25TH MBI LAKE ARROWHEAD RESEARCH CONFERENCE

OCTOBER 17-19, 2003

UCLA LAKE ARROWHEAD CONFERENCE CENTER
LAKE ARROWHEAD, CA

25TH MBI Lake Arrowhead Research Conference
October 17-19, 2003

Friday

3:00 **Welcome: Jim Bowie, Conference Chair**

Session I: Genomics I *(Jim Bowie, Chair)*

3:05-3:15 **Michael Strong (Eisenberg lab, MBIDP grad, C&MB trainee)**
Computational and biochemical analysis of the M. tuberculosis genome

3:20-3:30 **Debnath Pal (Eisenberg lab, Postdoc)**
Determining protein function from protein structure

3:35-3:50 **Eri Srivatsan (Surg-Gen Faculty)**
A candidate tumor suppressor gene from chromosome 11q13

4:00 **Check-in – Poster Set-up**

5:30 **Social Hour with hors d'oeuvres**
1st Poster Session (All posters)
[Note: A second poster session will be held Saturday afternoon, 5:30-6:30. All posters will be available for viewing all weekend, and poster awards presented Sunday morning.]

6:30 - 7:30	DINNER

7:30-8:10 SPECIAL LECTURE: **Heather Maynard (Chem & Biochem Faculty)**
Biology, Polymer Chemistry, and the Creation of New Materials

8:15-8:25 *Break*

Session II: Regulation - Signaling *(Jon Braun, Chair)*

8:25-8:35 **Marisa Baron (Bowie lab, Biochem grad, C&MB trainee)**
The Shank SAM domain and the architecture of the post-synaptic density

8:40-8:50 **Parthive Patel (Lengyel lab, MBIDP grad, C&MB trainee)**
Drosophila Rheb GTPase is required for cell cycle progression and cell growth

8:55-9:10 **Arnie Berk (MIMG Faculty)**
How activator-mediator interactions stimulate transcription

9:15-9:25 **Tina Miranda (S. Clarke lab, Biochem grad)**
Determining the role of protein methylation in cell signaling

9:30-9:40 **Nima Gharavi (Berliner lab, Pathology grad)**
OxPAPC induces IL-8 expression in endothelial cells through a novel TLR4-JAK/STAT pathway

9:45-MIDNIGHT	SOCIAL TIME (MOVIE – KARAOKE – DANCING – POKER – ETC.)

Saturday

8:00 - 8:45	BREAKFAST

Session III: RNA/RNAi *(Arnie Berk, Chair)*

9:00-9:10 **Jason Underwood (Black lab, MBIDP grad)**
Regulation of alternative pre-mRNA splicing by FOX proteins

9:15-9:30 **Chris Lee (Chem & Biochem Faculty)**
Alternate splicing in the human, mouse and rat genomes is associated with increased exon creation/loss

9:35-9:45 **Alissa Resch (Lee lab, Biochem grad)**
Assessing the impact of alternative splicing on domain interactions in the human proteome

9:50-10:10 *Break*

10:10-10:20 **Sean Curran (Koehler lab, Biochem grad)**
Defective mitochondria and deaf worms

10:25-10:35 **Gaynor Miller (Crosbie lab, Phys Sci grad)**
Alternate slicing of sarcospan: insights into muscular dystrophy

10:40-10:55 **Richard Gatti (Path & Lab Med Faculty)**
Non-classical splicing mutations in the ATM gene: misinterpreting potential missense mutations

11:00-11:15 **Rachelle Crosbie (Physiol Sci Faculty)**
Specific and potent RNA interference in terminally differentiated myotubes

12:00-1:00	LUNCH
1:00-5:30	*FREE TIME*

5:30 **Social Hour with hors d'oeuvres**
2nd Poster Session (All posters)

6:30-7:30	DINNER

7:30-7:35 **MBIDP Awards Presentation**
Paul D. Boyer Outstanding Teaching Awards
presented by Sabeeha Merchant, MBIDP Chair
Amgen Dissertation Year Awards
presented by awardee mentors, Wen Shi and Jon Braun

7:35-7:45 **Yinuo Li (Shi lab, MBIDP grad)**
Extracellular polysaccharides mediate pilus retraction during social motility of Myxococcus xanthus

7:50-8:00 **Helen Su (Braun lab, MBIDP grad)**
Uncovering the dynamics of T cell trafficking in vivo by positron emission tomography

8:05-8:15 *Break*

Session IV: Proteomics *(Jon Fukuto, Chair)*

8:15-8:30	**Julian Whitelegge (Psychr & Biobehav Sci Faculty)** *Intact protein mass measurements in proteomics*
8:35-8:45	**Eric Bortz (Sun lab, MBIDP grad, C&MB trainee)** *Proteomic ID of gammaherpesvirus virion proteins*
8:50-9:00	**Dan Boutz (Yeates lab, MBIDP grad, C&MB trainee)** *Thermophilic microbes with abundant intracellular protein disulfide bonds: genomic discovery and applications*
9:05-9:15	*Break*
9:15-9:25	**Sarah Yohannan (Bowie lab, Biochem grad, CBI trainee)** *Helix kinking in membrane proteins*
9:30-9:40	**Shilpa Sambashivan (Eisenberg lab, MBIDP grad)** *3D domain swapping zipper model as a model for amyloid fiber formation*
9:45-10:00	**Ken Houk (Chem & Biochem Faculty, PI of CBI Training Program)** *Why enzymes are proficient*

10:05-midnight	**Social Time (Movie – Karaoke – Dancing – Poker – Etc.)**

Sunday

8:00 - 8:45	**BREAKFAST**

9:00-9:05 **Poster Awards Presentation**

Session V: Genomics II *(Wen Shi, Chair)*

9:10-9:20	**Catherine Grasso (Lee lab, Biochem grad)** *Combining partial order alignment and progressive multiple sequence alignment increases alignment speed and scalability to very large alignment problems*
9:25-9:35	**Kirsty Salmon (Gunsalus lab, Postdoc)** *Systems biology approach to understanding E. coli transcription*
9:40-9:50	**Huiying Li (Eisenberg lab, Biochem grad)** *Genome-wide analysis of the protein functional linkages in Rhodopseudomonas palustris*
9:55-10:10	*Break*
10:10-10:25	**Steve Cole (Med-Hemat & Onc Faculty)** *New bioinformatic tools for identifying complex genetic influences*
10:30-10:45	**Yibin Wang (Molec Med/Anesthesiology Faculty)** *Distinctive roles of stress-activated MAP kinases in heart failure*
10:50-11:05	**Mike Teitell (Path & Lab Med Faculty)** *A new type of DNA methylation in mammalian cells*

11:10-11:15 **Closing Remarks: Steve Clarke, MBI Director**

12:00-1:00	**LUNCH – CHECK OUT**

Session Abstracts

(Listed Alphabetically by Speaker's Last Name)

Baron, Marisa
The Shank SAM domain and the architecture of the postsynaptic density

The efficiency of the postsynaptic density (PSD) depends on structural organization because of the massive amount of signal transduction that takes place there. Shank is a master scaffolding protein located in the PSD and contains a C-terminal SAM domain (sterile-a-motif). SAM domains are small protein modules known to polymerize. To understand the function of the Shank SAM domain in the PSD, I solved the crystal structure to 2.1 Å. The crystal structure and electron microscopy show that the Shank SAM domain is able to form a sheet of helical polymers. Mutating single amino acids within or between polymers solublizes the domain, supporting the authenticity of the sheet. Maltose Binding Protein fused to the C-terminal 400 amino acids of Shank can also form a sheet of polymers as shown by EM, demonstrating that the Shank SAM sheet can accommodate the entire Shank protein. Shank SAM contains a zinc binding site composed of two histidines, a glutamate and a chloride ion as shown by the crystal structure bound to zinc. The function of the zinc binding site is currently being investigated. Overall, this suggests that the Shank sheet may provide an architectural framework within the PSD where incoming signals can be organized and transmitted to intracellular proteins.

Berk, Arnie
How activator-mediator interactions stimulate transcription

In vitro and in vivo studies demonstrate that interactions between activation domains and the Sur2 mediator subunit stimulate pre-initiation complex assembly and Pol II initiation. These studies take advantage of sur2-/- murine ES cells in which most activation domains function normally, but adenovirus E1A conserved region 3 activation is completely defective and Elk1 activation in response to MAP kinase signaling is greatly attenuated. Considerable development and differentiation occurs in sur2-/- mouse embryos attesting to the ability of the Sur2-deficient mediator to support complex transcriptional programs. Using tethered template binding assays, we observe that VP16, E1A and ERK2-phosphorylated Elk1 (Elk1-P) activation domains, but not unphosphorylated Elk1 stimulate binding of Pol II and GTFs to promoter DNA. In contrast, only VP16 stimulates Pol II and GTF binding in sur2-/- extracts. In intact cells, serum stimulation of EGR-1 transcription is extremely defective in sur2-/- cells. The EGR-1 control region contains two critical serum response elements where MAP-kinase activated ternary complex factors, including Elk-1, bind cooperatively with SRF dimers, as well as binding sites for other transcription factors. Using ChIP assays, we find that histone H4 acetylation of nucleosomes associated with promoter DNA is not attenuated in sur2-/- cells in response to serum, but binding of Pol II, TFIIE and TFIIB to promoter DNA is diminished, as well as association of Pol II with 3' coding region. Binding of TBP to promoter DNA is only modestly affected. These results indicate that the interaction of E1A and Elk-P with Sur2 stimulates pre-initiation complex assembly as well as a post-recruitment step in Pol II initiation.

