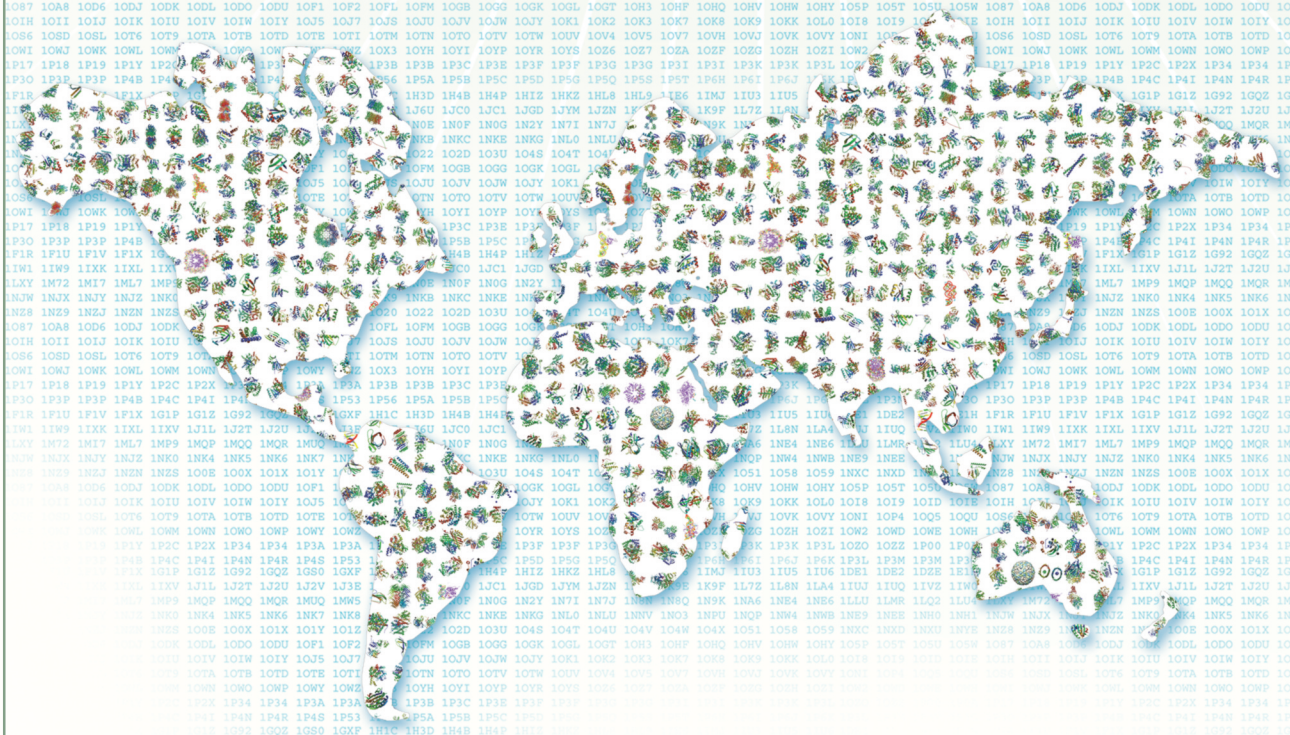
PDB40

Symposium

POSTER ABSTRACTS

October 28-30, 2011

Cold Spring Harbor Laboratory
Grace Auditorium



POSTER PRESENTATIONS

PDB40 posters will follow the numbers assigned in this abstract book. Poster locations will be numbered at the meeting site.

Posters should be displayed throughout the meeting. A dedicated poster session is scheduled for Saturday, October 29, 2011 at 5:30 p.m., but meeting attendees are strongly encouraged to explore the posters throughout the meeting.

Posters should fit within **1.22m x 1.22m** (4 ft X 4 ft). Supplies for hanging posters will be provided.

Poster Assignments

* denotes travel award recipient

Poster #	Author	Poster title
1*	Hiromi Arai	Crystal structure of a conformation-dependent rabbit IgG Fab fragment specific for a sequence-independent generic epitope associated with prefibrillar amyloid oligomers
2*	Ryoichi Arai	Domain-swapped dimeric structure of a *de novo* 4-helix bundle protein, WA20
3*	Pelin Armutlu	A TOG:αβ-tubulin complex structure suggests conformation-based mechanisms for a microtubule polymerase
4	Christopher Bahl	The *Pseudomonas* virulence factor Cif alters human ABC transporter trafficking and stability through epoxide hydrolase enzyme activity
5	Ramanuj Banerjee	Threaded dimer assembly of the novel Psu fold from Enterobacteria phage P4
6*	Avraham Ben-Shimon	Deciphering the arginine-binding preferences of Ser/Thr kinases by computational surface mapping
7	Talapady Bhat	Managing 2D and 3D ligand fragments from PDB and PubChem
8	Shveta Bisht	*Escherichia coli* Diaminopropionate ammonia lyase—Structure, substrate specificity and catalysis
9*	Spencer Bliven	A comprehensive comparison of protein structures and the correlation of structure with function
10	Matthias Buck	Dynamic protein-protein complexes—How alternative interactions create ensembles and how solution NMR and MD simulations can characterize them
11	Brianne Burkinshaw	Salmonella virulence protein SopB down-regulates activity of host Rho GTPase Cdc42 by mimicking a nucleotide dissociation inhibitor
12*	Arindam Chakrabarty	Unique crystal structure of protozoan parasite *Entamoeba histolytica* malic enzyme
13*	Tammy Cheng	Structural biology meets systems biology—Gauging the systemic impact of single nucleotide polymorphisms
14*	Alvaro Cortes	AtlasCBS—A web server to graphically represent the content of chemico-biological databases

Poster #	Author	Poster title
15	Charles Dann III	Treatment of cancer and autoinflammatory disease for over sixty years—Structural basis for folate and antifolate trafficking via human folate receptors
16*	Teresa De la Mora	Crystal structures of eIF4E in complex with Bn^7-GMP
17*	Bryan Der	Metal-mediated affinity and orientation specificity in a computationally designed protein homodimer
18	Jianping Ding	Regulation of the histone acetyltransferase activity of hMOF via autoacetylation of Lys274
19	Wolfgang Dostmann	A structural analysis of the regulatory domain from the cGMP-dependent protein kinase Iα
20	Joseph Dybas	Evolution of protein structures from the perspective of their super-secondary structure building blocks
21*	Clarissa Eibl	Crystal structure of NLRP4 pyrin domain
22	Richard Feldmann	The second log
23	Javier Fernandez-Martinez	Structure-function Map of a heptameric module in the nuclear pore complex
24	Barry Finzel	The use of distance geometry for substructure searching and local superposition in DNA, RNA, and protein-polynucleotide complexes
25	Andras Fiser	Predicting protein structure with supersecondary-structure building blocks and basic NMR data
26*	Szu-Chin Fu	Prediction of leucine-rich nuclear export signal containing proteins with NESsential
27*	Kyoko Furuita	NMR structural studies of the complex between lipid binding protein OSBP and ER membrane protein VAP-A
28	Miwako Homma	Autophosphorylation of CK2α at the N-terminal region is required for its catalytic activity
29*	Corey Hryc	Hybrid de novo modeling for 4 Å resolution cryo-EM maps of viruses
30	Valentin Ilyin	PFAT—Protein Functional Annotation based on TOPOFIT
31	Valentin Ilyin	N-TOPOFIT-DB, a DNA / RNA 3-D structural alignment database by N-TOPOFIT
32	Hitoshi Inada	Structure of human TRPV4 ankyrin repeat domain
33	Constance Jeffery	Moonlighting proteins
34*	Fangling Ji	Solution NMR and biophysical analysis of the cataract-associated R76S mutant of human γD-crystallin
35	Holly Jing	Discovery of novel inhibitors of therapeutic targets using fragment-based lead discovery methods
36	Jan Johansson	Prevention of amyloid β-peptide fibrillation—Insights from the BRICHOS domain structure
37	Chisa Kamada	Supercomputational life science in Japan
38	Tatsuki Kashiwagi	Crystallization of proteins in high magnetic force fields aiming at high resolution structures
39	Takeshi Kawabata	Chemical similarity-based docking to receptor proteins by searching maximum common substructure
40	Choel Kim	Structural and biochemical studies of the cGMP selectivity for cGMP dependent protein kinase
41*	Seung Joong Kim	Integrative structure determination of the components of the nuclear pore complex by X-ray crystallography, small angle X-Ray scattering, electron microscopy, NMR and comparative modeling

Poster #	Author	Poster title
42*	Serah Kimani	Unexpected reactions resulting from mutating catalytic residues in an amidase reveal the role of the catalytic unit
43*	Yu Kitago	Structural basis for peptide ligand recognition by LR11 Vps10p domain
44	Stefan Knight	Structural basis for control of spider silk assembly—A conserved N-terminal solubility relay
45	Adrian Laurenzi	Improving EST annotation with protein structure prediction software
46	Tali Lavy	The *GAL* regulon in *S. cerevisiae*—The Gal3p/Gal80p interaction
47	Thomas Leeper	Silver (I) mellatated protein structures derived by combined NMR and X-ray crystallographic analyses
48	Yang Lei	A maximum-likelihood approach to local real-space model fitting
49*	Alejandra Leo-Macias	In vitro reconstituted F-actin and fascin bundles studied by cryoelectron tomography and subtomogram averaging
50*	Yunfeng Li	Crystal structure and functional studies of the GerBC component of a *Bacillus subtilis* spore germination receptor
51*	Vidya Mangala Prasad	Towards a structure for the Rubella virus capsid protein
52	Manish Manish	Computational design of conformationally intact immunogen for the elicitation of protective antibody immune response
53	Stefano Maria Marino	Extreme evolutionary pressure determines Cysteine residues distribution in protein molecular surfaces
54	Yamuna Kalyani Mathiharan	Domain swapping and protein symmetry—Investigations on the stationary phase survival protein SurE as a model
55	William McLaughlin	A searchable set of predictions of protein functions with associated probabilities as derived from annotations of protein sequences and three-dimensional structures
56	Wenli Meng	Unfolded proteins can form long range contacts under strongly denaturing conditions in the absence of significant secondary structure
57	Thomas Moon	Autoinhibition and salt sensing are linked in the WNK1 kinase
58	Andrew Morin	The SBGrid Consortium—Enabling structurally biology computation
59*	Chie Motono	SAHG, a comprehensive database of predicted structures of all human proteins
60*	Kensuke Nakamura	Structural analysis of BAR domain of Arfaptin in complex with small GTPase Arl1
61	Janet Newman	Getting better at crystallisation by learning from the past
62*	Ilona Nudelman	Observation of intermolecular interactions in large protein complexes by 2D-double difference NOESY—Application to the 44 kDa interferon-receptor complex
63	Åsa Nylander	Structural studies of streptococcal adhesins
64	Sang Ho Park	Computational docking studies of histone deacetylase 8 (HDAC8) and its inhibitors by AutoDock Vina program
65*	Aviv Paz	Sugar transport inhibition of the sodium galactose transporter
66	Janez Plavec	NMR solution-state structures of monomeric and dimeric G-quadruplexes adopted by a sequence from N-myc
67*	Pavel Plevka	Mechanism of re-organization of flavivirus glycoproteins during maturation
68	Mario Pujato	The evolution of robustness involves balancing of local sequence and network architecture level mechanisms
69	Pavlina Rezacova	Crystal structures of the effector-binding domain of repressor CggR from *Bacillus subtilis* reveal ligand-induced structural changes upon binding of several glycolytic intermediates

Poster #	Author	Poster title
70	Florian Richter	A computational method to design backbone conformations featuring predefined interactions
71	Victoria Robinson	Elucidating the mechanism of action of the translation factor BipA
72	Aleksandr Sahakyan	Protein structures from the perspective of side-chain chemical shifts
73	Stefan Schoebel	LidA from Legionella is a Rab-supereffector
74*	Masaaki Shiina	X–ray crystallographic analysis of Runx1–CBFβ–Ets1–DNA complex assembled on the enhancer of T cell receptor α chain gene
75	Irena Sieglova	Structure of the effector-binding domain of arabinose repressor AraR from *Bacillus subtilis*
76*	Peter Stranges	Accurate design of a symmetric homodimer using β-strand assembly
77	Juergen Suehnel	The Jena Library of Biological Macromolecules—JenaLib
78	Juergen Suehnel	The Jena3D viewer for visualization of biological macromolecules
79*	Hiroshi Sugimoto	Structure and catalytic mechanism of iron-dependent enzymes
80*	Mitsuhiro Takeda	Recent progress in a stereo-array isotope labeling (SAIL) method for NMR structure determination
81	Todd Taylor	Discriminating thermophilic proteins from their mesophilic analogs
82	Eldon Ulrich	The Biological Magnetic Resonance Data Bank (BMRB)
83	Ihsanawati Wahab	Structural basis for the high alkaliphily of modular xylanase XynJ from *Bacillus* sp. 41M-1
84	Gary Wesenberg	The role of peptide-peptide stacking interactions in protein structure
85	Carrie Wilmot	In crystallo synthesis of the tryptophan tryptophylquinone cofactor of methylamine dehydrogenase
86*	Mousheng Wu	Structural basis of the Ca^{2+} inhibitory mechanism of *Drosophila* Na^{+}/Ca^{2+} exchanger CALX and its modification by alternative splicing
87*	Guozhou Xu	Structure, substrate specificity and mechanism of kinase activation of the IκB kinase β (IKKβ)
88*	Yusuke Yamada	Automation of data collection and processing at the Photon Factory macromolecular crystallography beamlines
89	Olga Yuzlenko	Membrane protein native state discrimination by a physical energy function
90	Huan Zhan	Influence of membrane dipole potential on peptide binding to lipid bilayers
91	Yan Zhang	Dephosphorylation regulation of RNA polymerase II in eukaryotic transcription
92	Yong Zhang	From NMR information to high accuracy biomolecular structures—A quantum chemical pathway
93*	Samira Zouhir	Structural study of a molecular switch implicated in quorum sensing in *Bacillus cereus*

Crystal structure of a conformation-dependent rabbit IgG Fab fragment specific for a sequence-independent generic epitope associated with prefibrillar amyloid oligomers.

Hiromi Arai[1], Charles Glabe[1], Hartmut Luecke[1,2,3,4]

[1]University of California, Irvine, Department of Molecular Biology and Biochemistry, Irvine, CA, 92697, [2]University of California, Irvine, Department of Physiology & Biophysics, Irvine, CA, 92697, [3]University of California, Irvine, Department of Computer Science, Irvine, CA, 92697, [4] University of California, Irvine, Center for Biomembrane Systems, Irvine, CA, 92697

Although rabbit antibodies are widely used in research, no structures of rabbit antigen-binding (Fab) fragments are known in the Protein Data Bank (PDB). M204 is a rabbit monoclonal antibody that recognizes a generic epitope that is common to prefibrillar amyloid oligomers formed from many different amyloidogenic sequences. Amyloid oligomers are widely suspected to be a primary causative agent of pathogenesis in several age-related neurodegenerative diseases, such as Alzheimer's disease (AD). The detailed structure of these amyloid oligomers is not known nor is the mechanism for the recognition of the generic epitope by conformation-dependent monoclonal antibodies. As a first approach to understanding the mechanism of conformation-dependent antibody recognition, we have determined the structure of M204 Fab (PDB ID 3NL4) at 1.54 Å resolution. The structure reveals structural features unique to rabbit Fab fragments, and it rules out "steric zipper" formation as a simple mechanism of antibody-antigen recognition. The details of a rabbit Fab structure should also be useful for exploiting the potential of rabbit monoclonal antibodies for the development of humanized rabbit antibodies as therapeutic agents.

Domain-swapped dimeric structure of a *de novo* 4-helix bundle protein, WA20

Ryoichi Arai*[1,2], Naoya Kobayashi[3], Akiho Kimura[3], Takaaki Sato[1], Anna F Wang[2], Jesse Platt[2], Luke H Bradley[2,4], Michael H Hecht[2]

[1]Shinshu Univ., Int. Young Researchers Empowerment Ctr., Ueda, 386-8567, Japan, [2]Princeton Univ., Dept. of Chem., Princeton, NJ, 08544, [3]Shinshu Univ., Dept. of Appl. Biol., Ueda, 386-8567, Japan, [4] Univ. of Kentucky College of Med., Depts. of Anat. & Neurobiol., Mol. & Cell. Biochem., Lexington, KY, 40536

To probe the potential for enzymatic activity in unevolved amino acid sequence space, we created a 3rd generation combinatorial library of *de novo* 4-helix bundle proteins. This collection of novel proteins can be considered an ‘‘artificial superfamily’’ of helical bundles. The superfamily of 102-residue proteins was designed using binary patterning of polar and nonpolar residues, and expressed in *Escherichia coli* from the library of synthetic genes. WA20, picked up from the library, is one of the most stable proteins in the superfamily of *de novo* proteins, and has rudimentary activities of peroxidase with heme, esterase, and lipase [Patel, S.C., et al., Protein Sci. 18, 1388 (2009)]. Here we report the crystal structure of WA20, determined by the MAD method. Unexpectedly, the WA20 crystal structure is not a monomeric 4-helix bundle but a dimeric 4-helix bundle. Each monomer comprises two long α-helices that intertwine with the helices of the other monomer. Together the two monomers form a domain-swapped 4-helix bundle dimer. In addition, there are two hydrophobic pockets, which may potentially work as substrate binding pockets. Small-angle X-ray scattering shows that the molecular weight of WA20 is about 25 kDa, the shape is a rod-like (the maximum diameter, D_{max} = ~8 nm), indicating that WA20 forms a dimeric 4-helix bundle in solution. These results suggest that our *de novo* protein library contains not only simple monomeric proteins but also functional multimeric proteins. It may open the way for applying *de novo* proteins to the self-assembly of functional building blocks for nanobiotechnology.
*R.A. is supported by PDBj travel grant from JSPS.

A TOG:αβ-tubulin complex structure suggests conformation-based mechanisms for a microtubule polymerase

Pelin Armutlu, Patrick Huddleton, Luke Rice

UT Southwestern Medical Center at Dallas, Biochemistry, Dallas, TX, 75390

Microtubules are highly regulated, dynamic polymers of αβ-tubulin that have essential roles in intracellular organization and chromosome segregation. Proteins in the evolutionarily conserved Stu2p/XMAP215/Dis1 family are the major regulatory factors that promote fast microtubule growth in vivo. These proteins use multiple tubulin-interacting TOG domains to bind unpolymerized αβ-tubulin and to selectively recognize one end of the microtubule, where they catalyze microtubule elongation. TOG domains are also present in a second family of microtubule-regulating proteins, the CLASPs. The molecular details of TOG domain function remain poorly understood, in part because the structural basis of TOG:tubulin interations has not yet been defined. Here we report the 2.9 Å resolution structure of a complex between the TOG1 domain from Stu2p and a polymerization-blocked mutant of yeast αβ-tubulin that we developed to facilitate crystallization. The structure reveals, contrary to expectation, that the TOG1 domain makes significant contacts with both α- and β-tubulin, and that the conformation of GTP-bound αβ-tubulin resembles a 'curved', microtubule-incompatible conformation observed previously in the structure of αβ-tubulin bound to a stathmin-like protein. The TOG1-interacting epitopes on α- and β-tubulin revealed by the structure do not overlap αβ-tubulin polymerization interfaces, and biochemical experiments demonstrate that TOG1 discriminates between αβ-tubulin conformations, binding preferentially to a curved one. This first structure of GTP-bound αβ-tubulin failed to reveal a postulated, partially straightened conformation of αβ-tubulin. Our findings have implications for the mechanisms by which Stu2p family proteins recognize microtubule ends and catalyze their elongation.

The *Pseudomonas* virulence factor Cif alters human ABC transporter trafficking and stability through epoxide hydrolase enzyme activity.

Christopher D Bahl, Dean R Madden

Dartmouth Medical School, Biochemistry, Hanover, NH, 03748

Pseudomonas aeruginosa is a Gram-negative opportunistic pathogen that causes chronic infection in the lungs of patients with compromised pulmonary function, and is especially prevalent for patients with cystic fibrosis (CF). *P. aeruginosa* has been shown to secrete a soluble protein, termed Cif (CFTR inhibitory factor), which inhibits the endocytic recycling of a select population of ABC transporters, including the cystic fibrosis transmembrane conductance regulator (CFTR). The result of Cif treatment on human airway epithelial cells in culture is a loss of CFTR mediated chloride efflux across the apical membranes. Cif promotes the mislocalization of CFTR to endosomes by blocking its recycling to the plasma membrane. Thus, CFTR is removed from the plasma membrane and shunted to the lysosomes for degradation. Since the majority of CF patients are colonized by *P. aeruginosa*, Cif represents a potential roadblock for the effectiveness of CFTR targeted therapies in a clinical setting. Additionally, elucidation of the Cif mechanism will illuminate novel pathways regulating ABC transporter trafficking.

We have determined the structure of Cif by X-ray crystallography and demonstrated it to be an epoxide hydrolase (EH) with unique substrate selectivity. Mammalian EHs can function as regulators of blood pressure and inflammation, while bacterial EHs have mostly been characterized for their ability to detoxify xenobiotic compounds. Cif is the first example of an EH serving as a virulence factor. Cif's EH enzyme activity utilizes a catalytic triad of residues Asp129, Glu153 and His297, with accessory residues His177 and Tyr239 coordinating the epoxide oxygen during ring opening. In the proposed enzyme mechanism, Asp129 undergoes nucleophilic attack at a carbon of the epoxide moiety of a substrate, forming an ester linked enzyme-acyl intermediate. In the second step of the reaction, a water molecule is activated by the charge-relay His297-Glu153 pair, and undergoes nucleophilic attack at the Cγ of Asp129. This hydrolyzes the ester group, liberating the hydrolysis product as a vicinal diol.

Mutation of any active site residue inhibits enzymatic hydrolysis of epoxide compounds. Furthermore, these mutations abolish the CFTR inhibitory effect seen on airway epithelial cells. Thus, it appears that EH activity is strictly required for Cif mediated mislocalization and degradation of CFTR, suggesting that Cif acts on a physiological epoxide with a role in regulating ABC transporter trafficking. It is possible that either the depletion of an epoxide target or the generation of a vicinal diol hydrolysis product is mediating this effect, and current work is aimed at elucidation of these putative compounds.

Threaded dimer assembly of the novel Psu fold from Enterobacteria phage P4

Ramanuj Banerjee[1], Seema Nath[1], Susmita Khamrui[1], Ranjan Sen[2], Udayaditya Sen[1]

[1]Saha Institute of Nuclear Physics, Crystallography and Molecular Biology Division, Kolkata, 700064, India, [2]Center for DNA Fingerprinting and Diagnostics, Laboratory of Transcription Biology, Hyderabad, 500001, India

Polarity Suppression protein (Psu) is a non essential capsid decoration protein from bacteriophage P4 and is involved in Rho dependent transcription termination in Escherichia coli. This 21 kDa protein exists as a dimer and this oligomeric state is also present on the capsid of the P4 virion. Psu inhibits the ATPase activity of the Rho hexamer, but does not effect its primary or secondary RNA binding activities. The interaction of Psu is direct, mediated by the C-terminal domain of Psu. Although biophysical characterization has been done, the structure and assembly of Psu has remained unknown until now. Here we focus on the 2 A crystal structure of the novel Psu dimer and shed light on the unique threaded dimer assembly which it attains in the crystal structure. Comparison with the previously determined electon micrograph of the P4+Psu capsid at 45 A resolution reveals the existence of the same dimeric structure on the capsid bridging the hole to prevent DNA leakage from the interior. Only a handful of proteins in nature form intrachain knots, but the formation of intermolecular threads are absent. It will be interesting to study the refolding kinetics of the dimer to reveal the appropriate mechanism as to how it crosses the high energy barrier to attain its energy minima in the threaded conformation. The site of interaction of Psu on the Rho hexamer is probably in close proximity to the secondary RNA binding channel of Rho. We show here the probable mode of interaction of Psu with rho through information based in-silico macromolecular docking studies.

