The complimentary use of small angle x-ray scattering with crystallography in the determination of biological macromolecular structures.

John Tainer, Greg Hura
Lawrence Berkeley National Laboratory

Project_Goals: Recent advances in small angle x-ray scattering (SAXS) technique and analysis have enabled shape prediction of proteins in solution. The SAXS technique is particularly powerful in combination with partial high resolution structures. SAXS can efficiently reveal the spatial organization of protein domains, including domains missing from or disordered in known crystal structures, and establish cofactor or substrate-induced conformational changes. Following a short introduction to SAXS, examples from data collected at SIBYLS, a dual SAXS and protein crystallography synchrotron beamline, will be drawn upon to demonstrate the complimentary use of SAXS with protein crystallography. I will also describe the recent implementation of a sample loading automation tool for true high throughput SAXS data collection. Several examples of the utility of high throughput SAXS will discussed in the context of the DOE/Gtl funded program project MAGGIE (Molecular Assemblies, Genes, and Genomics Integrated Efficiently).

        A core aim of the MAGGIE (Molecular Assemblies, Genes, and Genomics Integrated Efficiently) program project is to develop technologies for proteomic scale visualization of macromolecular structure. While determining function from structure remains a challenge, structure aids in understanding how macromolecules function. Detailed macromolecular structure determination from X-ray crystallography or NMR has provided a broad and deep survey of soluble biomolecules. Several techniques, both experimental and computational, exploit this information to provide significant insight into and prediction of structures which have not been probed by these often challenging atomic resolution techniques. Solution X-ray scattering (also known as SAXS: small angle X-ray scattering) has matured to provide structures at a 10Å resolution. This resolution is sufficient to elucidate a great deal of the architecture of macromolecules, how they interact and exchange products along a pathway. SAXS may be powerfully combined with information from atomic resolution and computational structure prediction methods. Here we detail our highthroughput data collection, data analysis, and data storage programs and strategies. Full SAXS data collection (including collection of buffer blanks and a 3 fold serial dilution) on a macromolecule of interest is complete in under 10 minutes. A computer processor limited data analysis tree has been developed requiring minimal human intervention. Our web accessible data storage utility BioIsis (www.bioisis.net) allows direct access to data at all stages of analysis. We demonstrate this strategy on 16 protein from pyrococus furiosis previously prepared for a crystallographic based structural genomic effort.

Redox balancing pathways in RubisCO knockout mutants of nonsulfur purple photosynthetic bacteria and the potential for enhanced biohydrogen production

Rick A. Laguna*[1] (laguna.2@osu.edu), James C. Liao[2], and **F. Robert Tabita**[1]

[1]Department of Microbiology, The Ohio State University, 484 W. 12th Avenue, Columbus, OH 43210 USA; [2]Department of Chemical Engineering, 5531 Boelter Hall, University of California at Los Angeles, Los Angeles, California 90095

Efficient bacterial metabolism and cellular integrity requires redox homeostasis. This is particularly important in nonsulfur purple (NSP) photosynthetic bacteria that generate reductant via light reactions, as well as via the oxidation of organic carbon compounds. In these organisms, redox balance is maintained via multifaceted and integrated regulatory networks that enable these organisms to respond to diverse and highly variable inter- and intracellular environments. When growing photoheterotrophically, especially with reduced carbon sources and in the absence of exogenous reductant, excess reducing equivalents are normally consumed via the reduction of CO_2. Other processes, such as nitrogen fixation, allow N_2 to compete with CO_2 as the terminal electron acceptor. Over the years (1-4), we have shown that NSP photosynthetic bacteria possess an array of metabolic and regulatory capabilities that allow for the utilization of alternative redox sinks when the primary electron sink, CO_2, is nullified via the inactivation or deletion of the RubisCO genes. RubisCO is the key enzyme of the Calvin-Benson-Bassham pathway of CO_2 assimilation. In order to grow photoheterotrophically, such RubisCO-compromised strains develop interesting strategies and alter their basic metabolic profile. For example, in many instances the derepression of nitrogenase synthesis occurs under normal repressive conditions. Such gain-of-function adaptive mutant strains have been obtained from *Rhodobacter capsulatus, Rhodobacter sphaeroides, Rhodospirillum rubrum*, and *Rhodopseudomonas palustris,* whereby such strains balance their redox potential via nitrogenase-catalyzed reduction of protons to hydrogen gas (1-4).