Bortz, Eric, Julian Whitelegge, Qingmei Jia, Ting-ting Wu, Z.H. Zhou and Ren Sun
Proteomic identification of gammaherpesvirus virion proteins

Murine gammaherpesvirus-68 (MHV-68) is a model for studying gammaherpesvirus infection and pathogenesis. While the capsid structure of gammaherpesviruses exhibits conservation with other herpesviruses, little is known about the structure and composition of the virion tegument. We have isolated extracellular MHV-68 and identified two morphologically distinct particles: enveloped icosahedral capsids devoid of visible viral DNA and containing only a low-density tegument region and denser enveloped, tegumented nucleocapsids characteristic of herpesvirus virions. We identified conserved and novel virion-associated proteins by liquid chromatography with tandem mass spectrometry (LCMSMS). These include five conserved nucleocapsid protein homologues, the ORF45 and ORF75c tegument protein homologues, and envelope glycoproteins B and H. ORF45 exhibits detergent sensitivity representative of a tegument component. Additional data suggesting the existence of novel virion-associated proteins is presented, including gene products from MHV-68 ORF's 20, 24, 28, 48 and

52, and four cellular proteins. Viral genes encoding virion-associated components are expressed in bacteria and mammalian cells and targeted in the MHV-68 genome for mutagenesis. The identification of conserved MHV-68 proteins previously not predicted to be associated with the gammaherpesvirus virion may provide an indication of the function of these gene products during the early stages of infection or in virion maturation and egress.

Boutz, Daniel R.

Thermophilic microbes with abundant intracellular protein disulfide bonds: genomic discovery and applications

We have recently shown that certain hyperthermophilic microbes contain an abundance of intracellular protein disulfide bonds. Here we provide experimental data and computational analyses to confirm and exploit this finding. Proteomic studies involving 2D gels run under non-reducing followed by reducing conditions have revealed that *Pyrobaculum aerophilum* cells contain large numbers of protein complexes held together by intermolecular disulfide bonds. Proteins of interest will be identified by mass spectrometry to identify potentially novel protein complexes. Various computational studies are also underway to (1) illuminate from a bioinformatics perspective how these cells maintain intracellular disulfide bonds, (2) develop special amino acid substitution tables for disulfide-rich organisms, and (3) improve protein fold prediction methods on the basis of disulfide bond information.

Cole, Steve

New bioinformatic tools for identifying complex genetic influences

Many genetic processes are controlled by combinatorial events (e.g., multiple transcription factors binding to a promoter, multi-hit oncogenesis). In addition, a single phenotype may arise from a variety of distinct combinatorial etiologies (e.g., multiple pathways to a common cancer phenotype). However, the statistical tools used to analyze genetic data are optimized to find unidimensional influences and often fail to identify multiple clusters of combinatorial influence. The causal "entity" in these cases is a disjunctive set of conjunctive events; e.g., cancer arises if gene A is altered AND gene B is altered, OR if gene C is altered AND gene D is altered, OR if gene E is altered AND gene F is altered, but not if any single alteration occurs the absence of others. We have adapted the machine-learning algorithm PRIM to identify such disjunctive sets of conjunctive predictors in data relating genetic characteristics to observable phenotypes. "Conditional PRIM" substantially outperforms conventional linear models and binary recursive partitioning "tree" models in identifying multiple clusters of causal variables (e.g., recovering ~80% of influential variables in representative Monte Carlo studies, vs. 20%-60% for other techniques). This talk presents C-PRIM's advantages in two genetic contexts: 1) detecting combinatorial effects of transcription factors on genome-wide expression profiles, and 2) identifying "latent classes" of complementary genetic alterations underlying oral cancer.

Crosbie, Rachelle

Specific and potent RNA interference in terminally differentiated myotubes

Double-stranded RNA interference (RNAi) is a potent mechanism for sequence-specific silencing of gene expression and represents an invaluable approach for investigating gene function in normal and diseased states, as well as for drug target validation. Here, we report that skeletal muscle myoblasts and terminally-differentiated myotubes are susceptible to RNAi. We employed an approach in which dsRNA is generated by cellular transcription from plasmids containing long (1 kb) inverted DNA repeats of the target gene, rather than using dsRNA synthesized *in vitro*. We show that gene silencing by this method is effective for endogenously expressed genes, as well as for exogenous reporter genes. Analysis of the expression of several endogenous genes and exogenous reporters demonstrates that the silencing effect is specific for the target gene containing sequences within the inverted repeat. Our method eliminates the need to chemically synthesize double-stranded RNA and is not accompanied by global repression of gene expression. Furthermore, we show for the first time that sequence-specific dsRNA-mediated gene silencing is possible in differentiated, multinucleated skeletal muscle myotubes. These findings provide an important molecular tool for the examination of protein function in terminally differentiated muscle cells and provide alternative approaches for generating disease models.

Curran, Sean P. and Carla M. Koehler
Department of Chemistry & Biochemistry

Defective Mitochondria and Deaf Worms

The protein translocation field has long relied on the budding yeast *Saccharomyces cerevisiae* as a model organism. The players, pathways, and mechanisms of mitochondrial protein targeting and translocation are conserved among eukaryotes. *Caenorhabditis elegans* has become an excellent system for both forward and reverse genetic techniques. Missing from this field are strong biochemical methodologies. We were interested in utilizing RNA interference (RNAi) in worms in tandem with the biochemical approach that we have establish in yeast to study protein import in isolated organelles. The Translocase of the Inner Membrane (TIM)9p and Tim10p facilitate the import of polytopic inner membrane proteins and are essential for viability in yeast. We therefore used RNAi to determine if Tim9p and Tim10p could be down-regulated in *C. elegans*. Indeed, feeding of double stranded RNA to a wild type *C. elegans* strain resulted in larval arrest and partial embryonic lethality. To determine if these phenotypes were a specific result of a defect in mitochondrial biogenesis, we have have adapted a mitochondrial isolation procedure from yeast to isolate mitochondria that are competent for in vitro import assays. Mitochondria purified from a large scale gene knockdown of the worm homologues of TIM9 and TIM10 resulted in specific defects in the import of mitochondrial inner membrane substrates. This is the first evidence that genetics and biochemistry can be used together in *C. elegans* to answer fundamental questions in cell biology.

Gatti, Richard A., Laura Eng, Gabriela Coutinho, Shareef Nahas
Department of Pathology and Laboratory Medicine

Non-classical splicing mutations in the atm gene: misinterpreting potential missense mutations

Ataxia-telangiectasia is an autosomal recessive neurological disorder caused by mutations in the ATM gene. In this report, we study ten A-T patients who have non-classical splicing mutations. These mutations fall into three types of categories: pseudo-exon insertions (Type II), single nucleotide changes within the exon (Type III), and mutations that disrupt or delete the conserved 3' AG splice motif (Type IV). Pseudo-exon insertions have been previously described by McConville et al. (1996) and Pagani et al. 2002. Single nucleotide changes within exons (Type III) can misinterpreted as missense mutations; we observed three examples of apparent missense mutations that actually caused aberrant splicing and the partial deletion of an exon. Mutations that disrupt the 3' splice site motif (Type IV) can be misinterpreted as classical splicing mutations, i.e., classical skipping of an entire exon (Type I); however Type IV mutations disrupt a splice motif and use another AG or GT that is located nearby within the exon, to form a new splice site. These mutations lead to partial deletion at the beginning of exons, creating frameshifts that eventually result in premature termination codons. Without screening cDNA, it is not possible to interpret the consequences of these mutations. This may lead to further misinterpretation of genotype-phenotype correlations.

Gharavi, Nima [1,2], Michael Yeh[1,2], Kimberly A. Walton[2], and Judith A. Berliner[1,2]
Department of Medicine[1]; Department of Pathology[2]

Ox-PAPC induces IL-8 expression in endothelial cells through a TLR4-JAK/STAT pathway.

Oxidized-1-palmitoyl-2-arachidonoyl-sn-glycero-2-phosphorylcholine (Ox-PAPC), a component of minimally modified low-density lipoprotein (mm-LDL) found in atherosclerotic lesions and other sites of chronic inflammation, induces endothelial cells to synthesize chemotactic factors, such as interleukin 8 (IL-8). We have previously shown that Ox-PAPC induction of IL-8 is mediated through a Toll-like Receptor 4 (TLR4) pathway. While most TLR4 ligands activate NFκ-B, the Ox-PAPC/TLR4 pathway is independent of NFκ-B, suggesting that an alternative TLR4 signaling mechanism may be involved. In this study, we provide evidence for the role of the Janus Activated Kinase (JAK)/Signal Transducer Activators of Transcription (STAT) pathway in Ox-PAPC-mediated TLR4 signaling. Ox-PAPC treatment of human aortic endothelial cells (HAEC) and human microvascular endothelial cells (HMEC) leads to the rapid phosphorylation and activation of JAK2. Treatment of these cells with AG-490, a JAK inhibitor, significantly inhibits IL-8 production by Ox-PAPC. Ox-PAPC treatment also causes the phosphorylation and activation of STAT3, a known downstream effector of JAK2. Previously, we

have shown that a mutation in the STAT3 binding element found within the IL-8 promoter significantly inhibits Ox-PAPC-induced IL-8 synthesis. In this study, we show that STAT3 phosphorylation induced by Ox-PAPC is subsequently followed by STAT3 translocation into the nucleus, suggesting that STAT3 may play a direct role in IL-8 synthesis. Finally, we have found that TLR4 associates with JAK2, as measured by immunoprecipitation. These results support our hypothesis that Ox-PAPC-induced IL-8 production is mediated through a TLR4-JAK/STAT pathway.

Grasso, Catherine

Combining partial order alignment and progressive multiple sequence alignment increases alignment speed and scalability to very large alignment problems

Partial Order Alignment (POA) has been proposed as a new approach to multiple sequence alignment, which can be combined with existing methods such as Progressive Alignment. This is important for addressing problems both in the original version of POA (such as order sensitivity) and in standard progressive alignment programs (such as information loss in complex alignments, especially surrounding gap regions). We have developed a new Partial Order ñ Partial Order alignment algorithm that optimally aligns a pair of multiple sequence alignments, and which therefore can be applied directly to Progressive Alignment methods such as CLUSTAL. Using this algorithm, we show the combined Progressive POA alignment method yields results comparable to the best available multiple sequence alignment programs (CLUSTALW, DIALIGN2, T-COFFEE), but is far faster. For example, depending on the level of sequence similarity, aligning 1,000 sequences, each 500 amino acids long, took 15 minutes (at 90% average identity) to 44 minutes (at 30% identity) on a standard PC. For large alignments, Progressive POA was 10 to 30 times faster than the fastest of the three previous methods (CLUSTALW). These data suggest that POA-based methods can scale to much larger alignment problems than possible for previous methods. The POA source code is available at http://www.bioinformatics.ucla.edu/poa.