Deciphering the Arginine-Binding Preferences of Ser/Thr Kinases by Computational Surface Mapping

Avraham Ben-Shimon, Masha Y.Niv

Institute of Biochemistry, Food Science and Nutrition, Faculty of Agriculture, Hebrew University of Jerusalem, Rehovot, 76100, Israel

Protein kinases are key signaling enzymes that catalyze the transfer of γ-phosphate from an ATP molecule to a phospho-accepting residue in the substrate. Unraveling the molecular features that govern the preference of kinases for particular residues flanking the phosphoacceptor is important for better understanding kinase specificities toward their substrates and for designing substrate-like peptidic inhibitors.
We applied ANCHORSmap, a new fragment-based computational approach for mapping amino acid side chains on protein surfaces, to predict and characterize the preference of kinases toward Arginine (Arg) binding, focusing on the intricate specificities at positions P-2 and P-5, commonly occupied by Arg in substrates of basophilic Ser/Thr kinases. The method accurately identified all the P-2/P-5 Arg binding sites previously determined by X-ray crystallography and produced Arg preferences that corresponded to those experimentally found by peptide arrays. The predicted Arginine-binding positions and their associated pockets were analyzed in terms of shape, physicochemical properties, amino acid composition, and in-silico mutagenesis, providing structural rationalization for previously unexplained trends in kinase preferences toward Arg moieties.

Managing 2D and 3D Ligand Fragments from PDB & PubChem

Talapady N Bhat

NIST, MML, Gaithersburg, MD, 20899

Today's community working on structure-based drug discovery must interact with a variety of standalone and applications to harness structural information. This limitation promotes the need to define and develop easy to use rule-based methods and publicly available resources to manage, intersect structural and inhibitor data available in many independently maintained resources such the PDB and the PubChem. We have developed rule-based, automated method for managing, querying and integrating disparate structural information on fragments and also complete inhibitors. The method is called Chemical Block Layered Alignment of Substructure Technique (Chem-BLAST). It first recursively dissects chemical structures into blocks of fragments of significance to drug design. Highly re-usable fragments are called 'roots' and they are defined a) using rules that operate on atomic connectivity and b) familiar concepts used to define fragments in drug design. Then we classify the all the ligands in the database in terms of the 'roots' to develop first Chemical Resource Description Framework (RDF) and then chemical ontologies in the form of a 'tree' made up of 'hub-and-spoke'. The hubs act as anchor points to define targets in a query to the database. At each anchor point a user is able to inspect its structures using molecular images (no need to hazel with naming conventions for structures) of the 'roots' and make informed decisions on a subsequent query. The corresponding spokes are used by the Web tool to generate query results. Using this method a user is able to gradually build the fragment and ligand of his choice starting from few 'roots'. This technique was applied for (a) both 2-D and 3-D structural data for AIDS (http://bioinfo.nist.gov/SemanticWeb_pr2d/chemblast.do); (b) to all the 3D structures from the PDB which is now available from the RCSB/PDB Web site (http://www.rcsb.org/pdb/explore/externalReferences.do?structureId=3GGT) and to integrate them with the fragments of a couple of million 2D structures from the PubChem http://xpdb.nist.gov/chemblast/pdb.pl . Full description of the Chem_BLAST along with recent results and illustrations will be presented.

Escherichia coli Diaminopropionate ammonia lyase: Structure, substrate specificity and catalysis

Shveta Bisht[1], Rajaram V.[1], Bharath S.R.[1], Kalyani J.N.[2], Savithri H.S.[2], Murthy M.R.N.[1]

[1]Molecular Biophysics Unit, Indian Institute of Science, Bangalore, 560012, India, [2]Department of Biochemistry, Indian Institute of Science, Bangalore, 560012, India

Diaminopropionate ammonia lyase (DAPAL) is a prokaryotic pyridoxal 5' phosphate (PLP) dependent enzyme that catalyses the conversion of diaminopropionate (DAP) to pyruvate and ammonia. DAP is a non standard amino acid structurally similar to serine and has been found to occur in both prokaryotes and higher organisms. Here, we report the first crystal structure of DAPAL. The structure of the enzyme from *Escherichia coli* was determined in two forms: tetragonal (a = 85.6 Å, c ~ 207 Å) and monoclinic (a = 85.7 Å, b = 94.1 Å, c ~ 94 Å and β = 110.0°). Structures of *Ec*DAPAL co-crystallized or soaked with substrates (DL-DAP, D-Ser and L-Ser) were also determined in the tetragonal form. These structures revealed a topology typical of fold type II PLP-dependent enzymes with some key differences. Structure from the tetragonal form had PLP anchored to Lys^{77} by a Schiff base linkage. However, the structure from the monoclinic crystals did not have density for PLP and therefore, depicts an apo form of the enzyme. The active site of the apo structure was more disordered than that of the PLP bound form. A unique quaternary arrangement of protomers, so far not observed in other fold type II enzymes, was observed in *Ec*DAPAL. The structure revealed a disulfide bond near substrate entry channel, which is also a unique feature of *Ec*DAPAL. This unique disulfide might be important for maintaining the active site in a geometry required for catalysis. DAPAL was found to degrade both the D and L isomers of DAP. Analysis of active site revealed that the enzyme is likely to follow a two base mechanism for degradation of both isoforms of the substrate. *Ec*DAPAL co-crystallized with DL-DAP revealed an aminoacrylate reaction intermediate trapped at the active site. Comparative analysis of structures of *Ec*DAPAL and its complexes, and other fold type II PLP-dependent enzymes provides unique insights into the mode of substrate binding and mechanism of catalysis. The importance of Asp^{120} and Asp^{189} in the catalytic reaction was confirmed by site directed mutagenesis experiments. These residues are at hydrogen bonding distance to the substrate bound at the active site and are likely to be involved in the abstraction of the Cα proton from D and L forms, respectively, of the substrate.

A comprehensive comparison of protein structures and the correlation of structure with function

Spencer E Bliven*[1], Andreas Prlic[2], Philip E Bourne[3]

[1]Bioinformatics, University of California, San Diego, La Jolla, CA, 92093, [2]San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA, 92093-0505, [3]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA, 92093-0743

Numerous studies have attempted to characterize protein fold space. Existing classification schemes such as SCOP and CATH tend to view protein structures as belong to a finite number of discrete protein folds. Others prefer to view protein space as continuum of possible structures, of which only a few have been sampled. However, previous debate about the topic has been primarily qualitative and philosophical, with examples being cited in support of both discrete and continuous viewpoints[1-3]. In this study we comprehensively compare protein structures, with the goal of quantitatively assessing the continuity or discreteness of fold space.

A comprehensive comparison of all protein structures in the Protein Database (PDB) was recently performed based on the FATCAT algorithm for structural comparison[4]. Conceptually, the pairwise protein comparisons form a graph over all proteins, where the edges in the graph correspond to the similarity between proteins. By analyzing this graph it is possible to quantitatively measure the similarity between protein folds. We find that this graph is highly connected, providing support to the hypothesis that fold space is continuous.

We validate the protein structure graph by comparing it with existing classifications of protein folds (SCOP classification). Additionally, we map proteins to functional annotations and analyze the effect of structural similarity on protein function.

[1] Sadreyev et al. Curr Opin Struct Biol (2009) vol. 19 (3) pp. 321-8
[2] Sippl. Curr Opin Struct Biol (2009) vol. 19 (3) pp. 312-20
[3] Govindarajan et al. Proteins (1999) vol. 35 (4) pp. 408-14
[4] Prlic et al. Bioinformatics (2010) vol. 26 (23) pp. 2983-2985

Dynamic Protein-Protein Complexes: How alternative interactions create ensembles and how solution NMR and MD simulations can characterize them

Liqun Zhang[1], Hyeong J Lee[1], Prasanta K Hota[1], Matthias Buck[1-4]

[1]Case Western Reserve University, Dept. of Physiology and Biophysics, Cleveland, OH, 44106, [2]Case Western Reserve University, Dept. of Pharmacology, Cleveland, OH, 44106, [3]Case Western Reserve University, Dept. of Neuroscience, Cleveland, OH, 44106, [4]Case Western Reserve University, Case Comprehensive Cancer Center, Cleveland, OH, 44106

It is now recognized that protein-protein interactions in solution are dynamic, especially if the binding affinities are only moderately strong. Dynamics, originate in part from the population of alternative protein complex structures, e.g. one bound state that is in equilibrium with one or several alternative configurations. Here we present two protein complexes that exhibit fluctuations in solution: the plexin RBD:Rac1 and the EphA2:SHIP2 SAM:SAM complexes. A number of techniques are used to shift the equilibrium and/or to demonstrate the population of alternate complex configurations, including ion-pair swap mutagenesis, disulphide cross-linking and Paramagnetic Relaxation Enhancement (PRE) NMR. An ensemble approach is required to calculate the structures even when a wide collection of NMR restraints (chemical shifts, PREs, NOEs and RDCs) are available. The origin of the alternate structures is revealed by the different patterns of sidechain contacts that are possible in the complexes. These interactions and the dynamics of the complexes are further analyzed by microsecond unrestrained molecular dynamics simulations carried out on the MD optimized supercomputer Anton. The functional significance of protein complex dynamics is discussed.

Salmonella Virulence Protein SopB Down-regulates Activity of Host Rho GTPase Cdc42 by Mimicking a Nucleotide Dissociation Inhibitor

Brianne Burkinshaw, Natalie Strynadka

University of British Columbia, Biochemistry, Vancouver, V6T1Z3, Canada

SopB is an essential Type III secreted Salmonella effector protein with phosphoinositide phosphatase activity and a GTPase binding domain that interacts with host Cdc42, an essential Rho GTPase that regulates critical events in eukaryotic cytoskeleton organization and membrane trafficking. Structural and biochemical analysis of the SopB GTPase binding domain in complex with Cdc42 shows for the first time that SopB structurally and functionally mimics a host guanine nucleotide dissociation inhibitor (GDI) by contacting key residues in the regulatory switch regions of Cdc42 and preventing Cdc42 nucleotide exchange.

Unique crystal structure of protozoan parasite Entamoeba histolytica malic enzyme

Arindam Chakrabarty, Debjyoti Dutta, Amit K Das, Sudip K Ghosh

Indian Institute of Technology, Dept. of Biotechnology, Kharagpur, 721302, India

Malic enzyme with ubiquitous distribution in nature has an immense importance in the biological system. The enzyme catalyzes oxidative decarboxylation of L-malate to pyruvate and carbon dioxide with reduction of NAD(P)+ to NAD(P)H using a divalent metal ion Mn2+/Mg2+. The cytosolic malic enzyme of amitochondriate protozoan parasite Entamoeba histolytica was cloned from its genomic DNA, purified to homogeneity and biochemically characterized. Here we report the crystal structure of NADP+ -dependent E. histolytica malic enzyme at a resolution of 2.9Å with a bound glycerol and sulphate molecule. The dimeric enzyme consists of two 53 kDa subunit which are perfectly superimposible to each other and each monomer consists of five distinct domains. N-terminal and C-terminal ends of the protein are folded to form a single domain A. Domain B acts as an interphase in dimer formation and bear a conserved signature 45GKI47 that forms a knot and lock the monomers with each other. This hinge region assists the easy movement of the domain A. Domain C consists of NADP+ binding Rossmann fold and the core structure includes two layers of α/β/α. Tyr 65 of one monomer and a Lys 120 of another monomer form the active site region of the enzyme. Metal binding site and the substrate binding site are present within the domain D which consists of three layers of α/β/α type Rossmann fold. Kinetic analysis of malic enzyme from E. histolytica shows its favour for NADP+ as coenzyme. A sulphate moiety that mimics PO4 group binds to 226GAG228 and 252DSK254 site elucidating GXGXXA and DSKG conserved motif respectively. Glycerol molecules are found to adhere within the crystal structure as the cryo solution used had 12% (v/v) glycerol. L-malate binding sites of each monomer are occupied by glycerol molecules. Positions of three glycerol molecule in a monomer show the possible channel for the entry of substrate within the active site. This is the first report of parasite malic enzyme crystal structure.

Structural Biology Meets Systems Biology: Gauging the Systemic Impact of Single Nucleotide Polymorphisms

Tammy Cheng[1], Linda Jeffery[2], Lucas Goehring[3], Yu-En Lu[4], Jacqueline Hayles[2], Bela Novak[5], Paul Bates[1]

[1]Cancer Research UK London Research Institute, Biomolecular Modelling Laboratory, London, WC2A 3LY, United Kingdom, [2]Cancer Research UK London Research Institute, Cell Cycle Laboratory, London, WC2A 3LY, United Kingdom, [3]Max Plank Institute, Dynamics and Self-Organization, Göttingen, 1037073, Germany, [4] University of Cambridge, Computer Laboratory, Cambridge, CB3 0FD, United Kingdom, [5]University of Oxford, Department of Biochemistry, Oxford, OX1 3QU, United Kingdom

Small changes in protein structure, such as non-synonymous single nucleotide polymorphisms (nsSNPs), can have a large impact on cellular behaviour. To understand how a change at the protein level eventually affects a cell's phenotypic outcome is, however, not trivial. This is because multi-level information needs to be considered in order to obtain analytical results with physiological meanings. With respect to the fact that the idea of integrating protein behaviour and pathway dynamics in terms of gauging the systemic impact of nsSNPs is still in its infancy, we investigated its practicality by formulating mathematical models to studying point mutations in two biological systems: G2 phase to mitosis transition in fission yeast and human MAPK pathway.

At the pathway level, we construct ordinary differential equations (ODEs) to simulate pathway dynamics. We perform sensitivity analysis, i.e. systematically perturbing each rate constant in the ODEs and measure its impact on the expression of a downstream reporter protein that is tightly linked to the phenotype. The sensitivity of the rate constants thus reflects the potential of each protein to alter the dynamics of a system.

At the protein level, we investigate the impact of point mutations on protein stability. Each stability change is considered as an absolute perturbation introduced into the rate constants associated with each mutant protein. Hence stability change at the protein level is transformed to changes of the expression profile that reflect changes of pathway dynamics – mathematically, we formulated this into SIF (systemic impact factor) score to quantify the likely phenotypic outcomes of point mutations.

In the yeast model, our *in silico* SIF scores correlate with *in vivo* cell lengths; in the human model, the SIF core clusters point mutations into subgroups that reflects their disease mechanisms. This suggests that systemic impact is a result of protein stability change and pathway perturbation.

Our study as a whole shows that it is beneficial to combine multi-level knowledge and perform multi-level mathematical modelling to understand the effects of small structural changes on cellular behaviour.

AtlasCBS: A web server to graphically represent the content of chemico-biological databases.

Alvaro Cortes[1,2], Federico Gago[2], Antonio Morreale[1], Celerino Abad-Zapatero[1,2,3]

[1]Center for Molecular Biology, Bioinformatics Unit, Madrid, 28049, Spain, [2]University of Alcala, Dept. of Pharmacology, Alcala de Henares , 28871, Spain, [3]Center for Pharmaceutical Biotechnology. University of Illinois., Dept. of Medicinal Chemistry and Pharmacognosy, Chicago, IL, 60607

The notion of Ligand Efficiency (LE), broadly defined as the affinity of a ligand towards its biological target (given by K_i, IC_{50}) scaled by the corresponding size (Mw, No. of non-hydrogen atoms), is gaining acceptance in the medicinal chemistry community as an efficiency yardstick. This is particularly true for ranking the efficiency of small fragments in approaches involving Fragment-Based libraries. A more encompassing definition of ligand efficiency was presented[1] that included two complementary Ligand Efficiency Indices (LEIs): BEI, SEI. The first one combines affinity with Mw (BEI = pK_i/(Mw/1000) and the second one affinity with polarity (SEI = pK_i/(PSA/100).
The combination of these two variables in a plane allows a mapping of the chemical series for a target in a plane as illustrated[2].

More generalized definitions of LEIs have been presented recently[3] suggesting that complementary pairs of variables (size, polarity) can be used to represent the content of Chemico-Biological Space in an atlas-like representation.

We present here the first AtlasCBS server as a proof-of-concept of the possibility of representing the content of various chemico-biological databases in a succession of electronic pages with different efficiency axes and at different scales, using the framework of LEIs as variables.

The web application is based on a Model-View-Controller architecture using standard web technology such as Java servlets and applets and AJAX technology that allows a feature-rich design to represent dynamically the contents of the databases.

The server is supported by a MySQL database and a Java Servlet container, while the web interface is composed of Applets and AJAX applications.

We propose that this representation of the content of various Chemico-Biological Databases can be very useful to navigate the Chemico-Biological Space and that when broadly tested and used could prove to be an invaluable tool to expedite drug discovery. A demonstration of the possibilities of the AtlasCBS concept and application will be presented at the meeting.

[1] Abad-Zapatero, C.; Metz, J.M. Drug Discovery Today. 2005, 10, 464-469.
[2] Abad-Zapatero, C. Expert Opinion in Drug Discovery. 2007, 2, 469-488.
[3] Abad-Zapatero et al. Drug Discovery Today. 2010, 15, 804-811.

Treatment of cancer and autoinflammatory disease for over sixty years – structural basis for folate and antifolate trafficking via human folate receptors

Charles E Dann III[1], Ardian S Wibowo[1], Kristen M Reeder[1], Joshua J Carter[1], Wuyi Meng[1], Faming Zhang[1], Manohar Ratnam[2]

[1]Indiana University, Department of Chemistry, Bloomington, IN, 47405-7102, [2]The University of Toledo College of Medicine, Department of Biochemistry and Cancer Biology, Toledo, OH, 43614

Human folate receptors (hFRs) are high affinity, glycosylphosphatidylinositol-anchored membrane surface proteins responsible for the uptake of both natural folate metabolites and therapeutic antifolate drugs. Antifolates have been approved by the FDA for treatment of patients with cancer and autoinflammatory disease since the late 1940s. While cytotoxicity is problematic at higher doses, classical and newer generation antifolates – most often methotrexate and pemetrexed, respectively – are still widely used to treat cancers derived from epithelial cell lineages (*e.g.* leukemia, choriocarcinoma, ovarian and breast cancers) and autoinflammatory diseases (*e.g.* rheumatoid arthritis, psoriasis, Crohn's disease). Humans possess at least three distinct transport pathways for the uptake of antifolates. As much of the antifolate toxicity could be ameliorated by designing therapeutics specifically taken up via the hFRs, we have set out to determine structures of hFRα and hFRβ both alone and in ligand complexes at pH values indicative of key physiological transport states. Our presentation, representing the first structures of folate receptors, will include at least six novel structural models illustrating the molecular determinants for folate and antifolate binding as well as the conformational changes at key stages of ligand transport. These structures will be the basis for the development of the next generation of antifolate therapeutics with the desired limited cytotoxicity profile in normal tissues.

Corresponding author contact: cedann@indiana.edu
http://www.indiana.edu/~dannlab/

Crystal Structures of eIF4E in Complex with Bn^7-GMP

Teresa De la Mora, Yan Jia, Chunqi Hu, Carston R Wagner, Barry C Finzel

University of Minnesota, Medicinal Chemistry, Minneapolis, MN, 55455

The eukaryotic Initiation Factor (eIF) is a multicomponent complex of proteins that can be assemble around the capped mRNA to initiate translation. Protein 4E (eIF4E) is the least abundant of the initiation factors in this complex and considered to be the limiting factor for recruitment of the ribosome to the mRNA start site [1]. It binds directly to the N^7-methylated guanosine cap at the mRNA 5' end. Overexpression of eIF4E in transgenic mouse models promotes tumorigenesis [2] and overexpression has been related to disease progression [3]. Efforts to explore the potential of eIF4E inhibitors to provide benefit as cancer chemotherapeutic agents has led to the discovery and characterization of Bn^7-GMP (7-benzyl-guanosine monophosphate), an analog of the mRNA cap that binds to eIF4E with a Kd of 0.8μM [4]. We have initiated structural studies with inhibitor molecules and mouse recombinant eIF4E [5] aimed at enhancing understanding of the binding of mRNA cap-analogs with eIF4E and to promote further compound development. Here, we report the first apo structure of the mouse recombinant eIF4E, and a surprising complex that appears to be a demonstration of crystal-facilitated organic chemistry resulting in a covalent adduct of two Bn^7-GMP molecules spanning adjacent binding sites in the novel $P2_1$ crystal form.

1. Marcotrigiano, J. & Burley, S.K. Advances in Protein Chemistry (2003) 61:269-297
2. Ruggero, D., et al. Nat. Med. (2004) 10:484-486
3. De Benedetti, A. & Graff, J. R. Oncogene (2004) 23:3189-3199
4. Jia, Y. et al. Eur J. Med Chem (2010) 45:1304-13
5. Ghosh, P., et al. Protein Expression and Purification (2008) 60:132-139

We thank the staff at GM-CAT at APS for access to beamline.

Metal-mediated affinity and orientation specificity in a computationally designed protein homodimer

Bryan S Der[1], Michael J Miley[2], Mischa Machius[2], Brian Kuhlman[1]

[1]University of North Carolina at Chapel Hill, Biochemistry & Biophysics, Chapel Hill, NC, 27599-3010, [2]University of North Carolina at Chapel Hill, Pharmacology, Chapel Hill, NC, 27599-7365

Many computationally designed protein-protein interactions have weak affinity and do not conclusively bind in the desired orientation. Incorporating metal binding sites at the target interface may be one approach for increasing affinity and specifying the binding mode, thereby improving robustness of designed interactions for use in research and medicine. Here we describe the Rosetta-based computational design of a protein monomer to form a metal-mediated homodimer. The design protocol uses a geometric hashing algorithm, RosettaMatch, to design zinc binding sites, then symmetric protein design to optimize the protein interface. Experimental characterization of this metal interface design, named MID1, validates that metal binding sites can improve affinity and orientation specificity in rationally designed protein interfaces: in the presence of zinc MID1 forms a tight dimer with a dissociation constant less than 30 nM, without zinc the dissociation constant is 4 μM. While a crystal structure of MID1-apo (without metal) revealed an unforeseen binding orientation (RMSD from model = 8.4 Å), crystal structures of MID1-zinc and MID1-cobalt demonstrate metal-driven rotation toward the intended binding orientation (RMSD from model = 2.1 Å and 2.3 Å, respectively). Atomic-level details of the modeled versus experimentally determined interface suggest that the computed protein-protein contacts are only partially accurate, though the metal-mediated interaction still features high affinity and the desired binding orientation.