The use of hydrogen as an energy source is attractive because the end-product is H_2O, as opposed to CO_2, a green house gas generated by the burning of fossil fuels. H_2 is normally produced by nonsulfur purple photosynthetic bacteria during nitrogen limiting conditions by the nitrogenase complex. However, the gain-of-function adaptive mutant strains of NSP bacteria are able to produce copious quantities of H_2. Here we have compared the amount of hydrogen produced from the RubisCO-compromised *R. capsulatus, R. sphaeroides, R. rubrum, and R. palustris* RubisCO mutant strains and have shown that the levels of H_2 produced are significantly greater than previously reported. We are also investigating the integrative control of CO_2 and N_2 fixation, specifically, how normal control of the nitrogenase complex is regained when a functional CBB cycle is restored in such strains.

In order to maximize hydrogen production, JCL is developing a modeling approach to allow us to target certain aspects of metabolism that may be altered via metabolic engineering. Metabolic engineering requires selecting enzymes for expression tuning in order to transition fluxes from the wild-type state to a targeted production state. In highly connected metabolic networks, flux distribution is controlled by multiple enzymes in the network, often indirectly linked to the pathway of interest. As such, the determination of the enzyme targets for over or under expression in the metabolic network has been

challenging. The goal of this work is to develop a method to achieve this task without a complete kinetic model. Stoichiometry-based methods have been developed to handle the effects of gene knockouts that change stoichiometry. Since expression tuning does not alter metabolic network stoichiometry, such approaches cannot give an appropriate treatment for expression level changes. Furthermore, flux calculation or flux measurements give the result, but not the cause of flux transition. With anecdotal exceptions, the key enzymes whose expression changes cause the flux transition cannot be identified from the flux map. On the other hand, the traditional Metabolic Control Analysis (MCA) requires either a complete kinetic model or extensive experimentation to determine the control coefficients, which only give local but not long-range predictions. The method postulates that the most efficient way to achieve the flux change is to alter a minimal amount of enzyme levels while allowing metabolite changes to help drive the flux transition. With this assumption and a data-based sampling of reversibility, the enzymes that need to be tuned to achieve a particular flux for production can be predicted.

Finally, the derepression of the nitrogenase complex and subsequent production of hydrogen is not the only mechanism by which RubisCO deletion strains are capable of maintaining redox poise. We have isolated adaptive mutant strains that use other means to do this. One such mutant that we are currently examining is an *R. sphaeroides* RubisCO strain that up-regulates key enzymes involved in sulfate reduction, serine biosynthesis, and cysteine biosynthesis.

1. Joshi, H. M., and F. R. Tabita. A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. *Proc. Natl. Acad. Sci. USA.* **93** (1996), 14515–14520
2. Tichi, M. A., and Tabita, F.R. Maintenance and control of redox poise in *Rhodobacter capsulatus* strains deficient in the Calvin-Benson-Bassham pathway. Arch. Microbiol. **174** (2000) 322-333.
3. Tichi, M. A. and Tabita, F.R. Interactive control of *Rhodobacter capsulatus* redox balancing systems during phototrophic metabolism. J. Bacteriol. **183** (2001) 6344-6354.
4. Dubbs, J. M., and Tabita, F. R. Regulators of nonsulfur purple phototrophic bacteria and the interactive control of CO_2 assimilation, nitrogen fixation, hydrogen metabolism and energy generation. FEMS Microbiol.Rev. **28** (2004) 353-376.