Houk, K. N. and Xiyun Zhang
Department of Chemistry and Biochemistry

Why Enzymes Are Proficient

Noncovalent host-guest binding affinities are usually in the range of 10^3-10^6 M^{-1}. Protein- ligand binding affinities are similar in magnitude, but can be tuned to > 10^9 M^{-1} in potent reversible enzyme inhibitors. The binding of transition states by enzymes, as measured by Wolfenden's proficiency, $k_{cat}/K_M/k_{uncat}$, is generally much larger, commonly exceeding 10^8 M^{-1} and achieving 10^{27} M^{-1} in some cases. Many explanations have been offered, such as electrostatic stabilization, ground state destabilization and desolvation, entropy traps, orbital steering, low barrier hydrogen bonds, covalent bonding with cofactors, tunneling, the "Circe effect", induced fit, dynamic coupling of protein fluctuations to motions in the transition state, reduction of reorganization energy by binding in near attack conformations (NACs), noncovalent cooperativity, and the spatiotemporal hypothesis.

We propose that proficient enzymes, those with $k_{cat}/K_M/k_{uncat}$ > 10^{11} M^{-1}, achieve 15-38 kcal/mol of transition state binding not by a concatenation of noncovalent effects, but by covalent bond formation between enzyme and transition state, involving a change in mechanism from that in aqueous solution. These interactions involve covalently bound transition states and intermediates, general acid/base catalysis, metal-substrate binding, and covalent binding to cofactors. A survey of established enzyme mechanisms for proficient enzymes is consistent with this conclusion. Estimations of uncatalyzed rates in water have extended proficiencies to over 1000 examples for which proficiencies are now known to within several orders of magnitude.

Lee, Christopher

Alternative Splicing in the Human, Mouse and Rat Genomes is Associated with Increased Exon Creation / Loss

One of the most interesting opportunities in comparative genomics is to compare not just genome sequences, but additional phenomena such as alternative splicing, using orthologous genes in different genomes to find similarities and differences between organisms 1,2. Recently, genomics studies have suggested that 40-60% of human genes are alternatively spliced 3-8, and have catalogued up to 30,000 alternative splice relationships in human genes 9. Here we report an analysis of 9,434 orthologous genes in human and mouse, which indicates that alternative splicing is associated with a large increase in recent exon creation / loss. Whereas most exons in the mouse and human genomes are strongly conserved in both genomes, exons that are only included in alternative splice forms (as opposed to the constitutive or major transcript form) are mostly not conserved, and thus are the product of recent exon creation / loss events. A similar comparison of orthologous exons in rat and human validates this pattern. While this says nothing about the very complex question of adaptive benefit, it does reveal that alternative splicing in these genomes has been associated with increased evolutionary change.

Li, Huiying, Matteo Pellegrini, David Eisenberg
Departments of Chemistry & Biochemistry and Biological Chemistry

Genome-wide analysis of the protein functional linkages in Rhodopseudomonas palustris

Understanding cellular metabolisms through the interactions of proteins is important for life science research and applications. The availability of over 100 fully sequenced genomes enables us to infer the protein functional linkages by analyzing the correlations in protein evolution across various species. Here we use *Rhodopseudomonas palustris*, one of the most metabolically versatile bacteria ever discovered, as the model organism to study protein-protein interactions and the regulations of various metabolic pathways. Computational methods, including phylogenetic profile, rosetta stone, gene neighbor and gene cluster method, have been applied to calculate genome-wide protein functional linkages. Functions of unknown proteins, interacting protein complexes and proteins functioning in the same pathways have been predicted. The functional linkage map of the genome also reveals putative new protein pathways. The multiple paralogs of protein complexes and pathways of same functions in the genome of *Rhodopseudomonas palustris* revealed by the functional linkage analysis may explain the metabolic versatility and the adaptability of the organism to various environments.

Li[1], **Yinuo**, Hong Sun[1], Xiaoyuan Ma[1], Ann Lu[1], Renate Lux[1], David Zusman[2], and Wenyuan Shi[1]
[1]Molecular Biology Institute and School of Dentistry
[2]Department of Molecular and Cell Biology, University of California, Berkeley

Extracellular polysaccharides mediate pilus retraction during Social motility of *Myxococcus xanthus*

Myxococcus xanthus is a Gram-negative bacterium with a complex life cycle that includes vegetative swarming and fruiting-body formation. Social (S)-motility (coordinated movement of large cell groups) requires both type IV pili and fibrils (extracellular matrix material consisting of polysaccharides and protein). Little is known about the role of this extracellular matrix, or fibril material, in pilus-dependent motility. In this study, mutants lacking fibril material and, therefore, S-motility were found to be hyperpiliated. We demonstrated that addition of fibril material resulted in pilus retraction and rescued this phenotype. The fibril material was further examined to determine the component(s) that were responsible for triggering pilus retraction. Protein-free fibril material was found to be highly active in correcting hyperpiliation. However, the amine sugars present in hydrolyzed fibril material, e.g., glucosamine and N-acetylglucosamine (GlcNAc) had no effect on fibril(-) mutants, but, interestingly, cause hyperpiliation in wild-type cells. In contrast, chitin, a natural GlcNAc polymer, was found to restore pilus retraction in hyperpiliated mutants, indicating that a polysaccharide containing amine sugars is likely required for pilus retraction. These data suggest that the interaction of type IV pili with amine-containing polysaccharides on cell and slime-trail surfaces may trigger pilus retraction, resulting in S-motility and slime-trailing behaviors.

Maynard, Heather D.
Department of Chemistry and Biochemistry and California NanoSystems Institute

Biology, Polymer Chemistry and the Creation of New Materials

Polymers that are analogs of natural macromolecules should be useful in drug delivery, nanotechnology, and sensors. Using controlled/living radical polymerization, we have developed general strategies to provide rapid access to polymers of this description. For example, universal block copolymer scaffolds with sequences of orthogonally reactive groups have been prepared. With this approach, many biopolymers with diverse functionality could be generated from a single polymer precursor, without the need to establish monomer and polymer synthesis conditions each time. In another example, polymers with end groups that react efficiently with biomolecules such as proteins have been synthesized. Synthetic strategies, characterization, and applications will be discussed.

Miller, Gaynor and Rachelle H Crosbie
Department of Physiological Science

Alternate splicing of sarcospan: Insights into muscular dystrophy

Sarcospan (SSPN), a 25kDa protein named for its multiple SARCOlemmal SPANning domains is an integral component of the dystrophin-glycoprotein complex (DGC) and is highly expressed in cardiac and skeletal muscles. The DGC is a structural complex that spans the muscle plasma membrane (sarcolemma) and links the extracellular matrix to the intracellular cytoskeleton. This structural linkage is critical for normal muscle function as clearly demonstrated by the many forms of muscular dystrophy that result from mutations within the DGC. Sarcospan has four transmembrane domains and is related to the tetraspanin superfamily of proteins. Northern blot analysis strongly suggests that sarcospan undergoes extensive alternate splicing.

I have recently identified a novel isoform of sarcospan from a human skeletal muscle cDNA library. The open reading frame is predicted to encode a protein of 15 kDa containing two transmembrane domains identical to the first two of the 25 kDa isoform. We have named this protein microspan (μSPN). μSPN is enriched in rabbit muscle membranes and indirect immunofluorescence demonstrates that μSPN localizes to the sarcolemma and the T-tubule. Further biochemical analysis of μSPN is required to confirm its membrane distribution and to ascertain whether μSPN is perturbed in muscle disease. Primary mutations within the sarcospan gene have not been identified in known cases of muscular dystrophy. However, these searches have only examined one isoform of sarcospan. Isolation and characterization of SSPN-related transcripts will enable the search for primary mutations in muscle disease to be broadened and may lead to the identification of novel muscular dystrophies.

Miranda, Tina

Determining the role of protein methylation in cell signaling

Protein arginine methylation is an apparently irreversible modification in which the guanidinium group of arginyl residues becomes methylated in a reaction with *S*-adenosylmethionine (AdoMet) as the methyl donor. Currently the mammalian protein methyltransferase family consists of six gene products. These six gene products not only function in the regulation of transcription and RNA transport, but also plays a role in signal transduction. Examples of this include the interaction of PRMT1 with ILF3 and the intracytoplasmic domain of the IFNAR1 chain in the type I interferon receptor and the interaction of PRMT5 with JAK2. TIS (transiently-induced sequences) 21 has also been shown to interact with PRMT1 and to modulate its activity. There are also signaling proteins such as the high molecular weight fibroblast growth factor-2 and STAT 1 that have been shown to be methylated at arginine residues. In collaboration with Dr. Newsham at the Henry Ford Hospital,we have recently found a new role for arginine methylation in cell signaling. Using the yeast two hybrid system we found DAL-1 (Differentially Expressed in Adenocarcinoma of the Lung)/4.1B to interact with PRMT3. DAL-1 is a tumor suppressor gene on human chromosome 18p11.3 whose expression is lost in > 50% of primary non-small cell lung carcinomas. DAL-1/4.1B is not itself a substrate for PRMT3-mediated methylation but its presence inhibits the methylation of a glycine-rich, arginine substrate sequence, GST-GAR, which contains 14 "RGG" consensus methylation sites. These findings suggest that modulation of post-translational methylation may be a novel mechanism through which DAL-1/4.1B affects tumor cell growth.

Pal, Debnath and David Eisenberg
UCLA-DOE Institute for Genomics and Proteomics

Determining Protein Function from Protein Structure

Structural genomics have brought in a curious situation where protein structures are being determined, with no information yet available on their functions. A quick glance at the Protein Data Bank gives us an emerging picture of the difficult challenges ahead. We have implemented a computational method for assigning function to structures, using the controlled vocabulary of the Gene Ontology dictionary. Useful clues derived from the similarity of features of the query protein to a database of known protein features are used to estimate the likelihood of a function (based on that feature). The useful features include: the fold of the protein domain, the sequence of amino acids, sequence motifs, and functional linkages computationally inferred to other proteins of the organism. A Bayesian scoring mechanism is used to update likelihood of the functions obtained from the various query-protein features. The final output is an updated set of weighted functions. This method can work not only with known three-dimensional structures, but also with sequence in conjunction with a fold-recognition algorithm.

Patel, Parthive H.