Regulation of the histone acetyltransferase activity of hMOF via autoacetylation of Lys274

Bingfa Sun[1], Shunling Guo[2], Qingyu Tang[1], Zhiqi Xiong[2], Chen Zhong[1], Jianping Ding[1]

[1]Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, State Key Laboratory of Molecular Biology, Shanghai, 200031, China, [2]Institute of Neuroscience, State Key Laboratory of Neuroscience, Shanghai, 200031, China

Human males-absent-on-the-first (hMOF, also called MYST1 or KAT8) is a MYST family histone acetyltransferase (HAT) responsible for global H4K16 acetylation in cells and plays an important role in early development and cancers, yet the regulatory mechanism of its activity is largely unknown. We determined the crystal structure of hMOF catalytic domain in apo-form, which combined with the mass spectrometry data reveal that acetylation occurs at a strictly conserved residue Lys274 at the catalytic active site. Our structural and biochemical data demonstrate that Lys274 is autoacetylated in an intra-molecular manner and the acetylation of Lys274 potentiates the HAT activity of hMOF, and indicate that this modification is critical in maintaining an active conformation of the catalytic site to allow proper binding/positioning of the substrate lysine. In addition, the in vivo cell biological data demonstrate that the overexpressed full-length hMOF is also autoacetylated at Lys274, and its autoacetylation level can be regulated by HDACs from multiple families, indicating that the acetylation level and hence the HAT activity of hMOF might be regulated in response to environment stimuli. Moreover, it is shown that the acetylation of Lys274 is critical for the regulation of global H4K16 acetylation level. Taken together, our results show for the first time that hMOF is autoacetylated both in vitro and in vivo, and provide structural basis for how autoacetylation at a single lysine residue could regulate the HAT activity of hMOF, which reveals a novel regulatory mechanism for HATs that might also apply to other members of the MYST family.

A structural analysis of the regulatory domain from the cGMP-dependent protein kinase Iα

Brent W Osborne[1], Andrew T Menke[1], Christian K Nickl[1], Donald K Blumenthal[2], Wolfgang R Dostmann[1]

[1]University of Vermont, Pharmacology, Burlington, VT, 05405, [2]University of Utah, Pharmacology, Salt Lake City, UT, 84112

The cGMP-dependent protein kinase (PKG) has two tandem cyclic nucleotide binding (CNB) domains which act as the primary intracellular receptor for cGMP. PKG exhibits a homodimeric rod-like structure which undergoes significant molecular rearrangements upon the binding of cGMP. However, a detailed structural analysis of the core regulatory elements inherent to PKG is still required. We recently solved a crystal structure of the two cGMP binding sites from PKG Iα in order to highlight the atomic details of the regulatory domain. This PKG78-355 structure is free of cGMP and presents the protein in an elongated conformation. A surprising dimeric arrangement between PKG78-355 protomers is orchestrated via hydrophobic contacts between a novel helical element C-terminal to the second cGMP binding site (the switch helix) and the opposite CNB domain B. Small angle X-ray scattering (SAXS) of PKG78-355 suggests an overall molecular dimension of ~130 Å, consistent with the maximal linear dimension observed in our crystal structure. Upon incubation with cGMP, PKG78-355 contracted to ~95 Å. This molecular compaction was not observed in a construct lacking the switch helix (PKG78-326), suggesting the additional importance of the switch helix in mediating cGMP-specific conformational changes inherent to the regulatory domain. Overall, these studies provide the first atomic resolution model of tandem cGMP binding domains and expand our understanding of the allosteric mechanisms surrounding PKG activation.

Evolution of protein structures from the perspective of their super-secondary structure building blocks.

Joseph M Dybas, Andras Fiser

Albert Einstein College of Medicine, Department of Biochemistry, Department of Systems and Computational Biology, Bronx, NY, 10461

We have established a library of protein super-secondary structure motif (S-motif) building blocks in order to study the evolution of protein structures from a modular perspective. Our goal is to elucidate the rules by which these modular units are combined to produce the array of structures in the fold space.

S-motifs are defined as two secondary structures, successive in primary sequence and connected by a loop. The S-motif library was developed by collecting and grouping all the constituent S-motifs from each structure in the PDB. The number of representative S-motifs in the library has remained essentially constant (~2300 clusters) since approximately the year 2000, despite the addition of novel folds into the SCOP database in the last 11 years. Thus, the S-motif library seems to be nearly completely saturated and the collection of S-motif building blocks is sufficient to combine to produce any existing or novel fold in the protein fold space. This completeness and saturation of the library, along with the relatively small number representatives, makes the S-motifs applicable to the study of protein modularity.

There is a characteristic distribution of S-motifs in the fold space. Many of the S-motifs are unique and present in one or a very few number of folds, whereas a smaller but significant number of S-motifs are ubiquitous and found in a relatively large number of folds. This characteristic is similar for the distribution of consecutive and non-consecutive pair-wise associations of S-motifs in the fold space. The unique single S-motifs and associations of S-motifs are being used to address the problem of classifying structures into folds and structural families. However, the ubiquitous S-motifs and associations of S-motifs, where similar structural motifs are conserved within differently classified folds, suggest that there is a distinct structural overlap between some folds and the notion of a discrete fold space should be reevaluated. Further, we hypothesize that the conservation of S-motifs throughout fold space can inform a systematic evaluation of the evolution and phylogeny of protein folds.

Ongoing efforts exploit the relationship between the S-motifs and protein folds to address problems in protein design and structure prediction. For instance, the S-motif library is currently used in an algorithm that employs NMR chemical shift data to select S-motifs to predict local geometry of unknown structures.

Crystal Structure of NLRP4 Pyrin Domain

Clarissa Eibl[1], Manuel Hessenberger[1], Sandra Puehringer[2,3], Kay Diederichs[4], Robert Schwarzenbacher[1]

[1]University of Salzburg, Molecular Biology, Salzburg, 5020, Austria, [2]Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Macromolecular Crystallography, Berlin, 12489, Germany, [3]Freie Universitaet Berlin, Biology and Chemistry, Berlin, 14195, Germany, [4] University of Konstanz, Biology, Konstanz, 78457, Germany

Nod like receptors (NLRs) are cytoplasmatic immune receptors that recognize, akin to the membrane-bound TLRs, pathogens and initiate host defense pathways [1]. Their crucial function in immune signaling is exemplified by their implication in various genetic disorders [2].
For most of the 22 human NLRs, precise function and structure are unknown and information on activating ligands and the signaling cascades they initiate is missing.
Although NMR structures for the pyrin domains (PYDs) of human NLRP1, NLRP7 and mouse NLRP10 were reported they have not yet revealed the mechanism of their differential interaction modes

In this work we present the first crystal structure of a PYD, more precisely the PYD of NLRP4 (PDP ID 3PJH), which we solved by MAD technique to 2.3Å resolution.
NLRP4 PYD forms a typical death domain fold with six α-helices tightly packed around a hydrophobic core, which is characteristically for PYDs. Although the primary sequence conservation between PYDs is generally low, structural comparison display high similarities to the other PYDs. The biggest differences between NLRP4 PYD compared to the PYDs of NLRP1, NLRP7 and NLRP10 lays in α-helix 3. Whereas α-helix 3 is well defined in NLRP4 and NLRP7 it is replaced by a flexible loop in NLRP1 and NLRP10. Additional differences can be seen in the electrostatic surface potential where NLRP4 displays a huge positive charged surface and exclusively exhibits a second positive charged patch.
We assume that the differences in α-helix 3 and the electrostatic surface are critical for diverse signaling of distinct members of the NLR receptor family.
Future work will clarify the role of these crucial immune receptors in innate immunity, inflammation, and related diseases.

References

1. Mariathasan, S., et al., Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf. Nature, 2004. 430(6996): p. 213-8.
2. Shaw, P.J., M.F. McDermott, and T.D. Kanneganti, Inflammasomes and autoimmunity. Trends Mol Med, 2011. 17(2): p. 57-64.
3. Fiorentino, L., et al., A novel PAAD-containing protein that modulates NF-kappa B induction by cytokines tumor necrosis factor-alpha and interleukin-1beta. J Biol Chem, 2002. 277(38): p. 35333-40.
4. Jounai, N., et al., NLRP4 negatively regulates autophagic processes through an association with beclin1. J Immunol, 2011. 186(3): p. 1646-55.

The Second Log

Richard J Feldmann

Global Determinants, Inc., Derwood, MD, 20855

In the early 1970s, I helped to collect, organize and bring into the open the second log (i.e. 10 to 100) of protein and nucleic acid structures. It turned out that then the way to make a collection was to first collect crystallographers, so I went around the world meeting them and asking them to contribute their structures. Thirty years later in the early 2000s, I found myself processing the second log of genomic structures.

Why is it that the collection of facts (i.e. this sequence of amino acids has this structure or this sequence of nucleic acids has these genes) is so much easier than understanding how the components interact to determine the structure? Why is behavior so much more difficult to understand than structure? What aspect of our Western culture leads us to this impasse?

Once a process is initiated, people rush along doing their thing for decades in the past and probably for decades in the future. When is there enough data? Jane Richardson was able to use my 2-log structure collection to do her Anatomy and Taxonomy monograph and yet now there are more than 74,888 structures in the PDB without any real understanding of protein folding. The global culture is rushing towards determining the genomic sequence of every individual but nobody understands what this data means.

Structure-Function Map of a Heptameric Module in the Nuclear Pore Complex

Javier Fernandez-Martinez*[1], Jeremy Phillips*[2], Matthew Sekedat[3], Ruben Diaz-Avalos[4], Javier Velazquez-Muriel[2], Josef Franke[1], Rosemary Williams[1], David Stokes[4], Brian T Chait[3], Andrej Sali[2], Michael Rout[1]

[1]The Rockefeller University, Laboratory of Cellular and Structural Biology, New York, NY, 10065, [2]University of California, San Francisco, Department of Bioengineering and Therapeutic Sciences, San Francisco, CA, 94107, [3]The Rockefeller University, Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, New York, NY, 10065, [4] The New York Structural Biology Center, The New York Structural Biology Center, New York, NY, 10027

The nuclear pore complex (NPC) is a multiprotein assembly that serves as the sole mediator of nucleocytoplasmic exchange in eukaryotic cells. Here, we use a new integrative approach to determine the structure of an essential component of the yeast NPC, the ~600 kDa heptameric Nup84 complex, to a precision of 1.5 nm. The structure was solved by satisfaction of spatial restraints derived from a diverse set of negative stain EM and nucleoporin domain mapping data. Phenotypic data was mapped onto the complex to generate a functionally informative structure, allowing us to find regions that stabilize the NPC's interaction with the NE membrane and that connect the complex to the rest of the NPC. We suggest a scenario for the evolutionary origin of the Nup84 complex through a series of gene duplication and loss events. Our work demonstrates that integrative approaches based on low resolution data can generate functionally informative structures at intermediate resolution.

*These authors contributed equally to this work.

The Use of Distance Geometry for Substructure Searching and Local Superposition in DNA, RNA, and Protein-Polynucleotide Complexes

Jeffrey R Van Voorst, Barry C Finzel

University of Minnesota College of Pharmacy, Medicinal Chemistry, Minneapolis, MN, 55455

A software suite and database is described that enables complex substructure searching across a wide selection of macromolecular structures in the Protein Data Bank. By using backbone distance geometry encoding distances between all alpha-carbon atoms, it is possible to identify not only individual segments of structure with distinct geometry, but also higher order assemblies of segments that share a similar geometry. When the searchable distance geometry also encodes distances between adjacent molecules in crystallographic structures, queries can identify similar scaffolds engaged in similar in intermolecular contacts.

We propose extensions to our searchable protein database that extend substructure searching to structures involving DNA and RNA. We show that a similar distance geometry representation of polynucleotide backbones arising from ribose C1' atoms correlates most strongly with substructural similarity. By including distance geometry encoding distances between nucleic acid residues in different polynucleotide chains, queries can return one-to-one matches to other examples of similar tertiary structure. We show that a unified database incorporating both protein and RNA components of distance geometry from crystallographic resources can be mined for examples of similar protein-RNA and protein-DNA "interaction motifs" using this software.

When conveniently integrated into a comprehensive molecular modeling environment (e.g. Coot), substructure searching can provide a powerful means to examine and validate local protein geometry, to accelerate electron density fitting, or to overlay and compare divergent protein structures with only localized similarity. Such an integration will provide a powerful tool to aid in the interpretation of macromolecular interactions and assembly.

PREDICTING PROTEIN STRUCTURE WITH SUPERSECONDARY-STRUCTURE BUILDING BLOCKS AND BASIC NMR DATA

Andras Fiser, Vilas Menon, Joeph Dybas

Albert Einstein College of Medicine, Systems and Computational, Bronx, NY, 10461

Recently we have established a supersecondary structure library (Smotifs) from the building blocks of known protein structures. We have observed that this library has saturated over time, and since about the late ninties all new folds can be decomposed into Smotifs that are part of earlier solved experimental structures. It appears that new folds do not require the emergence of novel Smotifs but a novel combination of existing Smotifs. This provides a hypothesis that all protein structures must be possible to build from the available set of Smotif building blocks. In recent years, structural modeling approaches have seen dramatic successes in predicting protein structures based on sequence data alone. However, these methods have limited success rates when the sequence similarity of the protein to be characterized to any given possible template structure is low or undetectable. As a result, predicting the structures of such proteins requires additional information besides the sequence. Here, we present a method that takes chemical shift and residual dipolar coupling data (both of which are easily obtainable from NMR experiments) to generate backbone structures without relying on any sequence signal.

The basis for this method is a library of supersecondary structure building blocks (Smotifs), each consisting of two successive secondary structures connected by a loop. Given the success of molecular fragment replacement approaches to structure prediction, our goal was to identify and assemble these relatively large, structurally meaningful Smotif building blocks to predict full protein structures. Starting from the experimental data for the query protein, we select approximately 5-20 candidates for each putative Smotif of the final structure, based on local loop configuration (using chemical shift data) and secondary structure orientation (using RDC data). Since a 150-residue protein, on average, contains about 10 Smotifs, we can fully enumerate and score combinations of these Smotifs. We found that a relatively simple scoring function with three global measures (compactness, statistical pair-potential, backbone hydrogen bond potential and solvation) is sufficient to identify accurate solutions.

We tested our method using experimental data for a variety of proteins (calmodulin, thioredoxin, integrasse like domain, tetratricopeptide repeaet domain and sensory rhodopsin), obtaining models within ~3A RMSD of the experimental solution structure. These results show that, by identifying large building blocks, we can model significant portions of unknown proteins without the need for sequence signals. In addition, the degree of success in predicting the structure of sensory rhodopsin suggests that this method shows promise in modeling membrane-bound proteins, a severely under-represented class of proteins in the set of solved structures in PDB.

Prediction of leucine-rich nuclear export signal containing proteins with NESsential

Szu-Chin Fu[1], Kenichiro Imai[2], Paul Horton[2]

[1]University of Tokyo, Department of Computational Biology, Graduate School of Frontier Science, Chiba, 277-8561, Japan, [2]National Institute of Advanced Industrial Science and Technology (AIST), Computational Biology Research Center (CBRC), Tokyo, 135-0064, Japan

The classical nuclear export signal (NES), also known as the leucine-rich NES, is a protein localization signal often involved in important processes such as signal transduction and cell cycle regulation. Although 15 years has passed since its discovery, limited structural information and high sequence diversity have hampered understanding of the NES. Several consensus sequences have been proposed to describe it, but they suffer from poor predictive power. On the other hand, the NetNES server provides the only computational method currently available. Although these two methods have been widely used to attempt to find the correct NES position within potential NES-containing proteins, their performance has not yet been evaluated on the basic task of identifying NES-containing proteins. We propose a new predictor, NESsential, which uses sequence derived meta-features, such as predicted disorder and solvent accessibility, in addition to primary sequence. We demonstrate that it can identify promising NES-containing candidate proteins (albeit at low coverage), but other methods cannot. We also quantitatively demonstrate that predicted disorder is a useful feature for prediction and investigate the different features of (predicted) ordered versus disordered NES's. Finally, we list 70 recently discovered NES-containing proteins, doubling the number available to the community.

NMR structural studies of the complex between lipid binding protein OSBP and ER membrane protein VAP-A

Kyoko Furuita[1,5], JunGoo Jee[1,2], Harumi Fukada[3], Masaki Mishima[1,4], Chojiro Kojima[1,5]

[1]Nara Institute of Science and Technology, Graduate School of Biological Sciences, Ikoma, Nara, 630-0192, Japan, [2]Tokyo Metropolitan University, Center for Priority Areas, Hachioji, Tokyo, 192-0397, Japan, [3]Osaka Prefecture University, Graduate School of Life and Environmental Sciences, Sakai, Osaka, 599-8531, Japan, [4] Tokyo Metropolitan University, Graduate School of Science and Engineering, Hachioji, Tokyo, 192-0397, Japan, [5]Osaka University, Institute for Protein Research, Suita, Osaka, 565-0871, Japan

Oxysterol-binding protein (OSBP) is a cytosolic receptor of cholesterol and oxysterols such as 25-hydroxycholesterol, and has been implicated to play a role in vesicle transport, lipid metabolism, and signal transduction. OSBP is usually localized at the endoplasmic reticulum (ER). The ER localization of OSBP is regulated by the complex formation between OSBP and type-II integral ER protein VAMP-associated protein-A (VAP-A), and essential for the stimulation of sphingomyelin synthesis by 25-hydroxycholesterol. OSBP is composed of N-terminal Pleckstrin Homology domain, central FFAT motif and C-terminal lipid binding domain, while VAP-A is composed of N-terminal Major Sperm Protein (MSP) domain, central coiled-coil domain and C-terminal transmembrane domain. Complex formation between OSBP and VAP-A is regulated by the interactions of the FFAT motif of OSBP with the MSP domain of VAP-A. Many lipid-binding proteins contain FFAT motifs, most of them, including OSBP, have acidic regions preceding their FFAT motifs.
In this study, to delineate the interaction mechanism between VAP-A and OSBP, we determined solution structure of the complex between the VAP-A MSP domain (VAP-A_{MSP}) and the OSBP fragment containing the FFAT motif ($OSBP_F$). VAP-A_{MSP} has an immunoglobulin-like β-sandwich fold consisting of seven β-strands and one α-helix. The FFAT motif of $OSBP_F$ adopts an extended β-strand-like conformation and is bent at the C terminus of the FFAT motif. Five of six conserved residues in the FFAT motif directly interact with VAP-A_{MSP}. Electrostatic and hydrophobic interactions and backbone-backbone hydrogen bonds stabilize the complex. By combining NMR relaxation and titration, isothermal titration calorimetry, and mutagenesis experiments with structural information, we further elucidated the detailed roles of the FFAT motif. Our results show that $OSBP_F$ is disordered in the free state, and VAP-A_{MSP} and $OSBP_F$ form a final complex by means of intermediates, where electrostatic interactions through acidic residues, including an acid patch preceding the FFAT motif of $OSBP_F$, probably play a collective role.

Autophosphorylation of CK2α at the N-terminal region is required for its catalytic activity

Miwako K Homma[1], Masaaki Oyama[2], Junko Yamaki[1], Yuko Hata[2], Yoshimi Homma[1]

[1]Fukushima Medical University School of Medicine, Department of Biomolecular Science, Fukushima , 960-1295, Japan, [2]Institute of Medical Science at the University of Tokyo, Medical Proteomics Laboratory, Tokyo, 108-8639, Japan

Protein kinase CK2, formerly known as casein kinase II, is a ubiquitous protein serine/threonine kinase that plays a central role in the regulation of variety of cellular processes including cell proliferation, transformation and apoptosis. This kinase is typically found in tetrameric complexes consisting of two catalytic α and/or α', and two regulatory β subunits. Although it has been considered to be a constitutively active enzyme, we have reported the cell-cycle dependent association of cellular CK2 with adenomatous polyposis coli (APC) protein that regulates CK2 activity, and have identified its downstream target as eukaryotic translation initiation factor 5 (eIF5). In this study, we demonstrate that CK2α becomes phosphorylated following stimulation of G0-arrested cells with FBS, using synchronized human normal fibroblasts. These phosphorylation levels are decreased when cells are treated with pharmacological inhibitors toward CK2, indicating that CK2 activities are partially required for phosphorylation of CK2α as well as β *in vitro*. Quantitative mass spectrometry analysis using recombinant CK2 reveals that it autophosphorylates multiple sites including its predicted phosphorylation sites as well as tyrosine residues in CK2α. Site-directed mutagenesis indicates that the catalytic activities of the recombinant CK2α proteins with phosphorylation site mutants at N-terminal region are reduced in vitro. The mechanism of autophosphorylation was investigated by measuring phosphoryl transfer to kinase-dead CK2 mutant, in which Lys68 around the ATP binding sites were converted to alanine, by wild type kinase tagged with His or FLAG, indicating that autophosphorylation occurs through an intramolecular mechanism. These results suggest an important role of the N-terminal phosphorylation in CK2α for its activity, and those residues including tyrosine may form part of a novel N-terminal structural unit that functions to sustain its enzymatic activity. Also, we performed immunoprecipitation of CK2 alpha complex followed by mass spectrometry analysis to identify CK2-associating proteins in relation to the progression of cell cycle.

Hybrid de novo modeling for 4 Å resolution cryo-EM maps of viruses

Corey F Hryc[1,2], Matthew L Baker[2], Dong-Hua Chen[2], Rui Zhang[1,2], Yao Cong[2], Frank DiMaio[3], David Baker[3], Wah Chiu[1,2]

[1]Baylor College of Medicine, Graduate Program in Structural and Computational Biology and Molecular Biophysics, Houston, TX, 77030, [2]Baylor College of Medicine, National Center for Macromolecular Imaging, Department of Biochemistry, Houston, TX, 77030, [3]University of Washington, Department of Biochemistry, Seattle , WA, 98195

Recent improvements in single particle electron cryomicroscopy (cryo-EM) have led to 3-D reconstructions of several viruses at near-atomic resolution (3.5 - 5 Å). At this resolution, model building tools have become essential to yield detailed information about protein folds, and molecular interactions in capsid stabilization and morphogenesis. These structural details help elucidate the functional roles of key domains and amino acids in virus assembly, morphogenesis and host cell infection.

To model proteins in a virus, a new protocol has been implemented using a combination of Gorgon for initial model construction and Rosetta for model optimization. To start constructing an initial model of a protein component without any structural homolog, Gorgon's *de novo* modeling tools are used to identify secondary structure elements (SSEs), generate a sequence to structure correspondence, place C-α atoms in SSEs and generate loops to connect SSEs. When all C-α atoms have been placed, the first step of model optimization uses Gorgon to correct bond length and to assure that atoms fit within the density. Following this, further model optimization for a single protein component in the virus is done with Rosetta. This fixes geometry and minimizes the energy of a single protein while using multiple constraints, such as the density map or homologous domains if available. Final optimization then uses Rosetta to eliminate any steric clashes among the proteins in the asymmetric unit.

In one example, the P22 bacteriophage capsid was modeled in two distinct functional states. This led to insights into movements of amino acid stretches, along with the shifting of their interactions in other parts of the neighboring proteins to accommodate this movement (Chen et al, 2011). Furthermore, a map of the P22 procapsid at 3.8 Å revealed the positively charged scaffolding protein interacting with the negatively charged N-terminal region from the coat protein. In a second example, Venezuelan Equine Encephalitis virus (VEEV) was modeled using this hybrid procedure, revealing key interactions between two glycoproteins and nucleocapsid proteins,(Zhang et al, 2011). This suggested a mechanism for the initial stage of nucleocapsid core formation and shed light on virulence attenuation, host recognition and neutralizing activities.