Integrated Nondestructive Spatial and Chemical Analysis of Lignocellulosic Materials during Pretreatment and Bioconversion to Ethanol

Gary Peter* (gfpeter@ufl.edu), Steve Blackband, Lonnie Ingram, Richard Yost

University of Florida, Gainesville, FL

Bioconversion of lignocellulosic biomass to ethanol involves four main operations: pretreatment, hydrolysis or saccharification, fermentation, and distillation. While improvement in each of these steps is required to make bioethanol more competitive with nonrenewable fuels, pretreatment and saccharification of the biomass are critical first steps. A number of pretreatment methods increase the yield of fermentable sugars over non-pretreated biomass. However, our fundamental knowledge of what changes pretreatment causes in biomass architecture, permeability, cellular structure, subcellular distribution, composition and organization of polymers in lignocellulosic cell walls and how these changes promote or inhibit digestion by cellulases remains limited. This limited knowledge exists despite rapid advancements in noninvasive, quantitative imaging technologies from other fields of research. Our _long-term goal_ is to develop a quantitative structural model for changes that occur in the organization and chemical composition of plant biomass during pretreatment, enzymatic degradation and bioconversion to ethanol or other products. The _objectives_ of this proposed work are to 1) use advanced high resolution magnetic resonance microscopy (MRM), micro and nano x-ray computed tomography (x-ray CT), electron microscopy and imaging mass spectrometry (IMS) to quantify changes in the architecture, porosity, permeability, surface area, pore size, interconnectivity, chemical organization and composition of corn stover and particularly Populus and pine wood chips during pretreatment and enzymatic degradation, and 2) integrate the results from these different quantitative imaging methods into a model for disassembly of the plant cell wall during pretreatment and bioconversion. To accomplish these objectives we will image the same biomass materials at various times during pretreatment and enzymatic degradation of cellulose. We expect to develop this model using MRM and x-ray CT images at about 5 uM resolution. The _rationale_ for this proposed research is that 1) the porosity and permeability of the cell walls limit degradation and better quantification of changes in surface area, pore size, interconnectivity, porosity and permeability, spatial arrangement of cellulose, hemicellulose and lignins during pretreatment and digestion will be essential information for researchers developing improved pretreatment and enzymatic degradation methods, 2) knowledge about the structures that are recalcitrant or more slowly degraded are needed to develop faster and more efficient degradation processes, and 3) such information will be valuable to plant geneticists whose aim is to alter the structure and composition of plant biomass to make them more amenable to bioconversion processes.

Development of Computational Tools for the Analysis and Redesigning Metabolic Networks, (Award No. DE-FG02-05ER25684), Costas D. Maranas, Penn State, PI and Anthony Burgard, Genomatica, Inc. (coPI).

Patrick F. Suthers, Vinay Satish Kumar, Young Jung Kim, and Costas D. Maranas[*]

Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802

The rapid progress in the generation of models of metabolism for an increasing number of organisms brings to the forefront two research questions. How can computations be leveraged to automatically assess the quality of the models and subsequently correct them by using the full complement of available experimental data ? How can we optimally combine labeling experimens with external flux measurements to elucidate the maximum number of fluxes in the model ? In this poster, we highlight progress towards both of these objectives.

Starting with the first question, our systematic metabolic model development follows four key steps. These include 1) identification of biotransformations using homology searches 2) assembly of reaction sets into a genome-scale metabolic model 3) network analysis and 4) evaluation and improvement of model performance when compared against *in vivo* data. A key challenge in the automated generation of genome-scale reconstructions is the elucidation of gaps and the subsequent generation of hypotheses to bridge them (Step 3). We have proposed optimization-based procedures to identify and eliminate network gaps in these reconstructions that we have named GapFind and GapFill. First we identify the metabolites in the metabolic network reconstruction that cannot be produced or consumed under any uptake conditions and subsequently we identify the reactions from a customized multi-organism database that restores the connectivity of these metabolites to the parent network using four mechanisms. This connectivity restoration is hypothesized to take place through four mechanisms: a) reversing the directionality of one or more reactions in the existing model, b) adding reaction from another organism to provide functionality absent in the existing model c) adding external transport mechanisms to allow for importation of metabolites in the existing model, and d) restoring flow by adding intracellular transport reactions in multi compartment models. To address the evaluation and refinement of the model performance when compared to *in vivo* data (Step 4), we have developed an optimization-based procedure (i.e., GrowMatch) to identify mismatches and propose corrective hypotheses for the model. The GrowMatch procedure identifies mismatches between the *in silico* and *in vivo* growth predictions of gene deletion mutants. GrowMatch generates hypotheses for minimally perturb the model (i.e., adding or removing functionalities) so as consistency with all experimental observations is achieved.