Drosophila Rheb GTPase is required for cell cycle progression and cell growth

Precise body and organ sizes in the adult animal are ensured by a variety of signaling pathways. In a screen to identify genes affecting hindgut morphogenesis in Drosophila, we identified a P-element insertion in dRheb, a novel, highly conserved member of the Ras superfamily of G-proteins. Overexpression of dRheb in the developing fly (using the GAL4:UAS system) causes dramatic overgrowth of multiple tissues; in the wing this is due to an increase in cell size; dRheb overexpression in cultured cells results in accumulation of cells in S phase and an increase in cell size. Using a loss-of-function mutation we show that dRheb is required in the whole organism for viability (growth), and for growth of individual cells. Inhibition of dRheb activity in cultured cells results in their arrest in G1 and a reduction in size. These data demonstrate that dRheb is required for both cell growth (increase in mass) and cell cycle progression; one explanation for this dual role would be that dRheb promotes cell cycle progression by affecting cell growth. Consistent with this interpretation, we find that flies with reduced dRheb activity are hypersensitive to rapamycin, an inhibitor of the growth regulator TOR. In cultured cells, the effect of overexpressing dRheb was blocked by the addition of rapamycin. These results imply that dRheb is involved in TOR signaling.

Resch, Alissa M.

Assessing the Impact of Alternative Splicing on Domain Interactions in the Human Proteome

We have constructed a database of alternatively-spliced protein forms (ASP), consisting of 13,384 protein isoform sequences of 4,422 human genes (www.bioinformatics.ucla.edu/ASP). We identified fifty protein domain types that were selectively removed by alternative splicing at much higher frequencies than average (p-value < 0.01). These include many well-known protein-interaction domains (e.g. KRAB; ankyrin repeats; Kelch) including some that have been previously shown to be regulated functionally by alternative splicing (e.g. collagen domain). We present a number of novel examples (Kruppel transcription factors; Pbx2; Enc1) from the ASP database, illustrating how this pattern of alternative splicing changes the structure of a biological pathway, by redirecting protein interaction networks at key switch points. Our bioinformatics analysis indicates that a major impact of alternative splicing is removal of protein-protein interaction domains that mediate key linkages in protein interaction networks. ASP expands the available dataset of alternatively spliced protein forms from 967 human genes (SwissProt release 40) to 5,021 human genes (non-redundant set, ASP + SwissProt), a five-fold increase.

Salmon, Kirsty
Systems biology approach to understanding E. coli transcription

A long-term goal of systems biology is the complete elucidation of the gene regulatory networks of a living organism. Today, with the availability of complete genome sequences and new genomic technologies, this goal is conceivable. Indeed, one long-term goal of DNA microarray technology is to infer and model regulatory networks on a global scale, beginning with specific networks all the way up to the complete circuitry of the cell. The work presented here is a first step towards that goal. We have employed DNA microarray technology to identify genes involved in the regulatory networks that facilitate the transition of *E. coli* cells from an aerobic to an anaerobic growth state. We also report the identification of a subset of these genes that are regulated by a global regulatory protein for anaerobic metabolism, FNR. Using this technology, we identified 2,820 genes that exhibit detectable expression levels in the presence and absence of oxygen, and 2,402 genes that exhibit detectable expression levels in wildtype or *fnr* cells grown anaerobically in minimal glucose medium. Analysis of these data demonstrated that the expression of over one-third of the genes expressed during growth under aerobic conditions are altered when *E. coli* cells transition to an anaerobic growth state, and that the expression of 712 (49%) of these genes are either directly or indirectly modulated by FNR.

Sambashivan, Shilpa and Yanshun Liu
3D domain swapping zipper model as a model for amyloid fiber formation

Amyloids can be described as insoluble protein aggregates of normally soluble proteins. Amyloidogenic proteins lack apparent sequence or structural homology, but a majority of these proteins has been implicated in fiber formation associated with the pathology of the disease state. Some of these amyloidogenic proteins are also characterized by expanded glutamine repeats, the most widely cited example being the polyQ expansion in the huntingtin protein from Huntington's Disease. Further, there are at least two instances in literature of proteins that are amyloidogenic and form domain swapped oligomers. Thus, it is highly likely that domain swapping might be one of the mechanisms of amyloid fiber formation. In order to address this question we use bovine pancreatic ribonuclease (RNaseA) as a model system. RNaseA is known to form two types of domain swapped dimers; the minor dimer- that is formed by exchanging the N-terminal -helix and the major dimer formed by exchanging the C-terminal -strand. Insertion of glutamine residues in the hinge-loop region of the N-terminus or C-terminus of RNaseA leads to fiber formation. Here we propose a 3D domain-swapping zipper model as a mechanism for formation of fibers in RNaseA. We believe that the fibers in RNaseA are stabilized from a combination of beta-sheet formation and domain swapping and ongoing biochemical assays seem to favor this model. We have also shown that the S-peptide (residues1-20) of RNaseA forms fibers and tubes that have a high beta-sheet content. We are trying to understand the mechanism of fiber formation in the S-peptide (residues 1-20) of RNaseA. As part of this endeavor we are in the process of crystallizing S-peptide residues1-14, 15-20 and 1-20 in order to obtain more structural detail regarding the mechanism of fibril formation.

Srivatsan, Eri S., Rita Chakrabarti, Kayvan Zainabadi, Veena S. Mysore and Bharathi Ramalingam
A candidate tumor suppressor gene from chromosome 11q13

Cervical cancer affects an estimated 500,000 women worldwide and 13,000 in the United States annually. In earlier studies we demonstrated that chromosome 11 contained a cervical cancer tumor suppressor gene. The mono-chromosome transfer method localized the suppressor gene to 11pter>11q23, which was then sub-localized to11q13. Microsatellite mapping of primary tumors using a number of 11q13 markers including those isolated from our laboratory mapped a minimal deletion to a 300kb interval of chromosome 11q13. These studies also identified homozygous deletion (loss of both allelic copies) of 11q13 sequences in a primary tumor, in a cervical carcinoma cell line, HeLa, and in four of the HeLa X fibroblast derived tumorigenic hybrids. The homozygous deletion was mapped to a 5.7kb sequence within the minimal tumor deletion. Genescan/Grail analysis of 400Kb genomic sequence overlapping the deletion interval showed at least 11 known genes. Of these, PACS1 (phosphofurin acidic cluster sorting protein-1) gene coding for a 114KD protein and spanning the homozygous deletion, shows the presence of an altered 9.0kb transcript in addition to the normal 4.5kb

transcript in the tumorigenic cell lines. Western blotting and immunofluorescence studies show a differential pattern of PACS1 expression in the tumorigenic cells in comparison to the fibroblast cell lines. PACS1 is involved in the transport of proteins through trans golgi network (TNG). Thus, if functional studies prove PACS1 to be the chromosome 11q13 tumor suppressor gene, then this will be a new mechanism of tumor suppression in normal human cells.

Strong, Michael, Parag Mallick, Tom Graeber, Matteo Pellegrini, Michael Thompson, and David Eisenberg.
Computational and biochemical analysis of the Mycobacterium tuberculosis genome

Utilizing the complete genome sequence of the pathogenic bacterium, *Mycobacterium tuberculosis*, we have applied a number of computational methods to identify functionally linked genes and proteins throughout the *M. tuberculosis* genome. We have applied a combination of the Rosetta Stone, Phylogenetic Profile, conserved Gene Neighbor, and Operon methods to identify potential operon organization on a genome-wide scale, as well as to aid in the inference of protein function for a number of previously uncharacterized proteins. Building on these computational methods, we have also developed a novel method for the visualization of genome-wide functional linkages among proteins, and have used this method to create "genome-wide functional linkage maps". These "maps" have enabled us to investigate functional linkages on a genome-wide scale, and have accelerated our understanding of protein connectivity in relation to genome organization. We have also applied methods of hierarchical clustering to cluster these "maps" based on the similarity of functional linkage profiles, which has aided in the identification of novel pathways and complexes in this deadly organism.
We have also developed a novel co-expression assay to identify protein- protein interactions, and have benchmarked this method by investigating the PE and PPE proteins of *M. tuberculosis.* The genome organization of the PE and PPE genes suggested possible operon organization, and we have demonstrated at the biochemical level a novel heterodimer formed between members of the PE and PPE protein families. Together, we demonstrate the combination of biochemical and computational methods to investigate a microbial genome on a genome-wide scale.

Teitell, Mike
A new type of DNA methylation in mammalian cells

Epigenetic regulation of mammalian gene expression encompasses multidimensional changes of the chormatin template to provide or limit access to specific DNA regulatory sequences by transacting factors. The most widely studied epigenetic modification is DNA methylation, which was originally described in mammals in 1948. Since that time most reports have focused on methylation of carbon-5 of the cytosine base in the canonical dinucleotide sequence *CG*. However, some reports have also alluded to non-*CG* methylation, particularly in early embryonic development.

Previously, our group described methylation of the internal cytosine in the canonical sequence *CCWGG* (W = A or T) in the B29 (IgB, CD79b) gene in mature B cell lymphoma cell lines. In corroborating studies, Agirre and colleagues have now described *CCWGG* methylation of the p53 tumor suppressor gene in acute lymphoblastic leukemia patient samples. In both studies, increased *CCWGG* methylation was associated with decreased gene expression. However, the extent of this novel type of DNA methylation and its physiologic relevance has not been examined and it is completely unclear if it will have a major role in regulating gene expression or additional aspects of chromatin configuration. To begin addressing these issues, we have performed a new technique called restriction landmark genome scanning (RLGS) to determine the extent of *CCWGG* metylation in lymphoma cell lines. In preliminary studies, we find *CCWGG* methylation is widely present in Jurkat T cells, particularly in *CG* hypomethylated "*CG* islands". Work in this novel area of epigenetic regulation and its implications in mammalian development and cancer will be briefly summarized here.