PFAT: Protein Functional Annotation based on TOPOFIT

Amit Upadhyay, Valentin Ilyin

Boston College, Biology, Chestnut Hill, MA, 02467

Structural genomics efforts have generated a large number of protein structures. These efforts have focused on proteins with low homology to existing structures in order to increase coverage. This has however resulted in deposition of many structures annotated as “hypothetical” or “unknown function”, which is is about one third of the structures solved by SG. These structures will be useful only after we can assign functions and understand their biological roles. A number of methods have been developed for functional annotation of proteins. Many present computational methods of functional annotation are based on the idea that proteins with similar sequence and structural features are likely to perform same function, but these methods cannot be used accurately for structural genomics proteins since target selection involves a 30% identity cutoff to protein families with a known structure. A number of structure based methods have been developed ranging from global comparisons to more local features like active site residues or DNA-ligand-binding motifs.

Here present a new method towards functional annotation of protein structures using the comprehensive TOPOFIT-DB database for structural alignment of proteins. Majority of methods attempt to balance between lower RMSD and higher number of aligned positions. TOPOFIT on the other hand identifies largest groups of residues having the save neighbors in the same spatial location common in both structures which may mathematically be referred to as the topological invariant. TOPOFIT detects a saturation point in the spatial tessellation graph referred to as the topomax point and then the corresponding Cα atoms and corresponding values between the aligned structures are reported. Such an objective methodology provides unambiguous identification and separation of the structurally invariant parts from the variable parts by identifying a precise border between the two. Studying such conserved invariant regions often reveal functionally critical areas of conserved tertiary structure. Identifying conserved residues is useful in assigning function. The PFAT method is also integrated with Catalytic Site Atlas (CSA), a database of catalytic sites for proteins in PDB most of which is manually curated and remaining identified by PSI-BLAST. Presented here PFAT web server combines TOPOFIT-DB with CSA in order to provide a valuable tool for protein function annotation. Our tool will complement the existing repertoire of methods by presenting a different perspective.

PFAT is free to the public and is available at http://topofit.ilyinlab.org/pfat/

N-TOPOFIT-DB, a DNA / RNA 3-D Structural Alignment Database by N-TOPOFIT

Amit Upadhyay[1], Daniil Ilyin[2], Snezhana Abarzhi[2], Peter Clote[1], Valentin Ilyin[1]

[1]Boston College, Biology, Chestnut Hill, MA, 02467, [2]University of Chicago, Division of Physical Sciences, Chicago, IL, 60637

Comparison of RNA 3D structures is crucial for understanding sequence-function relationship at the atomic level. It provides insight into the functional similarity of RNAs having different sequences, highlights structural variations based on sequence differences, and can detect flexibility of overall structure. Structure alignment in general is an NP-complete problem and therefore different heuristics are used to align 3D RNAs in real time. While there are currently several methods presented as web servers for pairwise 3D RNA alignments, the number of RNA 3D structures are growing fast and pre-calculated resources are needed.
We present N-TOPOFIT-DB, a new comprehensive database of RNA 3D alignments based on our N-TOPOFIT method. This method derives from our previously published TOPOFIT method, which aligns protein structures by using a Delaunay-Voronoi tessellation to represent molecular structure and subsequently compare structural patterns. The N-TOPOFIT method is accurate and very efficient. Our protein TOPOFIT-DB has currently almost 400 million alignments. The N-TOPOFIT-DB database, presented here, contains structural alignments between all available RNA and DNA 3D structures available in the PDB, with over 207,000 significant hits. It is updated on regular basis.
Given an input query, the web interface provides a list of structurally related RNAs/DNAs, outputting both a sequence alignment and two-dimensional alignment plot, along with a powerful visualization of superimposition in 3D with the Friend applet. All RNA/DNA 3D structures are collected in groups (families) with a representative centroid, which has been chosen automatically from the N-TOPOFIT alignments by Z-Score > 3 and coverage to both the query and subject at >80%.
The database is MySQL-based with flexible design, easy expandability, and easy retrieval of information. The N-TOPOFIT-DB database is free to the public and is available at http://topofit.ilyinlab.org/ntopo along with a server for pairwise N-TOPOFIT alignment of any RNA/DNA structures by user query.

Structure of Human TRPV4 Ankyrin Repeat Domain

Hitoshi Inada , Rachelle Gaudet

Harvard University, Molecular and Cellular Biology, Cambridge, MA, 02138

TRPV4 is a calcium-permeable cation channel playing important roles in various physiological functions such as osmosensation, mechanosensation, cell barrier formation, and bone formation. Recent studies reported that mutations in TRPV4 are associated with human inherited diseases including neuropathies and skeletal dysplasias due to constitutive activity of the channel. TRPV4 activity is regulated by binding of small molecules such as ATP and calmodulin to the ankyrin repeat domain (ARD) in its N-terminus. We determined structures of ATP-free and -bound forms of human TRPV4-ARD and compared them with available TRPV-ARD structures. The long Finger 3 loop is flexible and may act as a switch to regulate the channel activity. Comparisons of TRPV-ARDs structure also suggest an evolutionary relationship between ARD structure and ATP binding ability. Biochemical analyses of a large panel of TRPV4-ARD mutations associated with human inherited diseases showed that some mutations impaired thermal stability while others reduced ATP binding ability, providing possible molecular mechanisms for the diseases.

Moonlighting Proteins

Constance Jeffery

University of Illinois at Chicago, Biological Sciences, Chicago, MA, 60607

Moonlighting proteins comprise an interesting subset of multifunctional proteins in which the two functions are found in a single polypeptide chain. They do not include proteins that are multifunctional due to gene fusions, families of homologous proteins, splice variants, or promiscuous enzyme activities. Known moonlighting proteins include several different kinds of enzymes and other proteins and different combinations of functions. Recent crystal structures of some moonlighting proteins have provided clues to the molecular mechanisms of one or both functions, and in some cases how a protein can switch between functions. The ability of proteins to "moonlight" might be one reason why the human genome contains fewer protein encoding genes than had been predicted. We are preparing a curated, searchable database that contains information about the sequences, structures, and functions for known moonlighting proteins.

Solution NMR and biophysical analysis of the cataract-associated R76S mutant of human γD-crystallin

Fangling Ji[1], Jinwon Jung[2], Angela M Gronenborn[3]

[1]University of Pittsburgh, Structural Biology, Pittsburgh, PA, 15260, [2]University of Pittsburgh, Structural Biology, Pittsburgh, PA, 15260, [3]University of Pittsburgh, Structural Biology, Pittsburgh, PA, 15260

Although a number of γD-crystallin mutations have been associated with cataract formation, there is no clear understanding of the molecular mechanism(s) that lead to cataracts. As part of our ongoing studies, we investigated the recently discovered Arg76 to Ser (R76S) mutation that is correlated with childhood cataract in an Indian family. We expressed R76S γD-crystallin in E. coli and characterized the protein by CD and fluorescence spectroscopy, as well as determined its thermodynamic stability with respect to thermal and chemical denaturation. Surprisingly, no significant biochemical/biophysical differences were observed between the wild-type protein and the R76S variant. As expected, replacement of the positively charged arginine side chain by the neutral serine lowered the pI with an experimentally determined value of 6.8 compared to the wild-type value of 7.3. We also characterized the solution structure of R76S γD-crystallin by NMR. Using residual dipolar couplings (RDCs) we could show that the overall structure of the mutant is very similar to that of the wild type protein. Likewise, the dynamics behavior is also unaffected by the mutation and ^{15}N relaxation data reveal the same perturbation in R_2 values around His22 that were observed in wild type γD-crystallin. In kinetic unfolding/refolding experiments, the R76S mutant again exhibited essentially the same behavior as the wild type protein, including a similar degree of off-pathway aggregation suppression imparted by αB-crystallin. Overall, our results suggest that neither structural nor stability changes in the protein are responsible for the R76S variants association with cataract. Further studies will be necessary to evaluate the functional relevance of the R76S γD-crystallin mutation.

Discovery of novel inhibitors of therapeutic targets using fragment-based lead discovery methods

Holly Jing, John Badger, Pierre Bounaud, Barbara Chie-Leon, Cheyenne Logan, Robert Meadows, Vandana Sridhar, Vicki Nienabar

Zenobia Therapeutics, Inc., Structural Biology, La Jolla, CA, 92037

Fragment-based lead discovery (FBLD) is instrumental in the discovery of novel compounds targeting various diseases. FBLD allows efficient screening of chemical fragments upon which chemical groups can be added to further enhance ligand efficiencies and to improve pharmacokinetics and toxicity profiles of the compounds. A successful FBLD screen requires three fundamental components: 1) the therapeutic target in a highly purified form, 2) the fragment library containing high-quality compounds that cover the theoretical chemical space efficiently, and 3) the screening methods that bring together the target and compounds with high ture-positive rates at low costs. Meeting all these criteria could be challenging, and it requires in-depth knowledge of protein biochemistry, chemistry, and biophysics. Zenobia Therapeutics has established strong expertise in FBLD and has provided high-quality fragment libraries to many customers around the world. Using a combination of thermal shift assays, a proprietary nano-calorimetry technology, and X-ray crystallography, we have identified leads with strong potencies to target central nervous system and infectious diseases. A comparison of the state-of-the-art screening technologies, along with data on two therapeutic targets, will be presented on the meeting.

References:

1. Nienaber VL. Fragment-based Drug Discovery for Diseases of the Central Nervous System, in "Library Design, Search Methods and Applications of Fragment based Drug Design" (Editor: Bienstock RJ). 2011 (in press).
2. Nienaber VL, Richardson PL, Klighofer V, Bouska JJ, Giranda VL, and Greer J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. Nat Biotechnol. 2000 Oct; 18(10):1105-8.
3. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. Science. 1996 Nov 29;274(5292):1531-4.

Prevention of amyloid β-peptide fibrillation. Insights from the BRICHOS domain structure.

Hanna Willander[1], Glareh Askarieh[2], Jenny Presto[1], Sara Linse[3], Stefan D Knight[2], Jan Johansson[1]

[1]Karolinska Institutet, Alzheimer Research Centre, Stockholm, 14186, Sweden, [2]Swedish University of Agricultural Sciences, Molecular Biology, Uppsala, 75123, Sweden, [3]Lund University, Biochmistry and Structural Biology, Lund, 22100, Sweden

Amyloid diseases are associated with a specific form of protein misfolding, in which the protein aggregates into oligomers and fibrils of β-sheet structure. More than 27 amyloid diseases are known today, and among them Alzheimer's disease (AD), which is a progressive neurodegenerative disorder. The senile plaques found in brain of AD patients contain amyloid fibrils formed by amyloid β−peptide (Aβ), mainly consisting of 40-42 residues. The BRICHOS domain contains ~100 residues and is found in a diverse set of human membrane proteins associated with degenerative and proliferative disease, including lung fibrosis (surfactant protein C precursor; proSP-C) and familial dementia (Bri2).
Bri2 is expressed in neurons, and mutations in the Bri2 gene are linked to familial British and Danish dementia. These mutations lead to extensions of the proprotein and generation of extended peptides (ABri or ADan) that form amyloid.
The crystal structure of proSP-C-BRICHOS, determined at 2.1 Å, shows a homotrimer with a novel subunit fold composed of a five-stranded β-sheet flanked by α-helices and loops. A homology model for Bri2-BRICHOS was built using the X-ray structure of proSP-C BRICHOS as a template. Mapping sequence differences between the two BRICHOS domains to the structure reveals a number of changes from hydrophobic residues in proSP-C BRICHOS to charged residues in Bri2-BRICHOS.
We have found that the Brichos domains from both proSP-C and Bri2 prevent fibril formation of both Aβ40 and Aβ42, well below stoichiometric amounts and that the inhibition is concentration dependent. Kinetic experiments show a prolonged lag phase for Aβ fibril formation in the presence of BRICHOS proteins and circular dichroism experiments imply that Aβ40 is kept in a mainly unstructured state in the presence of BRICHOS.

References
1. Johansson H., et al., "The Brichos domain of prosurfactant protein C can hold and fold a transmembrane segment", Protein Sci, 2009 Jun; 18:1175-1182.
2. Willander H., et al., " High-resolution structure of an intramolecular chaperone. Implications for anti-amyloid activity of the BRICHOS domain", submitted, 2011.
3. Willander H. et al., "BRICHOS domains efficiently prevent fibrillation of amyloid β-peptide", submitted, 2011.

Supercomutational life science in Japan

Chisa Kamada, Erika Jinnai

RIKEN, HPCI Program for Computational Life Sciences, Kobe, 6500047, Japan

The Japanese Next-Generation Supercomputer–named “K” after the character京, which stands for 10 to the 16th power–is an essential tool for advancing science and technology. The potential K offers for expanding basic research in materials science and the life sciences is clear. The HPCI Strategic Program, K computer project, is a five-year program of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) aimed at fostering significant social and academic breakthroughs in five strategic fields: 1 Computational Life Science and Applications in Drug Discovery and Medical Development, 2. New Materials and Energy Generation, 3. Global change and the mitigation of natural disasters, 4. Industrial innovation, and 5. The origin of matter and the universe. In the field of Computational Life Science, we, RIKEN, have been developing applied computational sciences, collaborating with other universities and research institutes. In addition, we offer a number of technical workshops and seminars to the social and academic communities, and are working to develop human networks in computational life sciences**.**

Crystallization of Proteins in High Magnetic Force Fields Aiming at High Resolution Structures

Tatsuki Kashiwagi[1], Kazutoshi Takahashi[1], Mai Yamamoto[1], Kohki Ishikawa[1], Mototaka Suzuki[1], Keiichi Yokoyama[1], Seiichi Taguchi[2], Ei-ichiro Suzuki[1], Masaru Tanokura[3], Akira Nakamura[3], Jun Ohtsuka[3], Yoriko Sawano[3], Motosuke Kiyohara[4], Takahiro Ode[4], Akiko Kita[5], Noriyuki Hirota[6], Hitoshi Wada[6]

[1]Ajinomoto Co., Inc., Institute for Innovation, Kawasaki, 210-8681, Japan, [2]Hokkaido University, Graduate School of Engineering, Sapporo, 060-8628, Japan, [3]The University of Tokyo, Graduate School of Agricultural and Life Sciences, Tokyo, 113-8657, Japan, [4] Kiyohara Optics Inc, , Tokyo, 160-0022, Japan, [5]Kyoto University, Research Reactor Institute, Sennan, 590-0494, Japan, [6]National Institute for Materials Science, , Tsukuba, 305-0003, Japan

High resolution crystal structures of proteins are indispensable for protein engineering and structure based drug design. However, crystallization is a bottleneck to the determination of protein structures, requiring tremendous trial-and-error experiments. The low gravity environment, as attained in space, is reported to produce crystals of high integrity due to the lack of gravity-induced convection. A similar effect may be possible when protein samples are crystallized in magnetic fields. If this idea works, it will provide us with more opportunities for experiments at lower costs.
We are now constructing a protein crystallization system with high magnetic force fields. The protein samples are crystallized in a superconducting magnet where magnetic force should suppress the convection of diamagnetic protein solutions. We have applied two industrially important enzymes to this system. One is a microbial transglutaminase (MTG)[1] that catalyzes an acyl transfer reaction between the γ-carboxyamide groups of peptide-bound Gln residues and the primary amino groups. The other enzyme, phosphoketolase[2] is a thiamine diphosphate-dependent enzyme in sugar metabolism. The results of the crystallization in high magnetic force fields of these enzymes will be presented. In the case of MTG, we will demonstrate the improvement of thermostability obtained by rational mutagenesis.
The final goal of our system is *in situ* observation of crystal growth under many conditions. We will also report current developments of the optical probe and crystallization cell of the system.
This work is supported by “Development of Systems and Technology for Advanced Measurement and Analysis Fund” of the Japan Science and Technology Agency.
[1] T. Kashiwagi *et al.*, J. Biol. Chem., 277, 44252-44260 (2002).
[2] K. Takahashi, *et al.*, FEBS Lett., 584, 3855-3861 (2010).

Chemical similarity-based docking to receptor proteins by searching maximum common substructure

Takeshi Kawabata, Haruki Nakamura

Institute for Protein Research, Osaka University, Suita, Osaka, 565-0871, Japan

A binding conformation of a ligand to a receptor protein often tells us the essential interaction between the ligand and the protein, it can be used for screening another bioactive molecule in the chemical library. Methods to predict binding conformations are classified into two approaches: docking based on receptor protein structure and docking based on binding ligand structure (chemical similarity-based docking). Recently, the second approach becomes feasible, because an increasing number of 3D structural data of small compound-protein complexes are accumulated in the PDB database.

We recently developed the program *kcombu* for finding one-to-one atomic correspondence between given two chemical structures [1]. Our program searches maximum common substructure (MCS) using the build-up heuristic algorithm. It provides classical connected and disconnected MCSs, as well as topologically constrained disconnected MCS (TD-MCS) allowing a few gaps in connected substructures. If we know the 3D structure of one chemical compound (template) on the receptor protein, that of another similar chemical compound (target) can be predicted using the template 3D structure and an atomic correspondence between the template and target compounds.

In addition to the MCS search engine in the *kcombu* program, we implemented a 3D modeling procedure by superimposing and deforming the 3D structure of the target molecule onto that of the template molecule. We employed three methods following Marialke et al. [2]. The first method is a rigid-body superimposing of corresponding atom pairs using the well-known RMSD algorithm. The second method is a “stamping” the dihedral angles of the template molecule into those of the target molecule. The third method is a gradient-based optimization with rotatable bond dihedral angle. A difference of positions of corresponding atom pairs is employed as the objective function. We evaluated the prediction performance of our method using 154 binding ligand 3D structures on CDK2 proteins registered in PDB. The result shows that the combination of the three methods yielded the best prediction, and the correspondence of TD-MCS was better than C-MCS, especially for weakly similar compound pairs. Roughly speaking, if the template and target compounds have more than 60 % *tanimoto* similarity, 3D conformations of 70 % of the target compounds can be predicted correctly (rmsd<2.0 angstrom). This method will be available through the KCOMBU web server (http://strcomp.protein.osaka-u.ac.jp/kcombu/).

[1] Kawabata,T. Build-up algorithm for atomic correspondence between chemical structures. *J. Chem. Inf. Model.* **2011**, in press.

[2] Marialke, J., Tietze, S., Apostolakis, J. Similarity based docking. *J. Chem. Inf. Model.* **2008**, 48, 186-196.

Structural and biochemical studies of the cGMP selectivity for cGMP dependent protein kinase.

Choel Kim[1,2], Jeong Joo Kim[1], Dar-Chone Chow[2], Gilbert Huang[1], Peter Zwart[3]

[1]Baylor College of Medicine, Pharmacology, Houston, TX, 77030, [2]Baylor College of Medicine, Biochemistry and Molecular Biology, Houston, TX, 77030, [3]Lawrence Berkeley National Laboratory, The Berkeley Center for Structural Biology, Berkeley, CA, 94720

Background:
Cyclic GMP is a crucial second messenger that translates extracellular signals into a variety of cellular responses. As a central mediator of the Nitric Oxide-cGMP signaling cascade, which regulates vascular tone, platelet aggregation, nociception and hipocampal/cerebellar learning, cGMP-dependent protein kinase (PKG) mediates most of the effects of cGMP elevating drugs, such as nitric oxide-releasing agents and phosphodiesterase inhibitors. Thus PKGs represents an important drug target for treating hypertensive diseases and erectile dysfunction.
The fidelity of the NO-cGMP signaling pathway is largely dependent on PKG's ability to selectively bind cGMP over cAMP. Although both cGMP and cAMP bind and activate PKG, cGMP preferentially activates PKG 60-100 fold better than cAMP; yet, little is known about the molecular features required for the cGMP selectivity of PKG. We have investigated the mechanism of cyclic nucleotide binding to PKG by determining crystal structures of the amino-terminal cyclic nucleotide-binding domain (CNBD-A) of human PKG Iβ in the absence and presence and of cGMP or cAMP and by measuring binding affinities of CNBD-A and CNBD-B to both cGMP and cAMP.

Results:
The crystal structures of CNBD-A with bound cAMP or cGMP reveal that cAMP binds in either syn or anti configurations whereas cGMP binds only in a syn configuration, with a conserved threonine residue anchoring both cyclic phosphate and guanine moieties. The structure of CNBD-A in the absence of bound cyclic nucleotide was similar to that of the cyclic nucleotide bound structures. Surprisingly, isothermal titration calorimetry experiments demonstrated that, while CNBD-A binds both cGMP and cAMP with a high affinity and shows only a two-fold preference for cGMP, CNBD-B binds cGMP with a low affinity, but shows a two hundred-fold preference consistent with previous data.

Conclusions:
Our findings suggest that CNBD-A binds cGMP in the syn conformation through its interaction with Thr193 and an unusual cis-peptide forming residues Leu172 and Cys173. These studies provide the first structural insights into cyclic nucleotide binding to PKG and pave for PKG specific activator. Unexpectedly, our ITC results showed that CNBD-B with a relatively low affinity, not CNBD-A with a high affinity to both cyclic nucleotides, provides the cGMP selectivity. We hypothesize that this selectivity is mediated by a conserved acidic residues with the C-helix of CNBD-B interacting specifically with the guanine moiety of cGMP.

Integrative Structure Determination of the Components of the Nuclear Pore Complex by X-ray Crystallography, Small Angle X-Ray Scattering, Electron Microscopy, NMR and Comparative Modeling

Seung Joong Kim[1], Parthasarathy Sampathkumar[3], Javier Fernandez-Martinez[4], Jeremy Phillips[1], Dina Schneidman[1], Tsutomu Matsui[2], Hiro Tsuruta[2], Michael Sauder[3], Stephen K Burley[3], Michael Rout[4], Andrej Sali[1]

[1]University of California, San Francisco, Department of Bioengineering and Therapeutic Sciences, San Francisco, CA, 94158, [2]SLAC National Accelerator Laboratory, Stanford Synchrotron Radiation Lightsource, Menlo Park, CA, 94025, [3]Eli Lilly and Company, New York SGX Research Center for Structural Genomics, San Diego, CA, 92121, [4] The Rockefeller University, Laboratory of Cellular and Structural Biology, New York, NY, 10065

The Nuclear Pore Complex (NPC, ~50 MDa) is the sole passageway for the transport of macromolecules across the nuclear envelope. The NPC plays a key role in numerous critical cellular processes such as transcription, and many of its components are implicated in human diseases such as cancer. Previous work (ref 1, 2) defined the relative positions of its 456 constituent proteins (nucleoporin or Nups), based on spatial restraints derived from biophysical, electron microscopy, and proteomic data. Further elucidation of the evolutionary origin, transport mechanism, and assembly of the NPC will require higher resolution information. As part of an effort to improve upon the resolution and accuracy of the NPC structure, we set out to determine the atomic structures of the NPC components. Because it proved difficult to determine the atomic structures of whole Nups by X-ray crystallography alone, we are relying on multiple datasets that are combined computationally by our Integrative Modeling Platform (IMP) package (http://salilab.org/imp). In particular, we developed an integrative modeling approach that benefits from crystallographic structures of fragments of the protein or its homologs, Solution Small Angle X-ray Scattering (SAXS) profiles of the protein and its fragments (ref 3), NMR, and negative stain Electron Microscopy (EM) micrographs of the protein. Each dataset is converted into a set of spatial restraints on the protein structure, followed by finding a model that satisfies the restraints as well as possible using a Monte Carlo / molecular dynamics optimization procedure. The approach will be illustrated by its application to yeast Nup133.