We have applied GapFind, GapFill amd GrowMatch to existing metabolic reconstructions such as the most recent *E. coli* and *S. Cerevisiae* models revealing many gaps and ways of restoring them. We have also deployed these tools during the construction phase of a metabolic model for *Mycoplasma genitalium.* With a genome size of ~580 kb and approximately 480 protein coding regions, *Mycoplasma genitalium* is one of the smallest known free-living organisms.

The reduced genomic content of *M. genitalium* has led researchers to suggest that the molecular assembly represented by this organism may be a close approximation to the minimal set of genes required for bacterial growth. GapFind and GapFill were used to connect three metabolites by reversing the directions of three reactions, 21 metabolites by adding 12 external reactions, and an additional 12 metabolites by adding uptake reactions. Using GrowMatch, we were able to improve the percent agreement between gene-essentiality studies *in silico* as compared to in vivo experiments from 74% to 82% We have also used the model to guide the design the composition of a defined medium.

Development of a complete and balanced metabolic model does not necessarily mean that all internal fluxes are known or changes in metabolism in response to environmental or genetic perturbations can be predicted. The gold standard in elucidating fluxes in metabolic models is the use of ^{13}C isotopic label tracing experiments. In this effort, we constructed a large-scale *E. coli* isotopomer mapping model including 393 fluxes and 214 metabolites leading to over 17,000 isotopomers. The model accounts for balances on cofactors such as ATP and NADH as well as the electron transport chain, full amino acid biosynthesis and degradation, and a detailed biomass equation. Experimental results of flux elucidation are presented for an *Escherichia coli* strain engineered to produce amorphadiene, a precursor to the anti-malarial drug artemisinin. These include a statistical analysis of fluxes determined for the system such as the minimal and maximal values of the fluxes given measurement noise.

One of the important considerations for analysis of the isotope model is how isotope measurements impact the elucidation of fluxes in large-scale metabolic reconstructions. This identifiability problem in metabolic flux analysis (MFA) with isotopic considerations is very difficult as isotopic balances yield nonlinear constraints. Here, we employ an integer programming (IP) framework for the mathematical analysis of metabolic pathways to answer this question. By using a degrees of freedom based optimization method it is possible to exhaustively identify all combinations of isotope labeling experiments and flux measurement that completely resolve all flux values in the network. This approach results in an integer linear programming formulation while accounting for the case of partial measurements (e.g., when only some fragments are measured). These measurements, consisting of both partial or full isotope state determination, were assigned relative costs that allow the experimentalist to select the measurements that will be both sufficient and economical. The proposed framework has been tested on well-studied small demonstration examples. We present benchmarks of the proposed framework by applying it to medium-scale metabolic networks of *E. coli* and by revisiting our large-scale *E. coli* model to exhaustively identify all measurements options.

1,2,4-Butanetriol trinitrate (BTTN) is manufactured by the nitration of 1,2,4-butanetriol (BT). The challenges associated with chemical synthesis of BT will be discussed along with the creation of a biosynthetic pathway that allows a single microbe to catalyze the conversion of D-xylose into D-BT. Central to this created pathway is the discovery of the ability of *Escherichia coli* to catabolize D-xylonic acid and the role that the enzyme D-xylonate dehydratase plays in this catabolism. The BT biosynthetic pathway was assembled in an *E. coli* host and begins with oxidation of D-xylose to D-xylonic acid. D-Xylonate dehydrogenase, which is heterologously expressed in an *E. coli* host from the *Caulobacter crescentus xdh* locus, is recruited for this purpose. Two xylonate dehydratases encoded by *xjhG* and *yagF* loci, which were discovered to be native to *E. coli*, catalyze the conversion of D-xylonic acid into 3-deoxy-D-*glycero*-pentulosonic acid. Decarboxylation of 3-deoxy-D-*glycero*-pentulosonic acid to form 3,4-dihydroxy-D-butanal is mediated by heterologously expressed *mdlC* isolated from *Pseudomonas putida*. Final reduction of 3,4-dihydroxy-D-butanal to BT is catalyzed by an alcohol dehydrogenase native to the BT-synthesizing *E. coli*. BTTN is more stable than nitroglycerin and mixes effectively in a solvent-free process with nitrocellulose. These characteristics make BTTN an ideal replacement for nitroglycerin and a useful plasticizer in single-stage rocket motors.