Underwood, Jason
Regulation of alternative pre-mRNA splicing by Fox proteins

We are interested in the RNA element, UGCAUG, which acts as a pre-mRNA splicing regulatory sequence when present in the introns adjacent to alternatively spliced exons. This element has been shown to enhance the splicing of alternative exons in a number of model systems. Recently, a mammalian protein family was identified whose members bind to UGCAUG and show high homology to the *C.* elegans sex determination factor, fox-1. One human fox family member (hfox-1) is expressed only in muscle and brain. Through EST and genome mining and subsequent RT-PCR validation, we found that the fox pre-mRNAs are subject to many alternative splicing events themselves. We have shown through in vivo overexpression experiments in non-neural cell lines that hfox-1 can act through an intronic UGCAUG element to promote use of a test alternative exon. In particular, it can activate inclusion of the c-src N1 exon which is normally included only in neural cell lines. Through biochemical experiments, we find that fox proteins bind and gel shift UGCAUG-containing RNA probes and that fox proteins can be UV crosslinked to the UGCAUG element within a pre-mRNA intron. Both assays show that fox binding occurs in a sequence-specific manner. Taken together, these data suggest that fox proteins act as modulators of splice site choice through direct and sequence specific binding to intronic RNA sequences.

Brian Petrich, Manxiang Li and **Yibin Wang**
Division of Molecular Medicine, Department of Anesthesiology
Distinctive Roles of Stress-activated MAP Kinases in Heart Failure

Stress-activated MAP kinases, including JNK and p38, are activated in diseased hearts and are thought to play important roles in heart failure. We generated a number of transgenic mouse lines with targeted expression of upstream activators for each pathway in ventricular myocardium using a cre-loxP mediated gene switch system. Activation of p38, by either MKK3 or MKK6, leads to restrictive cardiomyopathy associated with loss of contractility and myocardium remodeling. MKK3 expression results in heterogenous myocyte drop-out and chamber dilation, while MKK6 results in myocyte hypertrophy and preservation of end-diastolic chamber size. At the cellular level, p38 activity is associated with repressed contractility without changes in intracellular calcium signaling. In contrast, activation of JNK activity by MKK7 induces lethal cardiomyopathy associated with loss of cell-cell gap junctions and dramatic repression of Cx43 expression. Furthermore, we have developed mouse models employing a tamoxifen inducible Cre to achieve temporally controlled transgene expression. Preliminary analysis showed efficient gene activation upon tamoxifen induction that will allow us to investigate primary and secondary effects of stress-activated MAP kinases in the pathogenesis of heart failure. In short, our data suggests that p38 and JNK mediate distinct stress-activated signaling pathways and may play specific roles in cardiac remodeling, and the inducible cre-loxP mediated gene-switch transgenic system represents an effective approach to achieve targeted and temporally regulated genetic manipulation for physiological and genomic studies in the intact heart.

Whitelegge, Julian
Intact protein mass measurements in proteomics

A well-resolved mass spectrum of a protein defines the native covalent state of the corresponding gene's product, and associated heterogeneity, and is thus a very useful measurement in proteomics. Over the past few years we have developed a suite of techniques for analysis of intact integral membrane proteins using electrospray-ionization mass spectrometry. The reverse-phase and size-exclusion liquid chromatography mass spectrometry systems (LC-MS) employed provide sufficient resolution to analyze samples of moderate complexity and range of abundance. This has been illustrated in a study that demonstrated full subunit coverage of the cytochrome b_6f complex from photosynthetic membranes (Whitelegge et al, 2002, Mol. Cell. Proteomics 1, 816-827). The intact masses that were measured for cytochrome *b* led us to report that this four trans-membrane helix protein was modified with a heme molecule, a conclusion that was recently confirmed by a high-resolution X-ray crystal structure of the complex. To extend the chromatographic technologies toward full membrane protein coverage in the context of a full proteomics project, a second dimension of chromatography has been applied prior to our aqueous/organic LC-MS systems. This 2D chromatography system will be

compared to 2D-gel technology both in terms of proteome coverage and fidelity of the analytical procedures toward preserving the native covalent state of proteins after extraction from cells. Fourier-transform mass spectrometry provides ultimate accuracy and resolution and thus the opportunity to perform 'top-down' proteomics. This technique may be the only way to view labile aldehyde modifications of proteins such as the retinal chromophore of bacteriorhodopsin, for example.

Yohannan, Sarah
Helix Kinking in Membrane Proteins

We are studying the effects of proline kinks in helices on stability and structure in membrane proteins, and have developed a method to predict kinks occurring at non-proline residues. There is a much higher frequency of kinks in transmembrane helices as compared to soluble proteins. Many of these kinks occur at proline residues, which distort regular helix geometry by preventing the intrahelical backbone hydrogen bond. We observe, however, that the three prolines in bacteriorhodopsin transmembrane helices can be changed to alanine with little structural consequences. This finding leads to a conundrum: if proline is not required for helix bending, why are prolines commonly present at bends in transmembrane helices? We propose an evolutionary hypothesis in which a mutation to proline initially induces the kink. The resulting packing defects are later repaired by further mutation, thereby locking the kink in the structure. Thus, most prolines in extant proteins can be removed without major structural consequences. We further propose that non-proline kinks are places where vestigial prolines were later removed during evolution. Consistent with this hypothesis, at 13 of 15 non-proline kinks in membrane proteins of known structure, we find prolines in homologous sequences. Our analysis allows us to predict kink positions with >94% reliability.

To further test our evolutionary hypothesis, we have examined ho easily proline is accommodated in transmembrane helices. While none of the mutants is as stable as wild type, many result in only modest reductions in stability. We are currently solving several crystal structures of these mutants to determine how the protein accommodates proline residues.

Poster Abstracts

(Listed Alphabetically by Retreat Participant's Last Name)

Heme oxygenase-1 overexpression inhibits the vascular inflammation of cardiac allograft acute rejection

Jesus A. Araujo[1,2], Lingzhong Meng[1], Aaron Tward[1], Diana Shih[1], Suhasini Iyer[3], Ronald Buelow[3], Jerzy Kupiec-Weglinski[1], Aldons J Lusis[1,2].
[1] UCLA School of Medicine, Los Angeles, CA
[2] UCLA Molecular Biology Institute, Los Angeles, CA
[3] Sangstat Medical Corporation, Freemont, CA

Heme Oxygenase-1 (HO-1) is an important enzyme that may play key anti-inflammatory and anti-oxidant roles in the immune cascade triggered by transplantation of MHC-incompatible heart transplants. We hypothesize that HO-1 overexpression would result in inhibition of vascular inflammation and prolonged cardiac allograft survival. We developed HO-1 transgenic (Tg) mice in the C57BL6/J background, using the native rat HO-1 promoter and gene. Transgene expression was demonstrated by RT-PCR analysis. HO-1 protein overexpression was confirmed by Western blotting, ELISA and enzyme activity assays. We have employed Tg mice as either organ donors or recipients and compared to wild type (WT) littermates in a well-defined murine acute rejection cardiac allograft model (C57BL6/J vs BALB/c). As donors, Tg hearts survived 10.5 ± 0.7 days (n=10), which was significantly longer as compared with WT hearts (6.5 ± 0.4 days; n= 6; p=0.0001). As recipients, Tg mice maintained BALB/c heart transplants for 26.8 ± 3.4 days (n=10), which was significantly longer as compared with WT hosts (6.3 ± 0.1 days; n=12; p<0.0001). Higher levels of COHb were noted in Tg mice. Graft survival was inversely associated with vasculitis and inflammatory cell infiltrates. However, only grafts transplanted into Tg hosts showed a paucity of CD4+ T cells and lack of immune activation as judged by absent CD25 expression at day 4 after transplantation. We conclude that HO-1 overexpression results in inhibition of the inflammatory response associated with acute graft rejection and host HO-1 rather than graft HO-1 overexpression leads to prolonged transplant survival and inhibition of alloreactivity.

Analysis of GAS11 in Mammalian Muscle

Janine M Bekker[1], Kent L Hill, Ph.D.[1,2], Rachelle H Crosbie, Ph.D.[1,3]
[1]Molecular Biology Institute, [2]Microbiology, Immunology, and Molecular Genetics, [3]Department of Physiological Sciences, University of California Los Angeles

The mammalian GAS11 gene was first identified as one of several genes up-regulated by growth arrest in cultured murine cells. Subsequently, GAS11 was found to be part of a large family of cytoskeleton-associated proteins present in diverse eukaryotic organisms from protozoa to humans. The best-characterized member of this family is trypanin from *Trypanosoma brucei*. Trypanin links the trypanosome's flagellum to the microtubule cytoskeleton and is required for flagellum based motility. Mammalian GAS11 contains a well-conserved 143 amino acid domain that localizes to the plus ends of microtubules when expressed as a GFP fusion protein in trypanosomes. This same fusion protein co-localizes with cytoplasmic microtubules in cultured mammalian cells suggesting that GAS11 may contain a unique microtubule association domain.

I have recently engineered recombinant fusion proteins representing full length and specific GAS11 truncations. Recombinant proteins have been utilized to determine that GAS11 directly binds microtubules *in vitro*. What's more, it appears that GAS11 self associates, potentially forming a filamentous network *in vitro*.

GAS11 message is ubiquitously expressed in mammalian tissue. Interestingly, it contains two muscle specific transcripts. I have generated polyclonal antibodies to various epitopes of GAS11 and established that multiple GAS11 isoforms are expressed in mammalian muscle tissue. These antibodies will be used for specific biochemical analysis of sub-cellular localization within mammalian muscle.

The cytoskeleton is integral in muscle development and function, therefore, further functional analysis of GAS11 as a novel cytoskeletal protein will provide insight into its importance in mammalian muscle.

Ionizing Radiation and Restriction Enzymes Induce Microhomology-mediated Illegitimate Recombination in Trans in Saccharomyces Cerevisiae

Cecilia Chan, Palaniyandi Manivasakam and Robert H. Schiestl
Molecular Toxicology Interdepartmental Program and
Departments of Pathology, Environmental Health, and Radiation Oncology

DNA double strand breaks are repaired by two pathways, homologous recombination in the presense of sequence homology and nonhomologous or illegitimate recombination in the absence of extended sequence homology. One mechanism of illegitimate recombination occurs between several basepairs of homology, also called microhomology-mediated recombination which is conserved from bacteria to mammalian cells. Here, we report that both ionizing radiation and restriction enzymes induce microhomology-mediated integration in trans at sites that are probably not damaged. Both, irradiation and restriction enzymes increased the frequency of illegitimate integrations. Irradiated yeast cells displayed 83% of microhomology-mediated illegitimate integration, compared to 23% of microhomology usage in spontaneous integration. Restriction enzymes mediated integration events not only at genomic restriction sites, but also at microhomologies at random non-restriction sites. These results suggest that DSBs caused by ionizing radiation and restriction enzymes may induce a genome wide microhomology-mediated illegitimate recombination pathway and facilitate integration at non-targeted sites.