1. F. Alber et al., “Determining the architectures of macromolecular assemblies”, Nature 450, 683-694 (2007).
2. F. Alber et al. “The Molecular Architecture of the Nuclear Pore Complex”, Nature 450, 695-701 (2007).
3. F. Förster et al. "Integration of Small-Angle X-Ray Scattering Data into Structural Modeling of Proteins and Their Assemblies". J Mol Biol 382 (4), 1089-1106 (2008).

Unexpected reactions resulting from mutating catalytic residues in an amidase reveal the role of the catalytic unit

Serah Kimani, Trevor Sewell

University of Cape Town, Electron Microscope Unit, Cape Town, 7700, South Africa

Nitrilase superfamily amidases catalyze the conversion of amides to their corresponding acids and ammonia; they perform essential metabolic roles and are useful in the manufacture of fine chemicals and pharmaceuticals. A conserved Cys, Glu and Lys (CEK) catalytic triad has been identified in these enzymes, and catalysis proposed to follow a 'ping pong bi bi' mechanism in which the substrate forms a thioester with the catalytic cysteine. Details of each step in the pathway remain to be clarified. A 'second', structurally conserved active site glutamate has been found to be essential for activity in amidases, but its role in catalysis has never been elucidated. This study was aimed at understanding the catalytic mechanism in amidases, particularly the role of the catalytic residues in the reaction process.

The four catalytic residues (C165, E61, E139 and K131) were mutated individually in the model amidase from Nesterenkonia species (NitN). The mutants were reacted with a range of short aliphatic amide substrates and the resultant proteins characterized by mass spectrometry and X-ray crystallography. Mutation of the catalytic cysteine allowed visualization of non-covalently bound substrates in the active site.

Mutation of the two catalytic glutamates and the lysine resulted in unstable mutants that were inactive, highlighting the importance of these residues in maintaining the protein fold and for enzymatic activity. Mass spectrometry showed that thioester acyl-enzyme intermediates could be trapped by mutating the 'first' glutamate (E61). In addition, incubation of the glutamate mutants with amide substrates resulted in two types of unexpected reactions: An SN2 substitution reaction leading to the displacement of the fluorine in fluoroacetamide was observed with the E61Q/L mutants, while Michael addition of acrylamide at the catalytic cysteine was observed with the mutants of both glutamate residues (E61Q/L and E139Q). These artifactual reactions not only emphasize the role that the two glutamates play in positioning the substrates but also demonstrate that the 'first' glutamate is not necessary to catalyze nucleophilic attack by the cysteine. Mutation of K131 to glutamine resulted in the formation of a covalently modified expressed protein that appears to have an adipamide thioester intermediate trapped in the active site, once again emphasizing the need for the intact catalytic assembly to catalyze the hydrolysis of the acyl intermediate.

These results support a Cys, Glu, Glu and Lys (CEEK) catalytic tetrad in which the two glutamates and the lysine function as a modular assembly.

Structural basis for peptide ligand recognition by LR11 Vps10p domain

Yu Kitago[1], Zenzaburo Nakata[1], Masamichi Nagae[1], Terukazu Nogi[1,2], Junichi Takagi[1]

[1]Osaka University, Institute of Protein Research, Suita, 565-0871, Japan, [2]Yokohama City University, Department of Nanosystem Science, Yokohama, 230-0045, Japan

LR11 (LDLR relative with 11 binding repeats), a member of the LDL receptor superfamily, is expressed abundantly in the human nervous system and implicated in Alzheimer's disease. Uniquely among LDLR family members, LR11 possesses a Vps10p (Vacuolar protein-sorting 10 protein) domain at the N-terminal of its ectodomain. In order to shed lights on the biological roles of LR11 in the intracellular protein sorting, the molecular mechanism of ligand recognition by LR11 Vps10p domain was analyzed by X-ray crystallography.
Recombinant human LR11 Vps10p domain (residues 1-753) with C-terminal His tag was produced in mammalian cells, purified, deglycosylated, and crystallized. From the crystal, X-ray diffraction data set was collected to 2.4Å resolution and the initial model was built from a molecular replacement solution using the structure of Vps10p domain of sortilin (PDB ID: 3F6K). The refined structure revealed that the major portion of the LR11 Vps10p domain is comprised of a ten-bladed β-propeller fold with a wide-open pore at the center. Previous studies reported that LR11 Vps10p domain is able to bind its own propeptide. We confirmed this binding using a fluorescence polarization assay with labeled propeptide. Finally, structure determination of the same domain in complex with propeptide ligand revealed the unique mode of peptide recognition by LR11 Vps10p domain, where the inside wall of the pore lined by the edge strand of the propeller blade provide binding site.

Structural basis for control of spider silk assembly: a conserved N-terminal solubility relay.

Glareh Askarieh*[1], My Hedhammar*[2], Michael Landreh[3], Cristina Casals[4], Jan Johansson[2], Stefan D Knight[1]

[1]Swedish University of Agricultural Sciences, Department of Molecular Biology, Uppsala, 75124, Sweden, [2]Swedish University of Agricultural Sciences, Department of Anatomy, Physiology and Biochemistry, Uppsala, 75124, Sweden, [3]Karolinska Intisitutet, Department of Medical Biochemistry and Biophysics, Stockholm, 17177, Sweden, [4] Complutense University of Madrid, Department of Biochemistry and Molecular Biology I, Madrid, 28040, Spain

Spider dragline silk is the strongest natural fiber known, with a unique combination of extensibility and tensile strength giving it a toughness never attained in synthetic fibers. Spider silk is produced in the spider's spinning apparatus by extrusion of a soluble high-concentration "dope" of spidroin proteins. Spidroins are large proteins (~4000 amino acid residues) consisting of repetitive poly Ala/Gly-rich segments flanked by conserved non-repetitive N- (~130 residues) (NT) and C-terminal (~110 residues) (CT) domains. We solved the X-ray structure of the NT domain of major ampullate spidroin 1 from the dragline of the nursery web spider *Euprosthenops australis*. The 1.7 Å structure reveals a homodimer of dipolar, antiparallel five-helix bundle subunits. Together with data showing pH dependent and salt-inhibitable aggregation of NT and of minispidroins containing NT, the structure suggests that the distinct charge distribution of NT helps regulate silk assembly by preventing premature aggregation at higher pH values and triggering polymerization as the pH is lowered along the spider's silk extrusion duct. Recent SAXS experiments suggest that an NT dimer very similar to that observed in our crystal structure exists in solution at pH <6.4. However, neither the dimer nor a monomer derived from the dimer structure, nor any mix of these, can fit data collected at higher pH values. The SAXS data further suggest the occurrence of a continued progressive stabilisation of the NT structure within the studied pH range. The existence of a less stable dynamic structure at higher pH values is strengthened by the observation of significant peak broadening in wt NT NMR spectra, most notably for residues in the dimer interface (unpublished), and HDX-MS data. The gradual change in SAXS signal with pH is significantly affected in the NTE79Q and NTD40NE84Q mutants, suggesting that these mutated (conserved) residues are critical for the structural transition.

References:
Askarieh *et al*. "Self-assembly of spider silk proteins is controlled by a pH-sensitive relay", *Nature* 465:236-238, 2010.
Landreh *et al*. "A pH-dependent Dimer Lock in Spider Silk Protein", *J Mol Biol* 404:326-336, 2010.

Improving EST annotation with protein structure prediction software

Adrian Laurenzi, Ram Samudrala

University of Washington, Department of Microbiology, Seattle, WA, 98109

Large public databases of Expressed Sequence Tags (ESTs) have been made available through projects such as the Gene Index Project at Harvard University by sequencing the expressed genes from a tremendous variety of organisms. There are over 60 million ESTs in GenBank representing well over half of all GenBank entries. To make efficient use of EST data computational techniques have been developed to analyze and organize EST databases. ESTs have been useful in discovering new genes, understanding gene expression and regulation, and constructing genome maps, all of which have important implications in medicine. However, the utility of EST data relies upon our ability to make accurate annotations that describe the functionality of the ESTs in the source organism. Presently most approaches used to annotate ESTs rely on sequence-based comparison methods such as BLAST. This is limiting because the function of a protein is dependent upon its tertiary (3-D) structure. Therefore, the ability to reliably predict the function of an EST could be improved if we were able to accurately predict the structure of the protein encoded by the EST. Because EST sequences often do not contain the entire coding region of the proteins they encode, we evaluated the quality of structures predicted by top-performing protein structure prediction software when given partial length sequences as input. Our results demonstrate that I-TASSER, Rosetta, and Protinfo can accurately predict the partial structures of proteins encoded by sequences that contain approximately 50% or more of the full-length protein sequence. This suggests that these methods would be useful in annotating at least a subset of sequences from EST databases. Application of protein structure prediction to EST data could enhance our ability to accurately predict EST function and enable researchers to design more informed bench experiments and expedite the discovery of new genes and gene networks with potential utility in medicine.

The *GAL* regulon in *S. CEREVISIAE*: the Gal3p/Gal80p interaction

Tali Lavy, Rajesh K Prakash, Hongzhen He, Leemor Joshua-Tor

Howard Hughes Medical Institute and Cold Spring Harbor Laboratory, Structural Biology, Cold Spring Harbor, NY, 11724

S. cerevisiae responds to galactose by activating the coordinated transcription of a family of genes known as the *GAL* genes. This transcription is regulated by the *GAL* regulatory switch, which consists of three key protein components: Gal4p, the transcriptional activator; Gal80p, the repressor; and Gal3p, the transducer. The interplay between Gal3p, Gal80p and Gal4p determines the on/off status of the switch. Gal4p binds the upstream activating sequences of *GAL* genes (UASgal) through its N-terminal domain and activates transcription through interaction of its C-terminal transcription activation domain (TAD) with the transcription machinery. During non-inducing conditions, Gal80p binds to the TAD of Gal4p to repress transcription. Repression is relieved upon induction with galactose, mediated by the transducer, Gal3p. We have shown previously that NADP also plays a role in activating the switch [1]. Gal3p forms a complex with Gal80p, α-D-galactose and ATP.
We have solved the crystal structure of the tertiary complex of Gal80p-Gal3p with α-D-galactose and ATP to 2.1Å resolution. The structure shows that the interaction between Gal3p and Gal80p occurs only when Gal3p is in a "closed conformation" in which the N- and C- terminal domains of the protein are closer together. This conformation is induced by α-D-galactose and ATP binding. The Gal3p interaction surface of Gal80p is different from the surface for Gal4p TAD interaction and the structure reveals the basis for super-repressor phenotypes of Gal80p. Gal3p is 75% identical in sequence to Gal1p, but lacks the galactokinase activity of the Gal1p enzyme. We show that the α-D-galactose conformation found in our structure is unique compared to what is found in other galactokinase structures and might be the basis for the lack of galactokinase activity in Gal3p.

[1] NADP Regulates the Yeast GAL Induction System. Kumar, Rajesh P., et al. 2008, Science, Vol. 319, pp. 1090-1092.

Silver (I) mellatated protein structures derived by combined NMR and X-ray crystallographic analyses.

Matthew J Panzner, Stephanie M Bilinovich, Christopher J Ziegler, Wiley J Youngs, Thomas C Leeper

University of Akron, Chemistry, Akron, OH, 44325-3601

In spite of the prominent role that diverse metal ions play in protein function, the majority of metallated protein structures in the pdb seem dominated by zinc, copper, and nickel. Furthermore, there is a particular paucity of protein structures in complex with biotoxic metals. This has lead to extensive confusion regarding the direct mechanism of toxicity by certain metal ions, e.g. silver, that exert their effects at concentrations too low to be explained by non-specific interactions. For example, our hypothesis is that silver (I) exerts its unique antimicrobial properties by inactivating specific protein components. However, our understanding of the coordination geometry and site specificity for silver(I) is quite limited. In this report we present the combined NMR and crystallographic analyses of three proteins that are demonstrated to bind to silver ions at unique sites: Lysozyme, Azurin, and Glutaredoxin. Each reveals different coordination geometries and ligating residues, consistent with the notion that silver is quite plastic in its coordination propensities. This diversity of geometries also suggests that a significantly larger number of structural examples are required before clear predictions about silver (I) site selection can be made. Our suggestions for structural targets particularly important in eukaryotic silver ion processing will also be presented.

A Maximum-Likelihood Approach to Local Real-space Model Fitting

Yang Lei, Ramgopal R Mettu

University of Massachusetts Amherst, Electrical and Computer Engineering, Amherst, MA, 01003

X-ray crystallography is perhaps the most commonly used method for protein structure determination in structural biology. To interpret the data, typically both the model and observed data are iteratively refined until a statistical criterion, such as R-free or the R-factor, is met. So-called ``reciprocal-space refinement'' attempts to minimize the difference between the model and observed structure factors, while ``real-space refinement'' attempts to minimize the difference between the model and observed electron density maps. Real-space fitting and/or refinement are particularly useful when selected parts of a protein structure must be fit into the electron density map. This is a basic component of nearly every modeling algorithm for X-ray crystallography, and is also useful in tasks such as sidechain placement or refinement of a homologous backbone. In practice, local real-space model fitting is usually alternated with global reciprocal-space model refinement.

In the case of reliable experimental phasing, we show how our maximum likelihood approach can potentially provide useful improvements for local real-space model fitting. Specifically, our contributions are as follows:

- Analysis of the error distribution in measurements made with isomorphous replacement and anomalous diffraction.

- Error propagation analysis from structure factor to electron density domain under the above experimental methods for phasing.

- Elimination of the redundancy (i.e. algebraic mutual dependence) among the electron densities which results from zero-padding when utilizing the discrete Fourier transform.

- An analytic representation of the real-space likelihood function which considers the contributions from the experimental data and the model.

- A confidence measure for evaluating the highest-scoring conformations that is derived from analyzing the likelihood statistics of the conformational choices.

To show the potential of our results, we study the problem of sidechain placement using experimental electron density maps of varying resolutions. We show that our approach indeed attains high accuracy in most cases. Moreover we show that our confidence measure can be used to detect and correct inaccurate predictions.

In vitro reconstituted F-Actin and Fascin bundles studied by Cryoelectron Tomography and Subtomogram Averaging

Alejandra Leo-Macias[1], Ruben Diaz-Avalos[3], Xin-Yun Huang[2], David Stokes[1,3]

[1]New York University, Skirball Institute of Biomolecular Medicine, New York, NY, 10016, [2]New York Structural Biology Center, Electron Microscopy, New York, NY, 10027, [3]Cornell University Weil Medical College, Physiology and Biophysics, New York, NY, 10065

Bundles of filamentous actin (F-actin), form primary structural components of a broad range of cytoskeletal processes including filopodia, sensory hair cell bristles and microvilli. Actin-binding proteins (ABPs) allow the cell to tailor the dimensions and mechanical properties of the bundles to suit specific biological functions. Fascin is a 55 kDa globular protein that organizes F-actin into well-ordered, tightly packed parallel bundles in vitro and in vivo inside the cells and plays a key role in the organization of two major forms of actin-based structures: cortical cell protrusions that mediate cell interactions and migration, and cytoplasmic microfilament bundles that contribute to cell architecture and to intracellular movements. In this work, we reconstitute unconstrained 3D actin/fascin bundles in vitro and analyze them by cryoelectron tomography. We characterize the bundles at the "macroscopic" level (number of filaments per bundle, length of filaments, interfilament spacing and crossbridge distances) and also perform an analysis to better characterize the actin/fascin interaction at the molecular level. Two different models have been proposed for that interaction whose validity we want to address here: the classical two-fold model, where fascin interacts with two actin filaments through two different binding sites, and a novel model consisting on a three-fold interaction, recently proposed based on a systematic mutagenesis experiment of all surface exposed residues on fascin, where critical residues for the bundle formation are located in three distinct regions of this protein. By subtomogram averaging, we have already obtained a preliminar average electronic density of the actin/fascin interaction, although its resolution is still far from allowing unambiguous fitting of the atomic structures. By increasing the resolution of our average map we expect to be able to perform a fitting reliable enough to help to shed some light on the way the interaction is actually taking place.

Crystal Structure and Functional Studies of the GerBC Component of a *Bacillus subtilis* Spore Germination Receptor

Yunfeng Li, Barbara Setlow, Parvathimadhavi Catta, Kerry-Ann V Stewart, Peter Setlow, Bing Hao

University of Connecticut Health Center, Department of Molecular, Microbial and Structural Biology, Farmington, CT, 06030

GerBC is the C component of GerB nutrient germinant receptors (GRs) that play a critical role in triggering the germination of dormant spores of *Bacillus* species in response to nutrient mixture of L-asparagine, D-glucose, D-fructose and K^+ ions (AGFK). The crystal structure of GerBC reveals an unusual interlocked dimer configuration in which each of the three distinct domains in the monomer possesses a unique fold. Secondary-structure prediction and structure-based sequence alignment suggest that the GerBC structure represents the prototype for C subunits of GRs from spores of all *Bacillales* and *Clostridiales* species and defines two highly conserved structure regions in this family of proteins. Single and multiple alanine substitutions were made in the residues in two conserved regions as well as the linker region connecting Domain I and II. The functional assays and kinetic analyses have shown that substitutions in these regions have profound effects on germination via the GerB GR and maintenance of normal GerBC levels in spores, and also have effects on L-valine germination via the GerA GR. Our findings provide the first structural view of the GR subunits and a molecular framework for understanding the architecture, conservation, and function of GRs.

Towards a Structure for the Rubella virus capsid protein

Vidya Mangala Prasad, Siyang Sun, Michael G Rossmann

Purdue University, Biological Sciences, West Lafayette, IN, 47907

Rubella virus (family *Togaviridae*) is a human pathogen which causes the Rubella disease. It has serious implications when infection takes place during pregnancy and can cause congenital rubella syndrome which results in birth defects. The rubella virus has a diameter of ~85nm and consists of a single stranded 40S RNA of positive polarity. It is known to encode 2 non-structural proteins (p150 and p90) and 3 structural proteins, namely, the capsid protein, the E1 and E2 glycoproteins. The nucleocapsid capsomere consists of two disulphide linked dimer of the capsid protein and its main function is to package the viral genome into nucleocapsids. Recently, studies have reported additional non-structural functions for the capsid protein which involves virus-host interactions and also cellular functions like inhibiting mitochondrial import which have not been reported earlier for any other virus encoded protein. This makes the capsid protein an interesting and important candidate for investigation. There have been some studies establishing the expression of the rubella capsid protein in different cell lines, with our lab pursuing expression and purification of the protein for structural purposes. To-date, a truncated form of the capsid protein has been purified and crystals of the capsid protein have been obtained. X-ray diffraction data to ~3Å has also been collected. Molecular replacement attempts with the related Sindbis virus capsid structure have not been successful. Co-crystallization with various heavy atoms and Se-Met are being carried out to get phase information using Multiple Anomalous Dispersion (MAD) technique. The capsid protein structure is still unknown and determining its crystal structure can provide useful insights to the function of the protein and also to the structure of the virus.

Computational Design of Conformationally Intact Immunogen for the Elicitation of Protective Antibody Immune Response

Manish Manish[1,3], Ram Samudrala[1], James I Mullins[1], Arturo Centurion[2], Rakesh Bhatnagar[3], Manuel Mendoza[4], Mirko Zimic[4]

[1]University of Washington, Department of Microbiology, Seattle, WA, 98109, [2]University of Washington, Harborview Medical Center, Seattle, WA, 98104, [3]Jawaharlal Nehru University, School of Biotechnology, New Delhi, 110067, India, [4] Universidad Peruana Cayetano Heredia, Department of Biochemistry, Lima, 430, Peru

Epitopes are the immunologically active region of antigens, and are being actively evaluated for the development of successful vaccination strategies. Antibodies interact with epitopes due to the non-covalent forces, which depend on 3- dimensional spatial arrangement of atoms in the interacting residues. During infection, antibody recognizes the structural element of protective epitope in the context of native antigen molecule. Hence, we aim to design conformationally intact immunogen for the elicitation of protective antibody immune response.

We aim to graft the most relevant epitopic regions in a soluble/stable carrier protein so that the resulting chimeric designed protein retains the original native conformation of the epitopes and presents it exposed to aqueous environment, and thus readily available to the immune system in a biologically relevant manner. We employ our computational protein structure prediction design strategies to achieve this objective.

The first stage of our strategy involves determining the conserved regions of surface proteins in the diseases of interest. For the next design stage, we employ our successful torsional angle based protein structure prediction simulation protocol to insert these epitopes sequences as "loops" into our carrier proteins at various positions, perform folding simulations, and examine the most stable configurations according to our knowledge based probability discriminatory functions. This stage of our method helps us to inform how and where these epitopes can be grafted onto the carrier. In the next stage, we have evaluated these structures for their similarity to the original antigen native structure using interatomic contact distances. Structures with these contacts are then further minimised using a larger set of contacts as a pseudoenergy function, to drive simulations where chimeric constructs with optimised antigen-like contacts are produced.

This design algorithm is able to retain the native 3 dimensional structure of epitope which cannot be achieved using conventionally synthesised linear peptide. Excluding the immune decoy elements and presenting these epitopes exposed to environment enable us to focus on eliciting protective antibody response. Hence, our approach can transform an immunogenic region (i.e., region which can induce antibody) to an antigenic region (i.e., region which can bind with antibody).

Extreme evolutionary pressure determines Cysteine residues distribution in protein molecular surfaces.

Stefano M Marino, Vadim N Gladyshev

Harvard Medical School, Medicine, Boston, MA, 02115

Cysteine (Cys) is an enigmatic amino acid. Although one of the least abundant, it often occurs in functional sites of proteins. Whereas free Cys is overall slightly polar, Cys in proteins is often buried and its classification on hydrophobicity scales is ambiguous. We hypothesized that deviation of Cys residues from the properties of free amino acid is due to their reactivity and addressed this possibility by combining extensive structure based computational analyses with (Blast based) evolutionary information. In this work, we examined 20,000 non redundant structures (from the PDB repository), and 200,000 homology models (built with Modeller). Compared to all other amino acids, Cys was characterized by the most extreme conservation pattern, with the majority of Cys being either highly conserved or highly degenerated. In addition, clustering of Cys with other Cys residue (α-carbon within 8 Å) was associated with high conservation, whereas exposure with low conservation. We found that Cys clustering was, by large, the strongest single composition-based feature influencing amino acid conservation in proteins. On the other side of the spectrum, isolated Cys were on average poorly conserved, significantly more so when exposed to molecular surfaces: indeed, these residues showed the highest degree of degeneration, among all amino acid types.
We investigated many structure-based descriptors related to chemical and physical properties of exposed and isolated Cys (e.g. pKa distribution, deviation from standard Henderson–Hasselbalch, effect of local geometries on force field parameters, etc); based on these and on previous results, we propose that the anomalous hydrophobic-like behavior and conservation pattern of Cys can be explained by elimination of isolated Cys from protein surface, not on the basis of their polarity, but due to the implication of having a dangerous surplus of highly reactivity free Cys thiols in protein surfaces. In this scenario, spatial clustering of Cys can serve a protective function, as it can prevent (or rescue, e.g. resolving Cys for oxidized catalytic Cys) unwanted oxidations. The findings presented here support the conclusion that Cys abundance is governed by Cys function in protein (rather than its sheer chemico-physical properties) and uncover fundamental features of this amino acid that limit its abundance on protein surface, as well as providing a rational for the high occurrence of Cys clusters in proteins.