Mechanisms for Transnational Coordination and Harmonization of Nanotechnology Governance

Gary E. Marchant, Kenneth W. Abbott, and Douglas J. Sylvester

Sandra Day O'Connor College of Law, Arizona State University, Tempe, AZ

Nanotechnologies are a rapidly developing set of emerging technologies being pursued by industry and governments around the world. While these technologies promise many benefits, they will also inevitably create some risks, and regulatory agencies in numerous countries are now considering regulatory oversight approaches for nanotechnology. This project is investigating models and approaches for coordinating or harmonizing international regulation of nanotechnology. The first step in the project is to create a publicly-accessible online database of transnational, national, and sub-national regulatory activities specific to nanotechnology. Examples of entries of this database, which is scheduled to go "live" by June 2008, will be presented. The second part of the project will be to identify and analyze nine different regulatory models for transitional oversight of nanotechnologies. The nine models that will be explored will be listed and briefly described.

Discovery of Novel Machinery for Lactate Utilization by *Shewanella oneidensis* MR-1

Grigoriy Pinchuk[1], Dmitry Rodionov[2*](rodionov@burnham.org), Chen Yang[2], Xiaoqing Li[2], Samantha Reed[1], Oleg Geydebrekht[1], Margaret Romine[1], Alex Beliaev[1] and **Andrei Osterman**[2,3] (osterman@burnham.org)

*Presenting author

[1]Pacific Northwest National Laboratory, Richland, Washington 99352; [2]Burnham Institute for Medical Research, La Jolla, California 92037; and [3]Fellowship for Interpretation of Genomes, Burr Ridge, Illinois 60527

Lactate is one of the major fermentative metabolism products for many microorganisms and is the most frequently used substrate for experimental studies of respiratory metabolism in *Shewanella oneidensis* MR-1. Consequently, the metabolism of lactate is a key component of the systems-level conceptual model (under development by the Shewanella Federation) linking electron transfer networks and central/peripheral carbon metabolism pathways of MR-1. Whereas physiological data has demonstrated a robust growth of *S. oneidensis* on both D- and L- forms of lactate, its genome does not contain orthologs of classical lactate dehydrogenases (LDH) such as D-LDH (gene *dld*) or L-LDH (gene *lldD*) of *E. coli*. We report here the discovery of a novel D- and L-lactate oxidative utilization machinery identified via a comparative genomic reconstruction of *S. oneidensis* MR-1 metabolism combined with physiological, genetic, and biochemical studies.

A hypothetical FeS-containing protein encoded by SO1521 was deemed a candidate for the missing D-LDH based on its presence in the putative operon with an ortholog of lactate permease (SO1522) and its remote homology with the FAD-containing D-LDH from yeast. This prediction was verified by analysis of a SO1521 targeted gene deletion mutant and by genetic complementation of *Escherichia coli* Δ*dld* mutant with a plasmid encoding SO1521. A detailed reconstruction and comparative analysis of *lactate utilization* subsystem including associated operons and regulons, across hundreds of bacterial genomes integrated in The SEED genomic platform (http://theseed.uchicago.edu/FIG/subsys.cgi) led to a conjecture that an adjacent three-gene operon SO1518-SO1520 comprised a previously uncharacterized enzymatic complex for the utilization of L-lactate. Two genes of this operon, SO1519 and SO1520 (previously annotated as a hypothetical FeS oxidoreductase and a ferredoxin-like protein, respectively) appear to form a core of this complex conserved in many divergent bacteria (e.g., uncharacterized operons *ykgEF* in *E. coli* and *yvfVW* in *Bacillus subtilis*). This prediction was validated by assay of targeted gene deletions in *S. oneidensis* and by genetic complementation and testing of the L-LDH enzymatic activity in *E. coli* Δ*lld* mutant overexpressing the SO1518-SO1520 operon. Furthermore, the inability of only two of 19 *Shewanella* sp. with completely sequenced genomes to grow with lactate as sole carbon source are consistent with the results of our comparative genome analysis of these species. These findings, in addition to the identification of previously unknown genes involved in lactate utilization in most *Shewanella* species, broadly impact our knowledge of this important aspect of carbon and energy metabolism in many other bacteria. Additional experiments are in progress to elucidate the details of the novel L-LDH complex in *S. oneidensis* and other species. This project is a component of the *Shewanella* Federation and, as such, contributes to the overall goal of applying the genomic tools to better understand the ecophysiology and speciation of respiratory-versatile members of this important genus.