Progress Towards the Total Synthesis of Guanacastepene A

Hiufung Chu, Xiaohui Du, Ohyun Kwon

Guanacastepene A was recently isolated by Jon Clardy and his colleagues from an unidentified fungus CR115 growing on the branch of *Daphnopsis americana* in Guanacaste, Costa Rica. This natural product exhibits antibiotic activity against drug resistant strains of *Staphylococcus aureus* and *Enterococcus faecalis*, which has been a tremendous threat to the human health. Therefore, guanacastepene A is a potential candidate as a new antibacterial agent. However, guanacastepene A also displays hemolytic activity against human red blood cells and thus, has limited its therapeutic application. Analogs of guanacastepene A that possess antibacterial activity without hemolytic effect will be a great value. Our goal is to complete the total synthesis of guanacastepene A and to establish a general synthetic scheme that allows us to elaborate the skeleton framework of guanacastepene A to its analogs.

EPR Iron Homeostasis Studies in Organisms Lacking Cu/Zn Superoxide Dismutase

Matthew H. Clement*, Chandra Srinivasan*, Sailaja Mantha**, Ting-Ting Huang**, Joan S. Valentine*, Edith B. Gralla*
*University of California, Los Angeles, **Stanford University

Previous studies in yeast (*S. cerevisiae*) lacking CuZn superoxide dismutase (*sod1Δ*) have shown elevated Fe(III) levels under aerobic growth conditions. This altered iron homeostasis can be observed by a whole cell electron paramagnetic resonance (EPR) method designed to quantitate EPR detectable iron (high spin, rhombic Fe(III), at g = 4.3) levels. The solvent exposed 4Fe-4S clusters of aconitase are one of the suspected sources of the elevated Fe(III) levels exhibited in *sod1Δ* cells. This 'free' or 'junk iron' liberated from 4Fe-4S clusters is then capable of participating in Fenton-type reactions which can lead to the formation of reactive oxygen species (ROS) causing oxidative stress. In collaboration with Sailaja Mantha and Ting-Ting Huang at Stanford, we have adapted the EPR method to study the Fe(III) levels in organ tissue homogenates of mice lacking *SOD1*. Preliminary studies by Mantha and Huang have investigated changes in iron homeostasis as well as aconitase activity in the knockout mice. Our collaborative study will aid in the elucidation of the source of elevated Fe(III) levels as well as its relationship to ROS production and oxidative stress in organisms lacking *SOD1*.

Papilin: A Versatile Basement Membrane Protein which interacts with ADAMTS Metalloproteases and Fibulin

Yali Chen (1), Andrei Kramerov (1), Irina Kramerova (1), Lynn Sakai (2), Doug Keene (2), Lisa Fessler (1) and **John Fessler** (1).
Molecular Biology Institute UCLA (1), and Shriners' Hospital, Portland OR (2)

The first molecular hall mark of a group of cells developing into an organ is the deposition of a basement membrane immediately adjacent to these cells. Basement membranes are 100 -1000 nm thick networks of specialized proteins. They are attached to epithelial and muscle cells by transmembrane receptors, and are essential for cell differentiation and survival. Our poster concerns the glycoprotein Papilin which we discovered in Drosophila and then in mice. Papilins are multidomain proteins. A characteristic subset of domains, which we call the papilin cassette, is shared with the non-catalytic portion of a group of metalloproteases, the ADAMTS enzymes (A Disintegrin And Metalloprotease with ThromboSpondin motifs). We established that papilins, which are not enzymes, interact with ADAMTS metalloproteases in vitro and are searching for their potential interactions during development in vivo. Papilin occurs in basement membranes throughout Drosophila and mice, is essential for Drosophila embryogenesis, and its overproduction causes aberrations of cell arrangements of Drosophila organs. We are characterizing the molecular interactions of different papilin domains with their basement membrane and receptor neighbors.

Single Molecule Detection of Protein-Protein Interactions

Natalie R. Gassman[1], Achillefs N. Kapanidis[1], Nam Ki Lee[1,*], Ted A. Laurence[1,#], Xiangxu Kong[1], and Shimon Weiss[1,2,3]
[1]Dept. of Chemistry and Biochemistry, UCLA, [2]Dept. of Physiology, David Geffen School of Medicine,UCLA, [3]CNSI, *presently Seoul National University , #presently Lawerence Livermore National Laboratory

Characterization of protein-protein interactions has become of great interest in the post-genomic era, and a variety of methodologies have been developed to assign activity to sets of proteins. Although these methodologies have greatly expanded working knowledge of protein functions and their interacting partners, there is still a large gap between proteins in the genome and their known interactions. By developing optical methods based on fluorescence burst and fluctuation spectroscopies to screen protein-protein interactions; we hope to detect oligomerization in identical macromolecules, as well as interactions between two or more different macromolecules. Detection of oligomerization can occur with a single photon channel where photon bursts emitted by free molecules can be differentiate from complexes, which have slower diffusion times and increased brightness. For protein-protein interaction studies, a greater sensitivity of detection is required and can be achieved by labeling the two different macromolecules with two distinct and separable fluorophores. Single molecule detection of the two fluorophores with distinct emission results in a more sensitive binding assay due to the ability to separate the emissions into two detection channels and resolve subpopulations of complexes, which display simultaneous photon bursts on both channels. These methods are applied to the elucidation of transcription in *Shewanella oneidensis* MR-1. Our lab is pursuing the protein-protein interactions involved in transcription of electron transport proteins by the prokaryotic enhancer binding protein, nitrogen regulatory protein (NtrC).

***Cre recombinase expressing helper dependent adenovirus application* in vivo**

Jose Gil, Oliver Dorigo, Sean Gallaher, Brent Tan, Arnold Berk

Cre recombinase is used to excise genes flanked by the LoxP recognition sequence. This system can be used to produce knock-out mice by crossing mice with the floxed genes with Cre Recombinase expressing mice. A major limitation is in studying developmentally lethal genes. This can be
overcome with the using a Cre expressing vector. We demonstrate a functional Cre Recombinase expressing Helper Dependent Adenovirus(HDA), in vivo. As with all HDA systems, the vector contains no adenoviral coding sequences. It only retains the ITR for DNA replication, and the packaging element. The vector drives Yellow Fluorescence Protein(YFP) as a reporter in addition to Cre recombinase. YFP has two uses. Firstly, it is used in the viral propagation step. The percentage of infected cells at each step can be determined by fluorescent microscopy. This determines when to scale up the number of cells to infect. The second use of YFP is to determine the percentage of tranduced target cells either in vivo or in vitro. Cre recombinase expressing HDA has been shown to be functional in Rosa26 reporter mice. Approximately 0.5-1E10 genomes of vector were injected via tail vein. Rosa26 mice contain a LacZ reporter gene for Cre recombinase. Upon Cre mediated excision of intervening stop codons, tranduced cells will express beta-galactosidase. Three to seven days post-injection, livers were harvested, fixed and stained with X-Gal. A significant proportion (approximately 30-40%) of cells was visibly blue. This result demonstrates the successful recombination event mediated by Cre from the HDA vector.

Patterning of Antibody by Photolithographic Approaches

Fang Gu[1], Chuya Yang[2], Carlo Montemagno[2], Wenyuan Shi[1,3]
[1] School of Dentistry, UCLA [2] School of Engineering [3]Molecular Biology Institute

Dental caries is the most epidemic disease. Caries risk is closely related to salivary levels of cariogenic bacteria such as *S. mutans*. Rapid diagnosis of these bacteria will help to prevent dental caries. On-site analysis is a powerful method for the diagnosis. This can be achieved using protein arrays on biochips. The successful development of protein chips requires a method for protein patterning on a solid surface. In this study, we describe a photolithography method for the surface patterning of antibody on a solid surface. Photoactivatable biotin was used to pattern biotin on a thermal oxide wafer by photolithographic method. Monoclonal antibody against cariogenic bacteria, *S. mutans* and *L. casei*, developed previously was conjugated to biotin. The biotinylated antibodies were immobilized on the wafer through a biotin-streptavidin system. The immobilized antibodies were shown to keep the bioactivity. This was used as a model system to establish the antibody array. It will provide basis for development of MEMS based hand-held diagnostic device for rapid detection of multiple cariogenic bacteria.

Recognition and Specificity in Early Stages of the Establishment of the Rhizobium-Legume Symbiosis

Nancy A. Fujishige, Katherine Jankaew, Christopher Butcher, Neel Kapadia, and **Ann M. Hirsch.** Department of Molecular, Cellular, and Developmental Biology. University of California, Los Angeles. Los Angeles, CA 90095-1606.

Recognition and specificity in the *Rhizobium*-legume symbiosis, in which Gram-negative rhizobia provide fixed-nitrogen to a host-legume, and in return, the legume provides carbon-containing molecules, has been well studied. The symbionts recognize each other by exchanging species-specific molecular signals; flavonoids from the plant and the products of the rhizobial *nod* genes, Nod factor. Plant lectins, carbohydrate-binding proteins that immunolocalize to root hair tips, are also thought to be involved in host specificity. In a previous study of transgenic alfalfa plants carrying either the soybean or pea lectin gene, we found that strains that normally do not nodulate alfalfa would do so, but only if the *nod* genes of the compatible *Sinorhizobium meliloti* were transferred to the heterologous strain. Furthermore, we observed greater attachment of the transconjugent *Bradyrhizobium japonicum* or *Rhizobium leguminosarum* bv. *viciae* to the appropriate transgenic root. Because root colonization is an important early step in the establishment of the symbiosis, we looked at the attachment stage in greater detail. We analyzed rhizobial strains for their ability to form biofilms, structured communities of microbes enclosed in a self-produced polymeric matrix and adherent to an inert or living surface. We developed assay conditions for biofilm formation of *R. leguminosarum* bv. *viciae* and *S. meliloti* on both inert surfaces and living roots. We found that several of the strains used in the study of the transgenic lectin plants were enhanced in biofilm formation. From screening *S. meliloti* mutants that have either symbiotic or non-symbiotic defects, we also found a number of mutants that are impaired in biofilm formation on both plant roots and on inert surfaces. When these genes were transferred to wild-type *S. meliloti*, the strains exhibited enhanced biofilm formation and nodulation.