Domain Swapping and Protein Symmetry: Investigations on the Stationary Phase Survival Protein SurE as a model

Yamuna Kalyani Mathiharan, Murthy M.R.N.

Molecular Biophysics Unit, Indian Institute of Science, Bangalore, 560012, India

Domain Swapping is a unique feature found in some proteins wherein a dimer or a higher oligomer is formed by the exchange of identical structural segments between protomers. Domain swapping is thought to have played a key role in the evolution of stable oligomeric proteins. We have addressed the role of domain swapping in the evolution of oligomeric structures and its importance for the stability and function of the oligomeric forms using *Salmonella typhimurium* stationary phase survival protein SurE (*St*SurE) as a model. *St*SurE is coded by a stress survival operon which is expressed during the stationary phase and when the bacteria are under other environmental stresses. The ubiquitous presence of SurE homologs in microorganisms is suggestive of their ancient origin. It is believed that these proteins are important for the survival of bacteria in hostile environments. *St*SurE appears to be a phosphatase specific to nucleoside monophosphates.

Structure determination of *St*SurE has revealed that it is a dimer with C- terminal helices of protomers swapped. The domain swapping observed in *St*SurE is also conserved among three other homologs. Structure of *St*SurE suggested that the hydrogen bonding between OD2 of D230 and NE2 of H234 may be a key feature of domain swapping. This interaction appears to impart rigidity to the hinge region connecting the swapped helices and the rest of the monomer. The resulting rigidity prevents the swapped helices from turning back and becoming a part of the same monomer. In order to examine the role of domain swapping in protein structure, symmetry, stability and function, single site H234A and double site D230A/H234A mutants were constructed and their three-dimensional X-ray structures were determined. In both H234A and D230A/H234A mutants, there is a drastic alteration in the dimeric structure. The H234A is a mixture of both swapped and unswapped forms, while the D230A/H234A is a predominantly unswapped form. Intriguingly, unlike the wild type *St*SurE, oligomers of the mutant proteins, although dimeric, were not symmetric. They were also enzymatically inactive. Domain swapping in *St*SurE thus appears to be important for the formation of a stable, symmetric quaternary organization as well as for the function of the protein. Similar mechanisms might underlay the evolution of oligomeric forms in other proteins as well.

A searchable set of predictions of protein functions with associated probabilities as derived from annotations of protein sequences and three-dimensional structures

Eichin S Julfayev[1], Ryan J McLaughlin[1], Yi-Ping Tao[2], William A McLaughlin[1]

[1]The Commonwealth Medical College, Basic Science, Scranton, PA, 18509, [2]Rutgers The State University of New Jersey, Chemistry & Chemical Biology, Piscataway, NJ, 08854

Annotations of protein structures from the community of protein scientists are presented in the Structural Biology Knowledgebase (SBKB). Annotation arrays of the protein structures in the SBKB and open biological databases were used to train classification models to discern between protein structures with and without a given Gene Ontology annotation. Patterns associated with the given GO annotation were derived, and recognition of these patterns within proteins of under-characterized function has garnered a means to predict novel assignments of protein function. An online searchable list of predictions of protein functions together with their associated assignment probabilities is presented. The aim is to support experimental characterization of the proteins, as the predictions may be used to form new hypotheses that may lead to new discoveries of protein function. As the community uses the predictions, it will help to complete the protein characterization cycle, make the cycle more efficient, and make the computational predictions become more accurate. Current predictions for protein structures of unknown function are presented as example cases of novel annotation assignments. The retrieval of known protein annotations from open sources is also described, and input is requested regarding user-defined functions.

Unfolded Proteins Can Form Long Range Contacts Under Strongly Denaturing Conditions in the Absence of Significant Secondary Structure

Wenli Meng[1], Nichloas Lyle[2], Rohit V Pappu[2], Daniel P Raleigh[1,3]

[1]State University of New York at Stony Brook, Department of Chemistry, Stony Brook, NY, 11794-3400, [2]Washington University, Department of Biomedical Engineering, St. Louis, MO, 63130, [3]State University of New York at Stony Brook, Graduate Program in Biochemistry and Structural Biology, and Graduate Program in Biophysics, Stony Brook, NY, 11794

The properties of unfolded proteins have come under increased scrutiny owing to the realization that unfolded states can contain residual structure which influences protein stability, folding and aggregation. The extent of interactions in the denatured state ensemble (DSE) involving residues distant in primary sequence is unclear, particularly under denaturing conditions. Global measures, such as the radii of hydration (R_h) or gyration (R_g), are generally consistent with expanded DSEs under denaturing conditions; however such experiments may be insensitive to transiently formed long range contacts since they reflect ensemble averages of global properties. Using the N-terminal domain of the ribosomal protein L9 (NTL9), we demonstrate that significant, long range contacts can form in the DSE under strongly denaturing conditions in the absence of significant secondary structure. NMR analysis shows that the urea induced DSE of NTL9 has a very low level of secondary structure and is expanded, however paramagnetic relaxation enhancement (PRE) experiments reveal extensive interactions involving residues which are distant in primary sequence. DSE PRE's are typically benchmarked against a Gaussian chain model, but we show, using excluded volume (EV) limit ensemble calculations, that the Gaussian chain underestimates the effects expected for a fully unfolded protein. The observed PRE effects cannot be accounted for by the Gaussian chain or the EV limit model, but are well fit by all atom molecular simulations of the DSE. The combined experimental and computational approach reveals extensive, transient clusters of residues which are distant in primary sequence. The presence of long range interactions will limit the conformation search early in the folding process and has implications for experiments which probe collapse by monitoring R_g.

Autoinhibition and Salt Sensing are Linked in the WNK1 Kinase

Thomas M Moon[1], Fernando Correa[1,2], Lisa Kinch[1], Melanie H Cobb[2], Kevin Gardner[1,2], Elizabeth J Goldsmith[1]

[1]The University of Texas Southwestern Medical Center at Dallas, Biochemistry, Dallas, TX, 75390, [2]The University of Texas Southwestern Medical Center at Dallas, Pharmacology, Dallas, TX, 75390

Protein kinases control diverse cellular pathways and have sustained interest in understanding how they are controlled and how they maintain specificity toward substrates. This research focuses on a 230kDa serine/threonine protein kinase known as WNK1 (with no lysine {k}). The protein was first cloned by Melanie Cobb's laboratory, isoforms of which have been associated with a monogenic form of hypertension as well as breast and prostate cancer. Disease mutations cause an upregulation of WNK1 expession which increases overall kinase activity. This fact may indicate a unique treatment option for hypertension in general. However, upregulation of WNK1 is linked to increased cellular mobility and chromosomal instability which may implicate WNK as a potential target for cancer research. Overall, WNK1 is a very poorly understood kinase. However, our initial bioinformatics analysis of the autoinhibitory domain of WNK revealed homology to a domain in WNK's associated substrate OSR1. Analysis of this domain by NMR reveals a well-folded domain that is stable above 310K in vitro over long time periods and has a similar structure to that of the OSR1-PF2 domain.
Examination of this domain in the longer construct which also contains the kinase domain shows substantial conformational changes when dialized from high to low salt. Differential scanning fluorimetry conducted on the kinase domain shows that the domain undergoes substantial increases in domain stabilization as the concentration of salt is increased. Further analysis of this phenomena has pointed toward evidence of anion sensing by the WNK1 kinase domain. This observation stands in stark constrast to the fact that we do not measure a similar effect with other protein kinases that our lab studies. The coupling of the information that we have gathered on the autoinhibitory and kinase domains point to an overall mechanism of salt sensing and downstream signaling control in the WNK1 kinase

The SBGrid Consortium: Enabling structurally biology computation

Andrew Morin[1,2], Ben Eisenbraun[1], Piotrek Sliz[1,2]

[1]SBGrid.org, Harvard Medical School, Boston, MA, 02115, [2]Harvard Medical School, Biological Chemistry & Molecular Pharmacology, Boston, MA, 02115

The SBGrid Consortium is a computing collaboration of over 160 structural biology laboratories at more than 60 academic institutions and 3 pharmaceutical companies in 13 countries. SBGrid administers and supports over 230 different software titles used in structural biology research with an emphasis on x-ray crystallography, nuclear magnetic resonance, electron microscopy and light imaging applications. SBGrid software is automatically configured, tested and deployed to Consortium member laboratories without need for end-user intervention. Members have access to all consortium software through a unified command line interface. SBGrid provides OS-level technical support for all software and can be deployed on a variety of computing platforms. The SBGrid Consortium also engages in structural biology computation outreach and education for the benefit of its members, hosting regular webinars, symposia and computing schools conducted by leading developers of structural biology software. As active intermediary between scientist software developers and users, SBGrid expands access, facilitates communication and lowers the costs of computational structural biology research. The Consortium is operated as a member supported, non-profit service center of Harvard Medical School.

SAHG, a comprehensive database of predicted structures of all human proteins

Chie Motono[1], Ryotaro Koike[2], Kana Shimizu[1], Matsuyuki Shirota Shirota[3], Takayuki Amemiya[4], Kentaro Tomii[1], Nozomi Nagano[1], Naofumi Sakaya[5], Kiyotaka Misoo[5], Miwa Sato[1], Akinori Kidera[4], Hidekazu Hiroaki[6], Tsuyoshi Shirai[7], Kengo Kinoshita[3], Tamotsu Noguchi[1], Motonori Ota[2]

[1]National Institute of Advanced Industrial Science and Technology (AIST), Computational Biology Research Center (CBRC), Tokyo, 135-0064, Japan, [2]Nagoya University, Graduate School of Information Science, Nagoya, 464-8601, Japan, [3]Tohoku University, Graduate School of Information Science, Sendai, 980-8579, Japan, [4] Yokohama City University, Department of Supramolecular Biology, Yokohama, 1-7-29, Japan, [5]Information and Mathematical Science Laboratory Inc., Life Science, Tokyo, 112-0012, Japan, [6]Kobe University, Graduate School of Medicine, Kobe, 650-0017, Japan, [7]Nagahama Institute of Bioscience and Technology, Department of Bioscience, Nagahama, 526-0829, Japan

Most proteins from higher organisms are known to be multi-domain proteins and contain substantial numbers of intrinsically disordered (ID) regions. To analyse such protein sequences, those from human for instance, we developed a special protein-structure-prediction pipeline and accumulated the products in the SAHG (Structure Atlas of Human Genome) database at http://bird.cbrc.jp/sahg/. With the pipeline, human proteins were examined by local alignment methods (BLAST, PSI-BLAST, and Smith-Waterman profile-profile alignment), global-local alignment methods (FORTE), and prediction tools for ID regions (POODLE-S) and homology modeling (MODELLER). Conformational changes of protein models upon ligand-binding were predicted by simultaneous modeling using templates of apo and holo forms. When there were no suitable templates for holo forms, and the apo models were accurate, we prepared holo models using prediction methods for ligand-binding (eF-seek) and conformational change (the elastic network model and the linear response theory). Models are displayed as animated images.
As of July 2011, SAHG contains 42,581 protein-domain models in ~24,900 unique human protein sequences from the RefSeq database. Annotation of models with functional information and links to other databases such as EzCatDB, InterPro, or HPRD are also provided to facilitate understanding the protein structure-function relationships.
This presentation is supported by PDBj.
ref. Motono, C. et al., *Nucleic Acids Res.,* 2001, **39**, D487-493.

Structural analysis of BAR domain of Arfaptin in complex with small GTPase Arl1

Kensuke Nakamura[1], Zhiqiu Man[2], Xie Yong[3], Hisayoshi Makio[4], Masato Kawasaki[4], Ryuichi Kato[1,4], Hye-Won Shin[2,5], Kazuhisa Nakayama[2], Soichi Wakatsuki [1,4]

[1]The School of Advanced Studies, School of High Energy Accelerator Science, Ibaraki, 305-0801, Japan, [2]Kyoto University, Graduate School of Pharmaceutical Sciences, Kyoto, 606-8501, Japan, [3]Institute of Medicinal Plant, Chinese Academe of Medical Sciences & Peking Union Medical College, Beijing, 100094, China, [4] High Energy Accelerator Research Organization, Institute of Materials Structure Science, Ibaraki, 305-0801, Japan, [5]Kyoto University, Carrier-Path Promotion Uint for Yound Life Scientists, Kyoto, 606-8501, Japan

Arfaptins are the effectors of small GTPase Arl1 and are localized in trans-Golgi. It has been shown that Arfaptin-2, in particular, is capable of tubulating liposomes in vitro, and frequently found on dynamic vesicular and tubular structures emanating from the Golgi. Arfaptin-2 owes this characteristic to its C-terminal Bin/Amphiphysin/Rvs (BAR) domain. Arfaptin-2, through the BAR domain, binds to two small GTPases of different families: Rac1 and Arl1. Previous structural studies have shown that one Rac1 molecule binds to a concaved face of Arfaptin-2 dimer (Tarricone, et al., Nature, 2001), possibly inhibiting the direct membrane association of Arfaptin-2; however, the effect of Arl1 on Arfaptin-2 has not yet been revealed.
In this work, we obtained the crystal structure of the BAR domain of Arfaptin-1 and Arfaptin-2 in complex with two molecules of Arl1. The Arl1 bound to each side of the BAR domain dimer, leaving the concaved face open for membrane association. Physiologically-lipidated N-terminus of Arl1 faced the same direction as the concaved face. The complex structure suggests membrane association of Arfaptin-2 in complex with Arl1, which is in accordance with the report of Arl1-dependent localization of Arfaptin-2 to the Golgi. We performed Surface Plasmon Resonance experiments to characterize the interaction between Arfaptin-2 and the two small GTPases. The results showed that the BAR domain does not bind to more than two molecules of small GTPases simultaneously. This is in accordance with competition experiments of Rac1 and ARF1, a close homologue of Arl1, to Arfaptin-2 (Tarricone, et al., Nature, 2001).

Getting better at crystallisation by learning from the past

Janet Newman[1], David Ratcliffe[2], Vincent J Fazio[1], Kerry Taylor[2], Pascal Vallotton[3]

[1]CSIRO, MSE, Parkville, 3052, Australia, [2]CSIRO, ICT, Acton, 2601, Australia, [3]CSIRO, MIS, North Ryde, 1670, Australia

How many crystallisation experiments have been set up to get the vast number of crystal structures available now through the PDB? Any initiatives we have for capturing crystallisation information are currently geared towards capturing the information about successful experiments, and even these are imperfect. There is a real difference in what information is needed for a local crystallisation problem (for example, how to grow crystals to solve a specific crystal structure) and a global crystallisation problem (for example, determining the radius of convergence of an optimisation experiment)

What do we need to do in order to improve? We need a robust way of describing an experiment, as well as a way of capturing the outcome of the experiment and finally a way of describing the outcome of the experiment in some quantitative manner.

At the Collaborative Crystallisation Centre, we, together with a number of groups around the world have initiated projects in ontology development to enable the description of a crystallisation trial, as well as setting in place the IT infrastructure and workflows required to allow for the analyses of the millions of images that are produced by automatic crystal imagers. By so doing we aim to work on this second 'global' set of crystallisation problems, and thus make the process of solving local problems easier.

Observation of intermolecular interactions in large protein complexes by 2D-double difference NOESY: application to the 44 kDa interferon-receptor complex.

Ilona Nudelman[1,3], Sabine R Akabayov[1,4], Tali Scherf[2], Jacob Anglister[1]

[1]Weizmann Institute of Science, Department of Structural Biology, Rehovot, 76100, Israel, [2]Weizmann Institute of Science, Chemical Research Support, Rehovot, 76100, Israel, [3]The Rockefeller University, Laboratory of Cellular and Structural Biology, New York , NY, 10065, [4] Harvard Medical School, Department of Biological Chemistry and Molecular Pharmacology, Boston, MA, 02115

NMR detection of intermolecular interactions between protons in large protein complexes is very challenging since it is difficult to distinguish between weak NOEs from intermolecular interactions and the much larger number of strong intramolecular NOEs. This challenging task is exacerbated by the decrease in signal-to-noise ratio in the often used isotope-edited and isotope-filtered experiments as a result of enhanced T2 relaxation. Here we calculate a double difference spectrum that shows exclusively intermolecular NOEs and manifests the good signal-to-noise ratio in 2D homonuclear NOESY spectra even for large proteins. The method is straightforward and results in a complete picture of all intermolecular interactions involving non exchangeable protons. Ninety seven such 1H-1H NOEs were assigned for the 44 KDa interferon-α2/IFNAR2 complex and used for docking these two proteins. The symmetry of the difference spectrum, its superb resolution and unprecedented signal-to-noise ratio in this large protein/receptor complex suggest that this method is generally applicable to study large biopolymeric complexes.

Structural Studies of Streptococcal Adhesins

Åsa S Nylander, Karina Persson

Umeå University, Odontology, Umeå, 90187, Sweden

In the oral biofilm (dental plaque) more than 700 species of micro-organisms have been identified. The formation of dental plaque starts with bacteria adhering to proteins in the salivary pellicle covering the surfaces in the oral cavity, 130 proteins have been identified in the acquired enamel pellicle. The bound bacteria then become new attachment sites for following micro-organisms through surface proteins and carbohydrates. Streptococcal species are early colonizers and adhere to the salivary pellicle through several cell-surface proteins. Antigen I/II (AgI/II) polypeptides are a family of surface adhesins expressed by streptococci that interact with an array of host proteins and other bacteria. SspB and SpaP, AgI/II from *Streptococcus gordonii* and *Streptococcus mutans* respectively, both have a molecular weight of roughly 170 kDa divided into several domains. SspB from the commensal *S. gordonii* can be bound by *Porphyromonas gingivalis* fimbria Mfa1 and thereby be responsible for the colonization of the periodontal pathogen. Despite the structural similarities between SspB and *S. mutans* SpaP, *P. gingivalis* is unable to bind SpaP. The uncharacterized protein SGO_0707 from *S. gordonii*, with a molecular mass of approximately 200 kDa, is also predicted to be an important multi-domain adhesin. Here we present the crystal structure of the C-domain of *S.mutans* SpaP at 2.2 Å resolution and the N-terminal part of *S.gordonii* SGO_0707 at 2.1 Å.

Computational Docking Studies of Histone Deacetylase 8 (HDAC8) and Its Inhibitors by AutoDock Vina Program

Sang Ho Park, Byung Woo Han

Seoul National University, College of Pharmacy, Seoul, 151-742, South Korea

Histone deacetylases (HDAC) play crucial roles in gene transcription, regulation of cell cycle, and carcinogenic processes. A variety of novel HDAC inhibitors has been developed, such as SAHA, trichostatin, oxamflatin, and so on. Recently, Prof. Nam's group has synthesized analogues of SAHA which have benzothiazole-containing hydroxamic acid, and has shown inhibitory effects on HDAC enzymes. To get preliminary information on the binding mode of these compounds to HDAC, we selected two representative compounds (compound 5b and 5c) for docking study. For the control docking experiment using AutoDock Vina Program, we used the crystal structure of HDAC8-SAHA complex (PDB ID: 1T69). As a result, stabilization energies of predicted binding modes were -4.4 kcal/mol, -6.1 kcal/mol, and -6.1 kcal/mol for SAHA, compound 5b, and 5c, respectively. The docking result indicates that compound 5b and 5c have higher affinity to HDAC8 than SAHA.

Sugar transport inhibition of the sodium galactose transporter

Aviv Paz, Akira Watanabe, Ernest M Wright, Jeff Abramson

David Geffen School of Medicine, University of California, Department of Physiology, Los Angeles, CA, 90065

Solute sodium symporters are a family of proteins that co-transport Na^+ with sugars, amino acids, inorganic ions, or vitamins. Members of this family are important in human physiology and disease. The co-transport of glucose and Na^+ by the symporter SGLT is the basis for oral rehydration therapy, and SGLTs are being targeted in drug trials for type II diabetes.
More than 40 years ago Jardetzky theorized the alternating access mechanism as a means for substrate transport across biological membranes. According to this model, substrate binds to one face of the membrane protein and an energy input drives a conformational change to drive the substrate to the opposite side of the membrane where it is subsequently released. Alternating access has become the accepted model in the field, yet after over 40 years there is still no complete mechanism for any gene family that describes with atomic resolution the global rearrangements the transporter undergoes to complete the cycle of substrate binding and release. Over the last three years, there has been a surge of crystal structures for different Na^+ symporters, displaying a 5-helix inverted repeat motif as originally reported for LeuT and subsequently for the Vibrio paraheanoliticus sodium-galactose transporter (vSGLT). Although the substrates, the number of driving ligands as well as the conformational states in which these structures were captured differ, these structures share a common assembly of 10 transmembrane helices formed through the association of an internal structural repeat of a 5 transmembranal helical bundle.
In this poster we present a screen of ten inhibitors, some of which had previously been shown to competitively inhibit hSGLT1, based on transport activity assays of ^{14}C-D-Galactose tracer into vSGLT proteoliposomes. Five novel vSGLT inhibitors were identified, three of which were further characterized by steady-state fluorescence to determine apparent inhibition constants. Crystallization trails for four of the leading inhibitors are underway to shed light on the inhibition mechanism and to attempt to shift the conformations of the crystallized vSGLT from the inward facing and inward occluded conformations, that were previously reported by the Abramson lab, into other conformations that will lead to a more complete understanding of the alternating access mechanism and of sugar transport.

NMR solution-state structures of monomeric and dimeric G-quadruplexes adopted by a sequence from N-myc

Marko Trajkovski[1], Metka Vivod[1], Mateus Webba da Silva[2], Janez Plavec[1,3,4]

[1]National Institute of Chemistry, Slovenian NMR Centre, Ljubljana, SI-1000, Slovenia, [2]University of Ulster, School of Biomedical Sciences, Colerine, BT52 1SA, United Kingdom, [3]University of Ljubljana, Faculty of Chemistry and Chemical Technology, Ljubljana, SI-1000, Slovenia, [4] EN-FIST Center of Excellence, RRP 02, Ljubljana, SI-1000, Slovenia

Guanine-rich DNA regions can in the presence of cations, including biologically most relevant K+ and Na+ ions, form four stranded structures called G-quadruplexes. G-quadruplexes consist of stacked G-quartets, each formed by assembly of four Hoogsteen hydrogen-bonded guanines. Stability of G-quadruplexes depends on a number of stacked G-quartets as well as on length and sequence details of loops that connect guanines involved in G-quartets. Orientations of the loops are tightly related to strand directionality of a G-quadruplex and give rise to high heterogeneity of G-quadruplex folding topologies.

The Myc family comprises c-myc, N-myc and L-myc proto-oncogenes, which encode proteins involved in mechanisms regulating cell growth, differentiation, proliferation and apoptosis. Under normal conditions the c-myc gene is expressed in all proliferating cells, while expression of N-myc gene is restricted to specific tissues during their development. Deregulation of N-myc is associated most notably with neuronal tumors. Amplification and level of expression of N-myc oncogene have been widely studied due to their relation to pathophysiology of cancers, most notably neuroblastoma.