Single-Cell Genomics of *Prochlorococcus*

Sébastien Rodrigue*[1] (s_rod@mit.edu), Rex Malmstrom[1], Matthew Henn[2], and **Sallie W. Chisholm**[1]

[1] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 15 Vassar street, Cambridge, Massachussetts 02139, USA.
[2] The Broad Institute, 7 Cambridge Center,Cambridge, MA 02142, USA.
*Presenting author

Complete genome sequencing of individual microbes is rapidly coming a reality through the combination of randomly-primed whole genome amplification and next generation sequencing platforms. One of the critical requirements in developing robust single-cell genomics pipelines is the elimination of background DNA from the sample before amplifying the genome of the single cell. This is particularly challenging when trying to study single microbial cells from the wild, such as the ubiquitous marine cyanobacterium *Prochlorococcus*. To overcome this challenge, we are developing a flow-cytometry based approach to remove single *Prochlorococcus* cells from the high concentrations of free-DNA normally found in the seawater. Our high throughput approach entails sorting individual cells into separate wells of a 96-well plate, amplifying extracted DNA with phi29 polymerase, and PCR screening of amplified genomic material using diagnostic primer sets. This pipeline not only provides products for whole genome sequencing of selected cells, but also enables multi-locus sequence analysis of natural populations without the need for isolation of different strains. To validate the method, we first sorted individual cells of the cultured *Prochlorococcus* strain MED4, and prepped them for sequencing using Solexa. We obtained genome coverage >80%, although there was substantial unevenness of the amplified genomic material. We are also initiating an effort to sort and amplify single *Prochlorococcus* cells collected in the tropical South Pacific ocean, where we expect to find interesting genomic variants.

Use of the D. radiodurans Repair System as a Possible Method for Assembling Synthetic Chromosomes

Sanjay Vashee*, Ray-Yuan Chuang*, Christian Barnes, Michael Montague, Hamilton O. Smith, and **J. Craig Venter** (PI)

*Co-Investigators

The J. Craig Venter Institue, Rockville, MD 20850

A major goal of our Institute is to rationally design synthetic microorganisms that are capable of carrying out the required functions. One of the requirements for this effort entails the packaging of the designed pathways into a cohesive genome. One of our approaches to this problem was to develop an efficient in vitro DNA repair system based upon *Deinococcus radiodurans* (Dr). This bacterium was selected because it has the remarkable ability to survive exposure to doses as high as 15, 000 Gy of ionizing radiation, which are otherwise lethal to almost all other organisms. Although hundreds of double-strand breaks are created during this exposure, Dr is able to accurately restore its genome without evidence of mutation within a few hours after exposure, strongly suggesting that the bacterium has a very efficient DNA repair mechanism.

Since the mechanism of DNA repair is not yet well understood in Dr, we undertook several general approaches to study this phenomenon. First, we sought to establish an endogenous extract that was capable of carrying out DNA repair. This extract can then be fractionated to isolate and purify all proteins that perform this repair. Second, we made use of information gathered from the sequenced genome. For example, homologues of *E. coli* DNA repair proteins, such as recA, recD and ruvA, etc. are present in Dr. Thus, we sought to characterize all these proteins both biochemically and genetically to probe their role in DNA repair. Representative results based on these studies are presented in the poster.