Simian virus 40 large tumor antigen is unable to transform murine embryonic fibroblasts lacking the mediator subunit sur2

Gregory A. Horwitz[1], Jennitte L. Stevens[2], Arnold J. Berk1[2]
Molecular Biology Institute[1] and Department of Microbiology, Immunology, and Molecular Genetics[2], UCLA

Transcription is controlled by the binding of gene-specific factors that affect transcription by interacting, either directly or indirectly, with components of the general transcription machinery. Sur2, first isolated from *C. elegans*, is a ubiquitous protein that acts late in a receptor-tyrosine kinase-Ras-Raf-MAP kinase signaling pathway. The human Sur2 homologue has been demonstrated to be part of a Mediator complex which stimulates transcription *in vitro* in response to both E1A and VP16 activation domains. Recently, our lab generated *sur2*$^{-/-}$ embryonic stem cells. Absence of Sur2 prevents activation by E1A-conserved region 3 (E1A-CR3) and the mitogen-activated protein kinase-regulated Ets transcription factor Elk1. Since Sur2 is a target of some Ras-MAP kinase pathways and the abnormal activation of Ras-MAP kinase pathways is thought to be involved in oncogenic transformation by many oncogenes, we sought to investigate the effect of the *SUR2* knockout on transformation by various oncogenes. Fibroblasts were prepared form E9.5 littermate embryos and infected with retrovirus vectors expressing SV40 large and small T-antigen, G12V Ha-Ras, or activated Raf. Transduced cells were initially assayed for anchorage-independent growth in soft agar. SV40 T-antigen transformed cells were subsequently also assayed for colony formation on tissue culture plates at low density, ability to form foci on a monolayer of wt MEFs, and tumor formation in nude mice. Fibroblasts lacking the Mediator subunit Sur2 are very greatly reduced in their ability to be transformed by SV40 large T-antigen. However, transformation by expression of constitutively active Ras or Raf in *sur2*$^{-/-}$ fibroblasts was not abrogated.

The crystal structure of an extraordinary, "open pore" ferritin from the hyperthermophilic Archaeon Archaeoglobus fulgidus

Eric Johnson[1], Duilio Cascio[2], Michael Sawaya[2], and Imke Schroeder[1]
[1]University of California at Los Angeles - Department of Microbiology Immunology and Molecular Genetics; [2] University of California at Los Angeles – Department of Energy Laboratory of Structural Biology and Molecular Medicine

Ferritins are important iron storage/detoxification proteins found ubiquitously in living organisms. Despite low sequence homology between species, the 3-dimentional structure of known ferritins is highly conserved. All ferritins previously characterized, are comprised of 24 subunits which assemble into a hollow, largely closed spherical shell with 4-3-2 symmetry. This report details the 2.1 Å resolution structure of the ferritin from the hyperthermophilic Archaeon *Archaeoglobus fulgidus*. The *A. fulgidus* ferritin (AfFTN) monomer has a high degree of structural homology with known prokaryotic and eukaryotic ferritins (r.m.s. deviation of ~ 0.9 Å), but the 24 subunit biological assembly represents the only known ferritin which packs to form a shell with tetrahedral (3-2) rather than octahedral (4-3-2) symmetry. The result of this unique packing is a shell containing 4 large (~ 45 Å) threefold pores arranged in a tetrahedral configuration, and four small threefold hydrophilic pores typical of other ferritins. Here we present the AfFTN structure and assembly in comparison with archetypal ferritins. Finally we offer a hypothesis explaining why the AfFTNs' unique structure is required in order for it to function properly in the physiological context of an anaerobic and hyperthermophilic organism.

Generation and Analysis of the Liver-Specific Farnesoid X-Activated Receptor (FXR) Transgenic Mouse

Florence Y. Lee[1] and Peter A. Edwards[2]
[1]Department of Biological Chemistry, [2]Departments of Biological Chemistry and Medicine, and the Molecular Biology Institute, UCLA

Farnesoid X-activated receptor (FXR) is a member of the nuclear receptor superfamily. Nuclear receptors are ligand-activated transcription factors which, upon binding of ligands, localize to response elements in the promoter region of target genes and regulate the transcription of such genes. Recently, bile acids were identified as the endogenous ligands FXR. In addition, the critical role of FXR in controlling cholesterol and lipid metabolism has also been established. Bile acids are catabolites of cholesterol. Excretion of bile acids comprises the major route for cholesterol elimination from the human body. Bile acids act through FXR to control their own synthesis via a tightly controlled feedback loop. FXR can also regulate very low density lipoprotein (VLDL), low density lipoprotein (LDL), and triglyceride metabolism in mice. In FXR-null mice, feeding of high cholesterol diet leads to elevated hepatic triglyceride levels and serum VLDL and LDL levels. The latter two lipoproteins are associated with increased risk of atherosclerosis. In an effort to study the therapeutic potential of FXR ligands in cholestatic liver diseases and hyperlipidemia, we created transgenic mice with liver-specific expression of a constitutively active form of FXR. Through preliminary studies of these mice, we noted that the transgene is expressed in the liver and some, but not all, previously established FXR target genes are up-regulated. We hope that future studies can shed light on the mechanism through which FXR exerts its effect on cholesterol and lipid metabolism.

Biochemical analysis of sarcospan-- a molecular facilitator in the dystrophin-glycoprotein complex

Karin Lipman, Emily Wang, Gaynor Miller, Rachelle Crosbie

The dystrophin-glycoprotein complex (DGC) provides critical linkage between the extracellular matrix and the intracellular cytoskeleton, maintaining stability during muscle contraction. Mutations in the DGC cause muscular dystrophy—a disease characterized by progressive degeneration of skeletal muscle. As the most recently identified component of the complex, sarcospan is tightly associated with the sarcoglycans subcomplex. Together sarcospan and the sarcoglycans interact with β-dystroglycan to stabilize α-dystroglycan at the muscle cell membrane. Sarcospan is closely related to the tetraspanin family of proteins, which promote lateral associations between integral membrane components in the membrane bilayer. Thus, we hypothesize that sarcospan acts as a molecular facilitator to stabilize protein-protein interactions within and across the sarcolemma, which may be critical to the function of the DGC as a whole. We have engineered 30 site-directed mutants, spanning the entire length of sarcospan, to identify the structural domains critical for these protein-protein interactions. Through expression in cultured mammalian cells, we have shown for the first time that sarcospan self-associates. We further proposed that cysteine residues facilitate thiol-mediated sarcospan oligomerization, and identified specific regions that caused aberrant hyper-oligomerization. When site-directed mutants of sarcospan were co-expressed with the sarcoglycans, we identified specific regions of sarcospan essential for interactions within the sarcoglycan subcomplex. Our research suggests a model whereby several sarcospan molecules may interact to form a large molecular web; mediating associations between the sarcoglycans and dystroglycans, and stabilizing the entire complex as a whole.

Identification of new substrates of adenovirus E1B/e4orf6 complex in the ubiquitination pathway

Yue Liu[1], Andrej Shevchenko[2], Anna Shevchenko[2], Tim Lane[1], Arnold J. Berk[1]
[1]Molecular Biology Institute, UCLA, [2]Max Planck Institute for Molecular Cell Biology and Genetics, Dresden, Germany.

Adenovirus E1B oncoprotein plays an important role in virus infection. Previous studies showed that in Adenovirus infected Hela cells, E1B and E4orf6 form an E3 ubiquitin ligase that targets p53 to degradation. In order to address whether or not E1B/E4orf6 target other cellular proteins besides p53, potential substrates were identified by comparing immunoprecipitated proteins from 293 cells and Ad5 infected 293 cells using an anti-E1B monoclonal antibody. This was followed by a Mass Spectrometry (MS) analysis. Since 293 cells contain high levels of E1B but no E4orf6, the proteins pulled down in 293 cells but disappear or shift in Ad5 infected 293 cells might be the potential substrates. By doing IP and MS, we found that Rad50, Eplin, Mre11, Hsp70, Drebrin E, Vimentin, p53, Myosin 1c, 1d, VI, α tublin, β tublin, β actin, and a hypothetical protein were pulled down by anti-E1B antibody in 293 cells. In 293 cells infected with Adenovirus, p53, Rad50, Mre11 disappeared while all the other proteins were pulled down. RNA helicase p68 was pulled down in Ad5 infected 293 cells but not 293 cells. Therefore, besides p53, Rad50 and Mre11 might be the new targets of E1B/E4orf6 complex mediated protein degradation pathway. Some of the other proteins pulled down in 293 cells might be ubiquitinated but not degradated by the E1B/E4orf6 complex. By doing IP and Western blot, we confirmed that Rad50, Mre11, Eplin β and Vimentin are E1B partners. By doing immunoflourensence assay, we found that E1B, p53, Rad50 and Mre11 co-localized in the same cytoplasmic discrete body in 293 cells. This implies that E1B might recruit p53 and Mre11-Rad50 complex to the same specific cytoplasmic structure to inactivate them and degrade them in cooperation with E4orf6.

Toll-like receptors induce a phagocytic gene program

Ryan O'Connell

Toll-like receptor (TLR) signaling and phagocytosis are hallmarks of macrophage-mediated innate immune responses to bacterial infection. However, the relationship between these two processes is not well established. Our data indicate that TLR ligands specifically promote bacterial phagocytosis, in both murine and human cells, through induction of a phagocytic gene program. Activation of this program is reliant on MyD88/IRAK4/p38-dependent signaling leading to upregulation of scavenger receptors. TLR ligands amplify both the percentage of phagocytes ingesting bacteria and the number of bacteria phagocytosed by individual macrophages. Our data describe a conserved mechanism by which TLRs specifically promote phagocytosis of invading bacteria.