Multidimensional heteronuclear NMR study has demonstrated that DNA oligonucleotide containing d[AG3CG3AG3AG3A] sequence originating from N-myc gene folds into G-quadruplex structures in the presence of K+, NH4+ and Na+ ions. A monomeric G-quadruplex formed in K+ ion containing solution exhibits three G-quartets and flexible propeller-type loops. The 3D structure with three single nucleotide loops represents a missing element in structures of parallel G-quadruplexes. Increase of K+ ion or oligonucleotide concentrations resulted in transformation of monomeric G-quadruplex to dimeric form. Dimeric G-quadruplex exhibits six stacked G-quartets and parallel strand orientations. Loops in dimeric form adopt propeller-type orientations. Dimeric G-quadruplex features a link between the third and the fourth G-quartets consisting of two adenine residues that are flipped out to facilitate consecutive stacking of six G-quartets.

Mechanism of re-organization of flavivirus glycoproteins during maturation

Pavel Plevka[1], Anthony J Battisti[1], Jiraphan Junjhon[1,2], Dennis C Winkler[3], Heather A Holdaway[1,4], Poonsook Keelapang[2], Nopporn Sittisombut[5], Richard J Kuhn[1], Alasdair C Steven[3], Michael G Rossmann[1]

[1]Purdue University, Department of Biological Sciences, West Lafayette, IN, 47907, [2]Chiang Mai University, Department of Microbiology, Chiang Mai, 50200, Thailand, [3]National Institutes of Health, National Institute for Arthritis, Musculoskeletal and Skin Diseases, Bethesda, MD, 20892, [4] Case Western Reserve University, Cleveland Center for Membrane and Structural Biology, Cleveland, OH, 44106, [5]National Science and Technology Development Agency, National Center for Genetic Engineering and Biotechnology, Bangkok, 10700, Thailand

Flaviviruses assemble as fusion-incompetent immature particles and subsequently undergo a large conformational change in their glycoprotein envelope producing infectious mature virions. Nevertheless, “mosaic” particles that did not complete the maturation process and contain separate regions with immature and mature structures constitute ~10% of virions released from infected cells.

Here we report, using cryo-electron tomography and single particle analysis of dengue 2 and West Nile viruses, that each mosaic particle had glycoproteins organized into two regions of mature and immature structure separated by a narrow transition zone. Furthermore, the icosahedral symmetries of the mature and immature regions in a specific mosaic particle had different orientations.

It is, therefore, apparent that the maturation is initiated from a nucleation center from which the mature conformation propagates across the particle surface. Since the relative orientation of the immature and mature regions in one particle can be arbitrary, the maturation mechanism does not define strict rules for the glycoprotein re-arrangement. Instead, the individual glycoproteins are allowed sufficiently free movement to assemble into one of the adjacent positions in the mature region.

The evolution of robustness involves balancing of local sequence and network architecture level mechanisms.

Mario Pujato, Thomas MacCarthy, Andras Fiser, Aviv Bergman

Albert Einstein College of Medicine, Systems and Computational Biology, Bronx, NY, 10461

Gene regulatory networks often exhibit robustness to both genetic and environmental perturbations, yet the underlying molecular level mechanisms are still poorly understood. Previous studies have shown robustness to be an emergent property of network evolution. We adopt a modeling approach that integrates two levels: local sequence information and network architecture. Using a set of 10 TF-DNA complexes from PDB as templates we modeled TF binding to all possible nucleotide sequences of 8bp in length by generating all 4^8 models of TF-DNA complexes. A distance dependent statistical pair-potential was used to determine the free energy of each modeled structure. For each TF we obtained free energy estimates for all possible interacting DNA binding sites. This allows an explicit sequence-level representation of upstream regulatory regions to determine the architecture of a gene regulatory network model. In the model, mutations at the DNA level cause changes on two levels: (a) at the local sequence level in individual motifs, e.g. by changing the binding affinity, and (b) at the network architecture level by creating and destroying motifs, which results in dynamically rewiring network connections. The regulatory network model is used to determine gene expression dynamics while at a higher, population dynamics, level the networks undergo cycles of reproduction, mutation and selection. We use this multilevel hierarchical model to investigate the mechanisms underlying the evolution of robustness in gene regulatory networks. Preliminary results show that in sparse (low connectivity) architectures a mixture of local sequence level and network architecture level changes are exploited, whereas in highly interconnected architectures robustness evolves entirely as a consequence of network architecture level changes. Our results suggest that the increase in the level of interconnectedness over time brings a concomitant shift towards greater complexity at the network level and a reduced importance of local sequence level changes, thus providing a potential explanation for an evolutionary path towards greater complexity.

Crystal structures of the effector-binding domain of repressor CggR from Bacillus subtilis reveal ligand-induced structural changes upon binding of several glycolytic intermediates.

Pavlina Rezacova[1], Irena Sieglova[1], Milan Kozisek[1], Zbyszek Otwinowski[2]

[1]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Structural Biology, Prague, 76610, Czech Republic, [2]The University of Texas Southwestern Medical Center at Dallas, Biochemistry, Dallas, TX, 75390-8816

Expression of genes in the gapA operon encoding five enzymes for triose phosphate interconversion in Bacillus subtilis is negatively regulated by the Central glycolytic genes Regulator (CggR). CggR belongs to the large SorC/DeoR family of prokaryotic transcriptional regulators, characterized by an N-terminal DNA binding-domain and a large C-terminal effector-binding domain. When no glucose is present in growth media, CggR binds to its target DNA sequence and blocks the transcription of genes in the gapA operon. In the presence of glucose, binding of the known effector molecule fructose-1,6-bisphosphate abolishes this interaction. We have identified dihydroxyacetone phosphate, glucose-6-phosphate, and fructose-6-phosphate as additional CggR ligands that can bind to the effector-binding site. Crystal structures of C-CggR, the C-terminal effector-binding domain of CggR, both unliganded as well as in complex with the four ligands at resolutions between 1.65 and 1.80 Å reveal unique ligand-specific structural changes in the binding site that affect the dimer interface. Chemical cross-linking shows that CggR oligomerization is mediated through its effector-binding domain and that binding of the different ligands differentially affects the distribution of oligomers. Our results suggest that CggR function may be modulated by various effectors in a complex fashion.

A computational method to design backbone conformations featuring predefined interactions

Florian Richter[1,2], David Baker[1]

[1]University of Washington, Department of Biochemistry, Seattle, WA, 98195, [2]University of Washington, Interdisciplinary Program in Biomolecular Structure and Design, Seattle, WA, 98195

Computational *de novo* Enzyme Design is the process of creating from scratch an enzyme that catalyzes a reaction of interest. Starting from a theozyme, which is a spatial model of the reaction's transition state together with the critical protein active site residues that stabilize it, and a library of protein scaffolds, sites in the library are found where the theozyme can be attached. Then, amino acid identities for the remaining residues that surround these attachment sites are determined through computational design algorithms.
In the current standard protocol, the backbone of the protein scaffold stays fixed. Experience has shown that with this fixed-backbone approach, active sites will only be found for theozymes containing a limited number(3-5) of amino acids. There are reactions however, for which more amino acids are needed to stabilize the transition state.
Here, a generalized method is presented to vary a protein scaffold's backbone conformation such that sites for theozymes with an arbitrary number of amino acids can be found. In this method, a partial version of the theozyme is first placed into a scaffold. Then, for the remaining side-chains of the theozyme, disembodied rotamers are placed according to their interactions with the partial theozyme. Nearby backbone segments of the scaffold are then remodeled to overlay with the backbone of the disembodied rotamers. As a proof of principle, this method is then used to redesign the backbone of *de novo* designed low-activity esterases, in hopes of generating variants with higher activity.

Elucidating the Mechanism of Action of the Translation Factor BipA

Victoria L Robinson[1], Alicia E Every[1], Balakrishnan S Moorthy[2], Ganesh S Anand[2]

[1]University of Connecticut, Molecular and Cell Biology, Storrs, CT, 06269, [2]National University of Singapore, Department of Biological Sciences, Singapore, 117543, Singapore

BipA is a highly conserved prokaryotic GTPase that is necessary for securing bacterial survival and successful invasion of the host. It has been implicated in controlling bacterial motility, modulating attachment and effacement (A/E) processes, avoidance of host defense mechanisms and upregulating the expression of virulence genes. In addition, BipA is involved in bacterial stress responses particularly those related to virulence, temperature and symbiosis. Our group discovered that BipA has two distinct ribosome binding modes. Under normal growth conditions, it binds to 70S ribosomes. However, it associates with 30S ribosomal particles in the presence of ppGpp, an alarmone signaling the onset of stress in bacteria. At 67 kDa, it is one of the larger members of the translational family of bacterial GTPases along with EF-G and LepA. As part of an interdisciplinary approach to answer fundamental questions about the mechanism of action of BipA, structural studies have been done to acquire an atomic-level picture of various ligand bound forms of the protein. Our crystallographic models indicate that although the gross morphological features of BipA, EF-G and LepA are similar, the C-terminal domain (CTD) of BipA has a unique fold. Using a series of deletion constructs, we demonstrated that this CTD is crucial, but not sufficient, to promote the interaction of BipA with the 70S ribosome. Moreover, removal of the CTD results in an increase in the GTPase activity of the protein and the loss of 70S stimulated GTP hydrolysis activity. Taken together, our structural and biochemical studies have defined an allosteric pathway that directs this differential binding of the protein to the ribosome.

Protein Structures from the Perspective of Side-Chain Chemical Shifts

Aleksandr B Sahakyan[1], Wim F Vranken[2], Andrea Cavalli[1], Michele Vendruscolo[1]

[1]University of Cambridge, Department of Chemistry, Cambridge, CB2 1EW, United Kingdom,
[2]Vrije Universiteit Brussel, Structural Biology Brussels, Brussel, 1050, Belgium

In contrasts to the ever-growing use for structure determination of chemical shifts for protein backbone atoms, there have been a few studies concerned with chemical shifts for side-chains. It would be highly desirable to extend such studies, since side-chains are excellent probes, which are naturally incorporated in proteins, of biomolecular structure, dynamics and recognition. In particular, proteins are rich of side-chain methyls and conjugated rings, many of which, owing to their hydrophobic nature, are more frequently placed at the inner core and protein-protein interfaces. This work presents the development and use of the structure-based predictors for methyl and aromatic side-chain chemical shifts by incorporating terms convenient for the introduction of restraining forces in molecular dynamics simulations. The predictors are tested for their usefulness in validation of protein structures and dynamical ensembles and are of great importance for exploring protein-protein and protein-ligand assembly and interactions. Web server and stand-alone implementations are created to facilitate the usage of the developed tools. Their performance is demonstrated to already provide an opportunity for their immediate implementation in restrained simulations to refine protein structure and dynamics from the side-chain perspective.

LidA from Legionella is a Rab-supereffector

Stefan Schoebel, Adam L Cichy, Roger S Goody, Aymelt Itzen

Max Planck Institute of Molecular Physiology, Department of Physical Biochemistry, Dortmund, 44227, Germany

Manipulation of intracellular transport is a means often exploited by bacterial pathogens to camouflage and replicate inside a eukaryotic host cell. Legionella pneumophila, the causative agent of Legionnaire's disease injects about 250 proteins into the host cytosol that reprogram the fate of the eukaryotic host cell in order to establish a replication permissive vacuole.
Rab proteins play a crucial role in the regulation of vesicular transport in eukaryotic cells and are targeted by Legionella proteins to redirect trafficking processes between the endoplasmatic reticulum and the Golgi apparatus to the Legionella containing vacuole (LCV).
The Legionella protein LidA (84 kDa) is important for intracellular growth of Legionella [1] and has been described as a Rab-effector [2]. However, the precise function of LidA is unclear. Recent publications reported unusual Rab-effector properties of LidA that prompted us to investigate its biochemical and structural basis.
We determined the crystal structure of the Rab-binding domain of LidA in complex with active Rab8 at 2.5 Å resolution. Additionally we performed a detailed kinetic analysis of the binding between various Rab proteins and LidA [3]. LidA uses a novel Rab binding mode compared to known mammalian Rab effectors to form a considerably extended binding interface. Rab8 is embraced by 4 α-helices and 2 pillar-like structures of LidA. Additional contacts to regions outside the switch regions (commonly involved in Rab:effector interactions) indicate an exceptional high stability of the Rab8a-LidA interactions. For Rab1b and Rab8a the affinities are extraordinarily high, but for the more weakly bound Rab6a, KD values of 4 nM for the inactive and 30 pM for the active form were determined. Rab1b and Rab8a appear to bind LidA with KD values in the low pM or even lower range, making LidA a Rab-supereffector.
So far, Rab-binding partners are known that have high affinity for only the GDP- or the GTP-state, but not for both. LidA is a remarkable exception in this respect since it can bind to Rabs even in the inactive, GDP-bound form with at least nanomolar affinity.

1. Conover, G.M., et al., The Legionella pneumophila LidA protein: a translocated substrate of the Dot/Icm system associated with maintenance of bacterial integrity. Molecular microbiology, 2003. 48(2): p. 305-21.
2. Machner, M.P. and R.R. Isberg, Targeting of host rab GTPase function by the intravacuolar pathogen Legionella pneumophila. Developmental cell, 2006. 11(1): p. 47-56.
3. Schoebel S., Cichy A.L., Goody R.S., Itzen A., LidA from Legionella is a Rab-supereffector. PNAS, 2011 (in press)

X–ray crystallographic analysis of Runx1–CBFβ–Ets1–DNA complex assembled on the enhancer of T cell receptor α chain gene

Masaaki Shiina[1], Keisuke Hamada[1], Taiko Inoue-Bungo[1], Mariko Shimamura[1], Akiko Uchiyama[1], Shiho Baba[1], Ko Sato[1], Masaki Yamamoto[2], Tahir H Tahirov[3], Kazuhiro Ogata[1]

[1]Yokohama City University Graduate School of Medicine, Biochemistry, Yokohama, 236-0004, Japan, [2]RIKEN, SPring-8 Center, Hyougo, 679-5148, Japan, [3]University of Nebraska Medical Center, Eppley Institute for Research in Cancer and Allied Diseases, Omaha, NE, 68198-7696

Transcription factors (TFs) act on various target genes in combination with other partner TFs in a cell, enabling complicated gene regulation. The underlying mechanism involves cooperative DNA binding, by which one recruits other partner nearby on a specific target enhancer. However, a structural insight for the binding cooperativity in higher-order TFs–DNA complexes is almost elusive. To reveal the underlying mechanism, we performed crystallographic analyses of complexes composed of Ets1, Runx1, CBFβ and DNA derived from the *TCRα* gene enhancer.

Ets1 and Runx1 are transcription factors that play important roles in differentiation of hematopoietic cells and are involved in leukemogenesis. It was reported that Ets1 activity is negatively regulated by the regulatory region (the exon VII) flanking its DNA binding domain (so called the ETS domain) and phosphorylation of the exon VII suppresses the DNA binding activity of Ets1 more strongly. Runx1, which binds to the *TCRα* gene enhancer with Ets1 side by side, counteracts the self-inhibition of DNA binding activity of Ets1, leading to the cooperative DNA binding.

In this study, we reveal the mechanism by which Runx1 and Ets1 cooperatively bind to the *TCRα* gene enhancer using crystallographic and biochemical analyses. We will also discuss structural insight of the effect of Ets1 phosphorylation on multiple TFs–DNA complex formation.

Structure of the effector-binding domain of arabinose repressor AraR from Bacillus subtilis

Irena Sieglova[1], Katerina Prochazkova[1], Petr Pachl[1], Milan Fabry[2], Pavlina Rezacova[1]

[1]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Structural Biology, Prague, 16610, Czech Republic, [2]Institute of Molecular Genetics Academy of Sciences of the CR, Structural Biology, Prague, 14220, Czech Republic

Abstract Arabinose repressor AraR in Bacillus subtilis negatively controls expression of genes in the metabolic pathway of arabinose containing polysaccharides. The protein is composed of two domains of different phylogenetic origin and function: the N-terminal DNA binding domain belonging to GntR family and the C-terminal effector binding domain, a GalR/LacI family member. We determined the crystal structure of the C-terminal effector-binding domain of AraR in complex with the effector L-arabinose at 2.2 Å resolution. The L-arabinose binding affinity was characterized by isothermal titration calorimetry and differential scanning fluorimetry, the Kd value was 8.4 ± 0.4 μM. Effect of L-arabinose on the protein oligomeric state was investigated in a solution and a detailed analysis of crystal identified dimer organization distinctive from other members of the GalR/LacI family.

Accurate design of a symmetric homodimer using β-strand assembly

Peter B Stranges[1], Mischa Machius[2], Michael J Miley[2], Brian Kuhlman[1]

[1]University of North Carolina Chapel Hill, Biochemistry and Biophysics, Chapel Hill, NC, 27599, [2]University of North Carolina Chapel Hill, Pharmacology, Chapel Hill, NC, 27599

Computational design of novel protein-protein interfaces is a test of our understanding of protein interaction biophysics and has the potential to allow modification of cellular physiology. Methods for designing new, high affinity, interactions that adopt a predetermined binding mode have proved elusive, suggesting that new approaches are needed which simplify the design process. Solvent exposed β-strands are thought of as 'sticky' and prone to self-assembly. β-strand pairing also stabilizes many naturally occurring protein complexes. Here, we use Rosetta to computationally redesign a monomeric protein to form a symmetric homodimer by pairing exposed β-strands to form an intermolecular β-sheet. One of the designed proteins was experimentally determined to form a dimer in solution with a dissociation constant less than 1 μM. A crystal structure of the designed homodimer closely matches the computational model (RMSD = 1.00 Å). This work demonstrates that β-strand pairing can be used to computationally design new protein interactions with high accuracy.

The Jena Library of Biological Macromolecules - JenaLib

Rolf Huehne, Juergen Suehnel

Leibniz Institute for Age Research - Fritz Lipmann Institute, Biocomputing Group, Jena, D-07745, Germany

The Jena Library of Biological Macromolecules – JenaLib (www.fli-leibniz.de/IMAGE.html) offers value-added information for all entries included in the Protein Data Bank (PDB) and Nucleic Acid Database (NDB).

Its features include:
- PDB/NDB atlas pages
- PDB sequence information extracted from atomic co-ordinates
- PDB/UniProt alignments that clearly indicate gaps, mutations, numbering irregularities and modified residues
- Integration of information on single amino acid polymorphisms (SAPs), PROSITE,motifs, CATH, SCOP, Pfam and the exon structure as well as Gene Ontology and species information with PDB data in one comprehensive database
- A platform-independent Jmol-based molecule viewer (Jena3D) that offers an integrated viewing of most of these data far beyond the information included in the PDB files alone
- SCOP, CATH and Pfam tree browsers
- Direct pattern search in protein and amino acid sequences
- A BLAST sequence homology search option
- Calculation of mean amino acid distances for individual structures and structured sets
- An easily customizable generation of entry lists
- A simple but powerful QuickSearch option that allows searching for PDB/NDB code, UniProt ID/Accession and other search terms in one input field.

Offering all these data and analysis tools in one place makes the JenaLib database an interesting resource for the dissemination of 3D structural information on biological macromolecules that complements information provided by the Protein Data Bank and other related resources.

The Jena3D viewer for visualization of biological macromolecules

Rolf Huehne, Juergen Suehnel

Leibniz Institute for Age Research - Fritz Lipmann Institute, Biocomputing Group, Jena, D-07745, Germany

The Jena3D viewer (jena3d.fli-leibniz.de) is a web-based interactive molecular viewer for three-dimensional biopolymer structures. Jen3D provides access to all structure entries deposited at the Protein Data Bank (PDB) and the Nucleic Acid Database (NDB). In addition local structure files can be uploaded and viewed. It is based on Jmol, an open source Java applet. The Jena3D viewer visualizes not only information included in the PDB files but also SCOP, CHATH and Pfam domains, single amino acid polymorphisms, PROSITE motifs and the exon structure mapped onto proteins. A versatile JavaScript interface allows the easy exploration of 3D structures. The Jena3D viewer is closely linked to the Jena Library of Biological Macromolecules (JenaLib).

Structure and catalytic mechanism of iron-dependent enzymes

Hiroshi Sugimoto, Yoshitsugu Shiro

RIKEN SPring-8 Center, Biometal Science Laboratory, Hyogo, 679-5148, Japan

Transition metal ions are widely distributed in biological systems in the form of metal-binding proteins and enzymes. They are involved in many important physiological actions such as amino acid metabolism, energy conversion and signal transductions. We analyzed some iron-dependent enzymes including heme-binding proteins using X-ray crystallography.
(1) Human indoleamine 2,3-dioxygenase (IDO): This heme-containing enzyme catalyzes very unique reaction in which both atoms of O_2 are incorporated into Trp without any supply of electron or proton. Structural analysis suggests no catalytic protein group is essential but strict geometry within the ternary complex is required [1, 2].
(2) Nitric oxide reductase (NOR): NOR is an iron-containing enzyme that catalyzes the reduction of nitric oxide (NO) to nitrous oxide (N_2O). Crystal structure shows that the non-heme iron (Fe_B) is coordinated by three His and one Glu ligands, but a His-Tyr covalent linkage common in cytochrome c oxidases is absent. Structure also revealed the proton transfer pathway from extracellular side to active center [3].
(3) Aliphatic aldoxime dehydratase (Oxd): The crystal of enzyme-substrate complex in resting state was converted into the Michaelis complex by X-ray photoreduction of the heme iron atom in the active site. Structure shows a unique reaction intermediate in which an organic substrate is directly bound to the heme iron [4].

[1] Sugimoto *et al. Proc. Natl. Acad. Sci. U.S.A.* 103, 2611 (2006)
[2] Chung *et al. J. Am. Chem. Soc.* 132, 11993 (2010)
[3] Hino *et al. Science* 330, 1666 (2010)
[4] Sawai *et al. J. Biol. Chem.* 284, 32089 (2009)

Recent progress in a Stereo-Array Isotope Labeling (SAIL) method for NMR structure determination

Mitsuhiro Takeda[1], Masatsune Kainosho[1,2]

[1]Nagoya University, Graduate School of Science, Nagoya, 464-8601, Japan, [2]Tokyo Metropolitan University, Center for Priority Areas, Tokyo, 192-0397, Japan

We have been developing a stereo-array isotope labeling (SAIL)-NMR method. This method utilizes a protein exclusively or residue-selectively composed of synthesized amino acids with an intended isotope labeling pattern. Here we present two topics related to the SAIL-NMR method. The first is a high-quality observation of aromatic peaks by using δ, e and ζ-SAIL phenylalanine (Phe) and δ and e-SAIL tyrosine (Tyr). Each of the various types of SAIL Phe and SAIL Tyr yields well-resolved resonances for the δ-, e-or ζ-13C/1H signals, respectively, which can readily be assigned by simple and robust pulse sequences. Since the δ-, e-, and ζ-proton signals of Phe/Tyr residues give rise to complementary NOE constraints, the concomitant use of various types of SAIL-Phe and SAIL-Tyr would generate more accurate protein structures, as compared to those obtained by using conventional uniformly 13C , 15N-double labeled proteins. The second topic is a newly developed method for monitoring the hydrogen exchange rate of side-chain hydroxyl (OH) and sulfhydryl (SH) groups in proteins. In this approach, carbon atoms attached to OH/SH groups enriched by 13C are directly observed in an H2O/D2O (1:1) solution. Due to proton/deuterium isotope shift effect, the carbon atom attached to slowly exchanging OH/SH groups gives isotopomer-resolved peaks, while those attached to rapidly exchanging OH/SH groups averaged peaks. For the identified slowly exchanging OH/SH groups, 1H-1H NOE restraints involving the OH/SH protons can be collected, which is of use to refine NMR protein structures and determine pairs forming the side-chain hydrogen bonds.
This presentation is supported by PDBj and JSPS

Discriminating thermophilic proteins from their mesophilic analogs

Todd J Taylor[1], Iosif I Vaisman[2]

[1]National Cancer Institute, Lab of Molecular Biology, Bethesda, MD, 20892, [2]George Mason University, Department of Systems Biology, Manassas, VA, 20110

The physical basis of thermostability in proteins has been the subject of a great many papers going back 30 years Some of the proposed mechanisms/indicators of increased thermostability include: a more highly hydrophobic core, tighter packing or compactness, deleted or shortened loops, greater rigidity, higher secondary structure content, greater polar surface area, fewer and/or smaller voids, smaller surface area to volume ratio, fewer thermolabile residues, increased hydrogen bonding, higher isoelectric point, and more salt bridges (ion pairs) and networks of salt bridges.
We have assembled two sets of thermophile-mesophile protein structure pairs-one containing chains from moderate thermophiles (132 pairs) and one containing chains from extreme thermophiles (149 pairs). Sequence identity between the members of a pair varies between 10% and 88%. The pairs have been chosen so that the structural superpositions are no worse than ~5 Angstroms, at least 80% of the residues from each align, and the proteins have the same functions. They were chosen to be as similar as possible except that one is stable and functional at high temperature and the other is not. We have then computed a large number of simple indices, one number for each chain (e.g. mean hydrophobicity of the core, surface to volume ratio), that measure some of the properties previously claimed to be associated with or to cause thermostability. To measure the strength of the association, we simply calculated the fraction of those pairs for which the index is systematically higher or lower for the thermophilic member of the pair. Many of the structure related indices were derived from the Delaunay tessellation of protein structures.
It turns out that it is possible to accurately discriminate between mesophile/ thermophile protein pairs. Purely sequence composition based indices , used previously on entire proteomes, are very good at discriminating. individual thermophile/mesophile protein pairs. Purely structure-based indices like surface area to volume ratio or measures of packing are not as good. Measures involving charged residues like the number of putative salt bridges on the surface are very good discriminators, particularly for extreme thermophiles. Mixed sequence structure descriptors derived from the Delaunay tessellation of protein structures are also good very discriminators.