Engineering Isobutanol Production in *E. Coli*

Thomas Buelter* (tbuelter@gevo.com), Renny Feldman, Peter Meinhold

Gevo, Inc., Pasadena, CA

Advanced biofuels, which include isobutanol, will deliver the performance of gasoline without the environmental impact and these biofuels will reduce our dependency on foreign oil. Isobutanol has a higher energy content per gallon than many first generation biofuels, it does not absorb water and can be transported through the existing oil and gas distribution infrastructure. Isobutanol can be used in gas-powered vehicles without modification or blending.

Gevo has licensed technology for the production of isobutanol from metabolically engineered *E. coli* cells from the University of California Los Angeles where it was developed by Prof. James Liao. In order to make isobutanol competitive on the fuels market it has to be produced with high yield. We are working on recombinant microorganisms that are engineered to convert biomass into isobutanol without byproducts. We use strain engineering and pathway engineering to increase the yield of the isobutanol production strain. First bottlenecks and limitations of the host strain are identified. Then metabolic engineering strategies focus on directing the carbon flux from the carbon source to isobutanol completely. Pathways that produce byproducts are deactivated and the flux to isobutanol is maximized.

Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol

John Frost

1,2,4-Butanetriol trinitrate (BTTN) is manufactured by the nitration of 1,2,4-butanetriol (BT). The challenges associated with chemical synthesis of BT will be discussed along with the creation of a biosynthetic pathway that allows a single microbe to catalyze the conversion of D-xylose into D-BT. Central to this created pathway is the discovery of the ability of *Escherichia coli* to catabolize D-xylonic acid and the role that the enzyme D-xylonate dehydratase plays in this catabolism. The BT biosynthetic pathway was assembled in an *E. coli* host and begins with oxidation of D-xylose and D-xylonic acid. D-Xylonate dehydrogenase, which is heterologously expressed in an *E. coli* host from the *Caulobacter crescentus xdh* locus, is recruited for this purpose. Two xylonate dehydratases encoded by *xjhG and YagF* loci, which were discovered to be native to *E. coli,* catalyze the conversion of D-xylonic acid into 3-deoxy D-*glycero*-pentulosonic acid. Decarboxylation of 3-deoxy-D-*glycero*-pentulosonic acid to form 3,4-dihydroxy-D-butanal is mediated by heterologously expressed *mdlC* isolated from *Pseudomonas putida.* Final reduction of 3,4-dihydroxy-D-butanal to BT is catalyzed by an alcohol dehydrogenase native to the BT-synthesizing *E. coli.* BTTN is more stable than nitroglycerin and mixes effectively in a solvent-free process with nitrocellulose. These characteristics make BTTN an ideal replacement for nitroglycerin and a useful plasticizer in single-stage rocket motors.

Applications of High Throughput Solution X-ray Scattering (SXS): Progress Toward Proteomic Scale Structural Biology

John Tainer

A core aim of the MAGGIE (Molecular Assemblies, Genes, and Genomics Integrated Efficiently) program project is to develop technologies for proteomic scale visualization of macromolecular structure. While determining function from structure remains a challenge, structure aids in understanding how macromolecules function. Detailed macromolecular structure determination from X-ray crystallography or NMR has provided a broad and deep survey of soluble biomolecules. Several techniques, both experimental and computational, exploit this information to provide significant insight into and prediction of structures which have not been probed by these often challenging atomic resolution techniques. Solution X-ray scattering (also known as SAXS: small angle X-ray scattering)has matured to provide structures at a 10Å resolution. This resolution is sufficient to elucidate a great deal of the architecture of macromolecules, how they interact and exchange products along a pathway. SAXS may be powerfully combined with information from atomic resolution and computational structure prediction methods. Here we detail our highthroughput data collection, data analysis, and data storage programs and strategies. Full SAXS data collection (including collection of buffer blanks and a 3 fold serial dilution) on a macromolecule of interest is complete in under 10 minutes. A computer processor limited data analysis tree has been developed requiring minimal human intervention. Our web accessible data storage utility BioIsis (www.bioisis.net) allows direct access to data at all stages of analysis. We demonstrate this strategy on 16 protein from pyrococus furiosis previously prepared for a crystallographic based structural genomic effort.