Ribosomal lysine (K) Methyltransferase 23a is a novel SET methyltransferase that dimethylates ribosomal protein L23a in yeast

Tanya Porras-Yakushi, Tina Branscombe Miranda, and Steven G. Clarke

In vivo studies have shown that the ribosomal protein of the large subunit Rpl23a in *Saccharomyces cerevisiae* is methylated at lysine residues (1). However, the gene encoding the methyltransferase responsible for the modification has not been identified (1,2). In this study we have shown that the lysine methyltransferase that modifies Rpl23a, which we refer to as Ribosomal Lysine (K) Methyltransferase 23a (Rkm23a) is encoded by a single gene in the yeast chromosome. Rkm23a was identified as a methyltransferase by *in vivo* labeling strains in which SET domain containing open reading frames had been deleted. When compared to the parent strain, a disappearance in methylation was observed in a 14kDa protein using the rkm23a knock out strain. Amino acid analysis of the band corresponding to the substrate demonstrated that the substrate was dimethylated at lysine residues. *In vitro* methylation experiments using purified ribosomes indicated that the substrate was a ribosomal protein. Finally, trypsin digestion and MALDI analysis confirmed the identity of the protein to be RPL23a. The presence of a SET domain in Rkm23a now implicates SET-containing methyltransferases in translation, in addition to transcription.

References:

(1) Lhoest, J., Lobet, Y., Costers, E. and Colson, C. (1984). Methylated proteins and amino acids in the ribosomes of *Saccharomyces cerevisiae. Eur. J. Biochem.* **141,** 585-590.

(2) Lobet, Y., Lhoest, J. and Colson, C. (1989). Partial purification and characterization of the specific protein-lysine N-methyltransferase of YL32, a yeast ribosomal protein. *Biochimica et Biophysica Acta.* **997,** 224-231.

A Neuronal Scaffolding molecule, Dlgh1, Functions to Construct the Immune Synapse

June Round[1], Tamar Tomassian[1], Viresh Patel[1], Stephen Schoenberger[3], M. Carrie Miceli[1,2]
[1]MIMG [2]MBI [3]La Jolla Institute for Allergy and Immunology

The reorganization of cholesterol rich lipid rafts present in the plasma membrane into a highly ordered and compartmentalized "immunological synapse" is imperative to obtaining productive T cell activation and proliferation. However, the molecular mechanisms regulating synapse formation still remain unclear. This study sets out to elucidate the role of a neuronal synaptic scaffolding molecule, Dlgh1/hDLG, in the formation of the synapse formed when a T cell engages an antigen presenting cell. We reveal that Dlgh1 localizes to the synapse and associates with key signaling molecules, Zap-70 and Wasp. We further show that Dlgh1 binds specifically to the SH3 domain of Lck and this domain is required for proper localization of Dlgh1 in the lipid raft compartment. Furthermore, using an RNAi knockdown approach in transgenic OT-1 T lymphocytes, we demonstrate that loss of Dlgh1 interferes with synapse formation as well as IL-2 and IFN-γ production. Based on these results we propose a model whereby Dlgh1 serves to form the immune synapse , and once localized properly serves to organize and traffic key molecules to ensure sustained TCR signaling and proper T cell activation.

Probing the Protein Partners in Huntington's Disease

Dyna I. Shirasaki, Allan J. Tobin and Joseph A. Loo
Department of Chemistry and Biochemistry, Brain Research Institute, UCLA

Although the function of huntingtin (Htt) currently remains elusive, "guilt by association" with a number of proteins has implicated a possible role in such cellular events as vesicular trafficking, transcriptional regulation, and the endosome/lysosome pathway. While previously discovered interactions have provided valuable insight, it is questionable as to how well they portray actual occurrences within the human brain. In order to obtain a more accurate picture, we are searching for Htt-protein interactions in PC12 cells have been stably transfected with a construct containing exon-1 of the huntingtin gene (containing either 25 or 103 glutamines) fused to a C-terminal EGFP tag. Immunoprecipitated Htt was separated on an SDS-PAGE gel and several bands were excised and the proteins digested with trypsin. Peptide mass fingerprints were obtained by MALDI-TOF mass spectrometry, and peptide sequencing was accomplished by nanoLC-MS/MS, nanoelectrospray-MS/MS and MALDI-MS/MS with a QqTOF-MS and an ion trap mass spectrometer. Preliminary sequence information has identified some of the interacting proteins to be GAPDH, tyrosine hydroxylase and HSP27. Western blots are currently being performed to confirm our results. Proteomics and mass spectrometry offer an alternative means for detecting huntingtin-associated proteins that enables us to search for interactions occurring throughout the cytoplasm as well as in the nucleus. It may also allow for increased detection sensitivity to proteins that may be in low abundance. Revealing the nature of huntingtin's protein partners may provide valuable clues as to the function of huntingtin, and possibly lead to a testable hypothesis for the disease mechanism.

Loss of BMP3 on Inbred Backgrounds Results in Delayed Lung Development

Bernadette So, Matthew Bahamonde, Karen Lyons

Bone morphogenetic proteins (BMPs) are a family of proteins within the TGFβ superfamily of growth factors. Although BMPs were first identified by their capacity to induce bone and cartilage formation, these proteins are involved in many other aspects of development. BMP3 is the most abundant BMP in bone. It is also the major protein found in osteogenin, which is capable of forming bone. On an outbred background, BMP3 null mice are viable and fertile, and show an increased bone density, which suggests that BMP3 acts to negatively regulate bone density. However, the phenotype of $Bmp3^{-/-}$ mice on inbred backgrounds demonstrates an additional role for BMP3 in development. Inbred $Bmp3^{-/-}$ mice appear cyanotic, gasp for air, and die within minutes to hours after birth. These breathing defects are not a result of bone or cartilage defects, as no cleft palate or treacheal defects were seen in histological or skeletal analyses. Histological analyses eliminated a diaphragm defect, but did show that $Bmp3^{-/-}$ lungs appeared developmentally delayed, both in the appearance of the lungs as well as the population of cells present in the lungs. Since BMP3 is highly expressed in the lung, this expression appears to be important for late lung development. The lung defects have been further analyzed at a molecular level using RT-PCR. Initial results support the hypothesis that the absence of BMP3 results in delayed lung development.

Novel Oncolytic Adenovirus with Linked imaging component in Prostate Cancer

Samuel Tahk1 and Lily Wu2.
[1]Molecular Biology Interdepartmental Program, [2]Department of Urology

Despite much research, the understanding of the prostate cancer and effective treatment remains a mystery. Recently, adenoviral vectors (AV) have been engineered to be oncolytic by targeting prostate tumors through prostate-specific antigen (PSA) promoters driving key transactivating early protein E1A. We are developing an oncolytic adenoviral approach that has the expression of E1A and E1B titrated by differing numbers of GAL4 binding sites for efficient lytic replication. We use the novel two-step transcriptional activation approach (TSTA) (Zhang, et al., 2002 Mol. Therapy 5:223) to enhance prostate-specific expression of E1A and E1B proteins bidirectionally. The TSTA system expresses a potent chimeric fusion GAL4-VP16 activator which in turn binds to GAL4 sites located between the promoters of E1A and E1B and activates their expression. Also, we are using a luciferase reporter system under the control of TSTA. The luciferase gene levels are anticipated to increase according to efficient viral replication. The oncolytic potential can then be monitored through light production. Luciferase levels are monitored through a charged coupled device (CCD) camera and correlated with viral replication. Together these systems are an effective tool to lyse prostate tumor cells and monitor effectiveness of treatment.

Transcription Initiation Complexes in Kinetoplastid Protozoa

Sean Thomas (MBI), Michael Yu, Nancy Sturm, and David Campbell (MIMG, MBI)

In kinetoplastids, a 39-nucleotide spliced leader (SL) RNA is *trans*-spliced to the 5' end of nuclear mRNAs before they can be translated. Thus the role of SL in gene expression is central to kinetoplastid biology. The SL RNA genes in *Leishmania* contain promoters with important sites at approximately –60 and –30, and are characterized to single nucleotide resolution. Gel shift assays show a complex being formed specifically on the –60 element. We are pursuing the identification and characterization of SL RNA transcription factors in detail by biochemical and genetic methods. We have cloned the genes for two potential homologs of transcription factors from *Leishmania*. One of these proteins is the global transcription factor TATA Binding Protein (TBP). Inclusion of anti-TBP antiserum in the gel shift assay resulted in a supershifted band suggesting that TBP interacts with the –60 element of the TATA-less SL RNA promoter. The second protein is similar to $SNAP_{50}$, a component of the Small Nuclear Activating Protein complex in humans. While studies in other kinetoplastids suggest that $SNAP_{50}$ associates with the SL RNA promoter, $SNAP_{50}$ has not been detected immunologically in the –60 element shift complex. Tagged forms of both proteins are being expressed in *Leishmania* with the goal of understanding their interaction and co-purifying other transcription factors.

Reconstructing human transcriptome by multi-assembly of human expressed sequences.

Yi Xing, Alissa Resch and Christopher Lee
UCLA-DOE Center for Genomics and Proteomics, Molecular Biology Institute and Dept. of Chemistry & Biochemistry

Recent evidence of abundant transcript variation in complex genomes suggests that cataloguing the complete set of transcripts from an organism is an important project. One challenge is the fact that most high-throughput experimental methods for characterizing transcripts (such as EST sequencing) give highly detailed information about short fragments of transcripts or protein products, instead of a complete characterization of a full-length form. We analyze this "multiassembly problem"—reconstructing the most likely set of full-length isoform sequences from a mixture of EST fragment data—and present a graph-based algorithm for solving it. In a variety of tests we demonstrate that this algorithm deals appropriately with coupling of distinct alternative splicing events, increasing fragmentation of the input data and different types of transcript variation (such as alternative splicing, initiation, polyadenylation, and intron retention). To test the method's performance on pure fragment (EST) data, we removed all mRNA sequences, and found it produced no errors in 40 cases tested. Comparing the protein sequences generated by our algorithm versus human-curated protein sequences from SwissProt, we found the method produced a single error out of fifty-seven cases tested. Using this algorithm, we have constructed an Alternatively Spliced Proteins database from analysis of human expressed and genomic sequences, consisting of 29,204 isoforms of 13,608 genes. ASP significantly expands available protein isoform data for the human proteome. Our data indicate that human transcript variation alters the protein product in 85% of cases, and the untranslated region in 15% of cases.