The Biological Magnetic Resonance Data Bank (BMRB)

Eldon L Ulrich[1], Alexander Anderson[1], Jessica Frosch[1], Danielle Henak[1], John Idso[1], Yannis E. Ioannidis[2], Francisca Jofre[1], Jordan Kutil[1], Miron Livny[3], Steve Mading[1], Matthew Murphy[1], Dimitri Mazuik[1], Christopher Schulte[1], Jonathon R Wedell[1], Melanie Ulrich[1], R.K. Wenger[3], Hongyang Yao[3], John L. Markey[1]

[1]University of Wisconsin, Dept. of Biochemistry, Madison, WI 53706, [2]University of Athens, Dept. of Informatics and Telecommunications, Athens, Greece 15784, [3]University of Wisconsin, Dept. of Computer Sciences, Madison, WI 53706

BMRB (www.bmrb.wisc.edu), as a member of the wwPDB, provides the scientific community with a growing archive of data on the structure, dynamics, and chemical properties of biological macromolecules important for understanding structure-function relationships. The BMRB repository contains >7200 entries for proteins, peptides, and nucleic acids. These entries include 1H, 13C, and 15N chemical shift values from peptides and proteins (>4,600,000) and nucleic acids (>60,000), J-coupling values (>26,000), R1, R1rho, R2 relaxation data and order parameters (>75,000), binding constants, chemical rate constants, and hydrogen exchange rate constants. Raw NMR data (time-domain spectra with acquisition and processing parameters) are available from over 180 structural studies of proteins and nucleic acids.

Structural Basis for the High Alkaliphily of Modular Xylanase XynJ from *Bacillus* sp. 41M-1

Ihsanawati Wahab[1,2], Guntur Fibriansah[1], Hirohito Umemoto[3], Rie Yatsunami[3], Nobuo Tanaka[1], Satoshi Nakamura[3], Takashi Kumasaka[1,4]

[1]Tokyo Institute of Technology, Department of Life Science, Yokohama, 226-8501, Japan, [2]Institut Teknologi Bandung, Biochemistry Research Division, Bandung, 40132, Indonesia, [3]Tokyo Institute of Technology, Department of Bioengineering, Yokohama, 226-8501, Japan, [4] Japan Synchrotron Radiation Research Institute/Spring-8, Structural Biology Group, Hyogo, 679-5198, Japan

Xylanase J from *Bacillus* sp. 41M-1 (XynJ) has a highest alkaliphily with an optimum pH of 9.0, which is hoped to be utilized in papermaking, food industries, and biofuel production. To understand its alkaliphily, we determined the crystal structures of XynJ in two crystalline forms from neutral and basic pH conditions. XynJ is a modular enzyme composed of a catalytic domain of family 11 glycoside hydrolase (GH11) accompanying a xylan binding domain of family 36 carbohydrate binding module (CBM) via a flexible linker. The structure of the catalytic domain is highly homologous to other GH11 enzymes, and the comparison of both forms shows the movement of the thumb-like loop as a common GH11 feature. On the alkaliphily of the enzyme, unique features were observed as salt bridges around acid-base catalyst Glu183 and nucleophile Glu93, and salt bridge network Glu16-Arg48-Glu177-Lys52. In addition, the xylan-binding CBM enhances the affinity for insoluble xylan and shows a novel role in increasing the total catalytic activity of XynJ at highly alkaline conditions.

The role of peptide-peptide stacking interactions in protein structure

Gary E Wesenberg[1,2], Julie C Mitchell[1,2]

[1]University of Wisconsin-Madison, Biochemistry, Madison, WI, 53706-1544, [2]University of Wisconsin-Madison, Mathematics, Madison, WI, 53706-1388

A comprehensive search of protein structures in the PDB was made for the presence of planar peptide-peptide stacking configurations. Similar to other stacking interactions between aromatic side chains, peptide units from neighboring backbones were found in planar orientations at distances within 4.0 Angstroms. Though the frequency of peptide-peptide stacking interactions was not high, cases were observed in which the interactions played special roles, such as within helix-helix interfaces, in both intra-molecular and inter-molecular settings. Of the four basic ways for peptide stacking to occur, only two were found in proteins. Defining the positive direction of the normal vector of the peptide plane as the direction of the cross-product of the CN bond with the CO bond, the observed peptide stacking in proteins showed these vectors in the opposite directions. The stacking geometry may be further characterized by the relative directions of the CN bond vectors (from C to N). Both parallel and anti-parallel orientations of the CN bonds were observed in proteins. Along with the proposed π-π stabilization of peptide stacking interactions, both of the observed types of peptide stacking directions often presented tandem $C_{\alpha}H...O=C$ hydrogen bonds between the peptides. In the parallel case, both hydrogen bonds occurred on the same end of the stacked peptides, whereas in the anti-parallel case, one hydrogen bond occurred on each end of stacked peptides. Dependencies on amino acid composition were examined, and a high percentage of the stacked peptides contained glycine and alanine, allowing more open access to the backbone for stacking. In an effort to estimate the stabilization energy of stacked peptide-peptide pairs in various observed configurations, quantum mechanical methods were used. Preliminary results using semi-empirical methods (MOPAC/PM6-DH+) gave typical stabilization energies of 5 kcal/mole.

In crystallo synthesis of the tryptophan tryptophylquinone cofactor of methylamine dehydrogenase

Carrie M Wilmot[1], Erik T Yukl[1], Lyndal M Jensen[1], Nafez A Tarboush[2], Victor L Davidson[2,3]

[1]University of Minnesota, Dept. Biochemistry, Molecular Biology & Biophysics, Minneapolis, MN, 55455, [2]University of Mississippi Medical Center, Dept of Biochemistry, Jackson, MS, 39216, [3]University of Central Florida, Burnett School of Biomedical Sciences, Orlando, FL, 32827

Methylamine dehydrogenase (MADH), a metabolic enzyme found in methylotrophic/autotrophic bacteria, contains a quinone cofactor, tryptophan tryptophylquinone (TTQ), derived from the post-translational modification of two Trp residues in the protein. MauG is a highly unusual *c*-type di-heme enzyme responsible for the completion of TTQ synthesis. The natural substrate for MauG (preMADH) is a 119-kDa protein precursor of MADH with a partially formed cofactor. MauG catalyzes a six-electron oxidation to complete TTQ biosynthesis, using either hydrogen peroxide or oxygen (with reducing equivalents) as the second substrate. The activation of molecular oxygen is highly unusual for a *c*-type heme system.

The crystal structure of the MauG-preMADH complex was solved to 2.1 Å resolution (PDB ID 3L4M), and the crystals support TTQ formation (PDB ID 3L4O) [1]. One of the MauG hemes has an unusual His-Tyr axial ligation that is required for activity (PDB ID 3ORV) [2]. The catalysis occurs via long-range electron transfer as the oxygen-binding heme is ~ 40 Å from the site of TTQ synthesis (PDB IDs 3PXT & 3PXW) [3]. This presentation will focus on the latest unpublished work; (1) the order of the three MauG-catalyzed 2-electron modifications that create TTQ, and (2) in crystallo trapping and characterization of the novel MauG oxidant that consists of a ferryl heme (Fe(IV)=O) and a heme in the Fe(IV) oxidation state. X-ray crystallography, mass spectrometry and single crystal spectroscopies have been coupled to site-directed mutagenesis to achieve these aims.

References:

[1] Jensen, L.M.R., Sanishvili, R., Davidson, V.L. & Wilmot, C.M. (2010). In crystallo posttranslational modification within a MauG/pre-methylamine dehydrogenase complex. Science 327: 1392-1394.

[2] Tarboush, N.A., Jensen, L.M.R., Feng, M., Tachikawa, H., Wilmot, C.M. & Davidson, V.L. (2010) Functional importance of tyrosine 294 and the catalytic selectivity for the bis-Fe(IV) state of MauG revealed by replacement of this axial heme ligand with histidine. Biochemistry 49: 9783-9791.

[3] Yukl, E. T., Goblirsch, B.R., Davidson, V.L. & Wilmot, C.M. (2011) Crystal structures of the CO and NO adducts of MauG in complex with pre-methylamine dehydrogenase: Implications for the mechanism of oxygen activation. Biochemistry 50: 2931-2938.

Structural basis of the Ca^{2+} inhibitory mechanism of Drosophila Na^{+}/Ca^{2+} exchanger CALX and its modification by alternative splicing

Mousheng Wu, Shuilong Tong, Jennifer Gonzalez, Vasanthi Jayaraman, John Spudich , Lei Zheng

The University of Texas Houston Medical School, Center for Membrane Biology, Biochemistry and Molecular Biology, Houston, TX, 77030

Na^{+}/Ca^{2+} exchangers play important roles in Ca^{2+} homeostasis in nearly all tissues. They trade intracellular Ca^{2+} with extracellular Na^{+} across the cell membrane, attenuating Ca^{2+}-mediated cellular signaling processes. The exchanger activity is tightly regulated by specific binding of intracellular Ca^{2+} on its two Ca^{2+} regulatory domains (CBDs). However, how Ca^{2+} binding induces protein conformational change of CBDs to control the exchanger activity remains unknown. To gain structural insight on Ca^{2+} regulatory mechanism, we studied the Na^{+}/Ca^{2+} exchanger CALX facilitating light-mediated Ca^{2+} homeostasis in Drosophila sensory neuronal cells. CALX activity is negatively regulated by specific Ca^{2+} interaction within its two intracellular Ca2+ regulatory domains CBD1 and CBD2. We determined crystal structures of the entire Ca^{2+} regulatory domain CBD12 from two alternative splicing isoforms, CALX 1.1 and CALX 1.2, exhibiting distinct regulatory Ca^{2+}-dependency, at 2.3 or 2.9 Å resolution, respectively. The structures show that CBD1 and CBD2 form an open V-shaped conformation with four Ca^{2+} ions bound on the CBD domain interface, confirmed by LRET analysis. The structures together with Ca^{2+} binding analyses of mutations support that the Ca^{2+} inhibition of CALX is achieved by interdomain conformational changes induced by Ca^{2+} binding at CBD1. The conformational difference between the two isoforms also indicates that alternative splicing adjusts the interdomain orientation angle to modify the Ca^{2+} regulatory property of the exchangers. This pair of structures provide a structural model to study Ca^{2+} regulatory mechanism and its alternative splicing modification of the Na^{+}/Ca^{2+} exchanger family.

Structure, Substrate Specificity and Mechanism of Kinase Activation of the IκB Kinase β (IKKβ)

Guozhou XU[1], Yu-Chih LO[1], Gennaro Napolitano[2], Xuefeng WU[2], Michael Karin[2], Hao WU[1]

[1]Weill Cornell Medical College, Department of Biochemistry, New York, NY, 10021,
[2]University of California at San Diego, Department of Pharmacology, La Jolla, CA, 92093

Inhibitor of κB (IκB) kinase (IKK) is required for the activation of NF-κB by phosphorylating its inhibitor protein IκB, leading to its proteasomal degradation. This then allows NF-κB to enter the nucleus for target gene transcription. IKK is a large protein complex containing the kinase subunit IKKα and/or IKKβ, and the regulatory protein NEMO (IKKγ). IKKβ contains an N-terminal kinase domain, a ubiquitin-like domain (ULD), predicted leucine zipper (LZ) and helix-loop-helix (HLH) domains and a C-terminal NEMO-binding domain (NBD). It plays a dominant role in the canonical NF-κB pathway by phosphorylating IκBα. IKKα shares more than 50% sequence identity with IKKβ and plays an indispensible role in the non-canonical NF-κB pathway. We have determined the crystal structure of IKKβ at 3.6 Å resolution. The structure reveals a tri-modular architecture that contains the kinase domain (KD), a ubiquitin-like domain (ULD) and a highly elongated, α-helical scaffold/dimerization domain (SDD). One surprising finding is that the LZ and HLH domains were falsely predicted; instead, this region of the structure forms an elongated helical bundle we named the scaffold/dimerization domain (SDD). IKKβ exists as dimers in solution and together with biochemical and mutagenesis studies, we showed that dimerization is mediated by the SDD domain. The IKKβ dimer structure has the shape of a pair of shears with the KD and the ULD forming the "handle" and the SDD being the "blade". The three domains intimately interact with each other, which makes IKKβ an integral trimodular unit. Our functional assays further elucidated that the ULD and SDD mediates a critical interaction with IκBα to restrict substrate specificity, and the ULD is required for catalytic activity. Interestingly, dimerization is no longer necessary to maintain the kinase activity once IKKβ is activated. However, dimerization of IKKβ is required for its activation upon overexpression. The structure suggests that other IKK family members IKKα, TBK1 and IKKi share the similar tri-modular architecture but likely have evolved different functions for the individual domains.

Automation of data collection and processing at the Photon Factory macromolecular crystallography beamlines

Yusuke Yamada, Masahiko Hiraki, Naohiro Matsugaki, Leonard Chavas, Noriyuki Igarashi, Soichi Wakatsuki

Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization, Tsukuba, Ibaraki, 305-0801, Japan

Recent advances in high-throughput techniques have highlighted the importance of macromolecular crystallography in the field of pharmaceutical researches. Highly intense and stable X-ray beams and fast detectors have been available in beamlines and a time required in the collection of a diffraction data set has been drastically decreased. Furthermore, introduction of sample exchanger in a beamline experimental station makes the experiment by users more easy, because users do not need to enter/exit the experimental hutch and be troubled in the sample mounting on the diffractometer. With these developments, pharmaceutical researches such as structure-based drug design (SBDD) and fragment-based drug discovery (FBDD), where enormous number of protein/small compound complexes should be examined, are valid in synchrotron facilities. However, to facilitate these researches, further developments especially in the automation of experiments are necessary.

A typical diffraction experiment at a macromolecular crystallography beamline consists of following steps; exchange of samples on the diffractometer, centering the sample in the beam, evaluating of the sample's diffraction to determine the data collection condition, collecting the diffraction data set and processing the data set. In the Photon Factory macromolecular crystallography beamline, we have developed fully automated system by making these steps automatic and organizing them. Especially, in the data processing, we have developed a platform to monitor experiments at the beamline, submit data processing jobs according to experimental and data processing results and store these results in the database. Because of its feasibility, we are applying this platform for automated structure determination for general use.

This presentation is financially supported by JSPS and PDBj.

Membrane protein native state discrimination by a physical energy function

Olga Yuzlenko, Themis Lazaridis

City College of the City University of New York, Department of Chemistry, New York, NY, 10031

An implicit membrane model (IMM1) was tested for its ability to discriminate the native from non-native conformations of membrane proteins. We examined 500 decoys for five different transmembrane proteins generated by the Rosetta-Membrane program. Membrane thickness was found to play a substantial role in discrimination. IMM1 showed the native conformation to have the lowest effective energy for two full length sequences (V-ATPase and bacteriorhodopsin). The correlation coefficient between RMSD and energy of the decoys was 0.4. Van der Waals interactions correlated better with RMSD from the native structures. The energy function was less succesful for the three other proteins. The lowest correlation was observed for lactose permease (ltpA) N-terminal domain (N6). For the transmembrane domain of fumarate reductase (fmr5) and Rhodopsin the native structures had similar energies as the lowest energy decoys. The computed interaction energy between the protein and structural elements missing in the decoys, e.g. the C6 domain in ltpA and two heme molecules in fmr5, was found to be large enough to account for lack of discrimination. This raises the question whether these domains would adopt by themselves the same fold they adopt in the context of the full-length sequence. Experiments in ltpA suggest that the N terminal subdomain may adopt a different fold by itself.

Influence of membrane dipole potential on peptide binding to lipid bilayers

Huan Zhan, Themis Lazaridis

City College of New York/CUNY, Department of Chemistry, New York, NY, 10031

The implicit membrane model IMM1 was extended to include the membrane dipole potential and applied to molecular dynamics simulations of the helical peptides alamethicin, hemagglutinin fusion peptide of influenza (HAFP), HIV fusion peptide (HIVFP), magainin, and the pre-sequence of cytochrome c oxidase subunit IV (p25). The results show that although the orientation of helices on lipid membrane is mainly determined by the distribution of hydrophobic residues, it could be influenced by the dipole potential due to the interaction between the potential and the helix backbone dipole. The results also show that the dipole potential may increase or decrease the peptide binding affinity, through the interaction of the potential with both the backbone dipole and ionic side-chains.

Dephosphorylation Regulation of RNA polymerase II in Eukaryotic Transcription

Yan J Zhang, Mengmeng Zhang, Yonghua Luo, Yong Zhang

University of Texas at Austin, Chemistry and Biochemistry, Austin, TX, 78712

In eukaryotes, the C-terminal domain of RNA polymerase II (CTD) orchestrates the temporal and spatial control of transcription and is involved in the epigenetic regulation of gene expression. Errors in CTD regulation can result in cell death, cancer and severe developmental defects. Two levels of regulation can be achieved by CTD modifications: global transcriptional mRNA processing and gene silencing/activation on an epigenetic level. To understand transcription regulation in eukaryotes, it is essential to comprehend how the CTD is regulated and how that correlates to different stages of transcription. We focus on the molecular mechanism of dephosphorylation of CTD Ser5, and investigate how CTD phosphatases modulate the phosphorylation state of CTD at this site. The regulation and activity of two CTD phosphatases are studied: Scp which epigenetically silences neuronal gene expression and Ssu72 which controls global transcription. We investigated the substrate recognition mechanism of Scp and Ssu72 towards the Ser5 position and provided structural explanation for the precision. Furthermore, the interaction of the phosphatases and their regulatory factors are studied. The first specific inhibitor for Scp is reported as the lead compound to direct neuronal gene expression. Finally, the effects of prolyl-isomerization on the activity of these phosphatases are being investigated.

From NMR information to high accuracy biomolecular structures: A quantum chemical pathway

Yong Zhang

Stevens Institute of Technology, Department of Chemistry, Chemical Biology, and Biomedical Engineering, Hoboken, NJ, 07030

NMR spectroscopic properties are sensitive probes to investigate biomolecules. Deciphering the NMR observable information through high accuracy quantum chemical calculation can provide valuable insights into structural investigations and related drug discovery. Recently, the use of the NMR chemical shift information of an FPPS.IPP.Risedronate complex (FPPS: farnesyl pyrophosphate synthase; IPP: isopentyl pyrophosphate; Risedronate: a drug molecule) together with quantum chemical investigations has assisted in refinement of the x-ray active site structure of this protein complex and revealed unprecedented atomic level diphosphate binding details in the protein. Results also helped the discovery of new drug leads. More recently, our investigation of the NMR and other spectroscopic properties of the myoglobin complex with HNO, a small molecule that has recently been recognized to regulate many biomedical activities, uncovered the first atomic level binding structures, which not only provides the best predictions of experimental spectroscopic measurements, but also offer novel insights into the reactivity and stability results. Quantum chemical investigations of NMR hyperfine shifts have also aided the structural investigations of some paramagnetic protein systems and suggest that the NMR spectroscopic tool may supply sensitive structural information of the binding of redox active metals in the proteins involved in neurodegenerative diseases.

Structural Study of a Molecular Switch Implicated in Quorum Sensing in *Bacillus cereus*

Samira Zouhir[1], Stéphane Perchat[2], Beatriz Guimarães[3], Magali Aumont-Nicaise[4], Didier Lereclus[2], Sylvie Nessler[1]

[1]CNRS, LEBS, Gif-Sur-Yvette, 91198, France, [2]INRA, UGME, Guyancourt, 78285, France, [3]SYNCHROTRON SOLEIL, PROXIMA-I, Saint-Aubin, 91190, France, [4] CNRS, IBBMC, Orsay, 91405, France

By regulating gene expression according to population densities, quorum-sensing allows bacteria to control several processes such as sporulation or virulence. Quorum-sensing usually involves production, secretion and detection of diffusible biomolecules, designated as pheromones, which are recognized by the responder bacteria. In Gram-positive species, like the *Bacillus cereus* group, this communication dominantly relies on the recognition of oligopeptides. Recently, a new family of quorum-sensing effectors named RNPP for Rap, NprR, PlcR and PrgX has been identified. All members contain several TPR motifs allowing interaction between proteins or protein and peptide. Rap proteins are involved in sporulation inhibition. PlcR and PrgX are HTH-type transcription factors regulating the expression of virulence genes and conjugative transfer genes respectively.
The project is focused on the quorum-sensing system: NprR/NprX, where NprR is the protein effector and NprX the signal peptide. NprR displays two functions: in the absence of NprX, NprR regulates sporulation whereas the NprR/NprX complex has been shown to regulate the transcription of 42 genes.
A structural study was undertaken to elucidate the molecular mechanism of this functional switch. NprR is composed of three domains: an N-terminal HTH-type DNA-binding domain, a central domain usually observed in the Rap phosphatases and the C-terminal TPR motifs. The crystal structure of a ΔHTH truncated form of NprR in complex with its peptide NprX-7 has been solved by SAD phasing using selenomethionine labeling. The crystal diffracted at 3.5Å, in space group P1. The cell contained 12 copies of NprR, each one binding one cognate peptide. The NprR/NprX complex forms an unexpected tetramer. Complementary studies in solution, using SAXS and DLS technologies, confirmed the tetrameric state of the complex, and demonstrated that, in the absence of the peptide, NprR is a dimer. Comparison of this new structure with those of the other members of the RNPP family gives us new insights into this original quorum-sensing system.