96

Label-Free, Real Time Monitoring of Biomass Processing with Stimulated Raman Scattering Microscopy

Brian G. Saar[1]* (saar@fas.harvard.edu), Yining Zeng,[2] Christian W. Freudiger,[1,3] Yu-San Liu,[2] Michael E. Himmel,[2] Shi-You Ding,[2] (Shi.you.Ding@nrel.gov) and **X. Sunney Xie**[1] (xie@chemistry.harvard.edu)

[1]Dept. of Chemistry and Chemical Biology, Harvard University, Cambridge, Mass.; [2]Chemical and Biosciences Center, National Renewable Energy Laboratory, Golden, Colo.; and [3]Dept. of Physics, Harvard University, Cambridge, Mass.

Project Goals: (a) To develop novel optical imaging technology based on coherent Raman scattering that is capable of real time, label-free chemical imaging and (b) to apply these techniques to image the process of biomass conversion to biofuels in real time in three dimensions in intact plant tissue. The new information available from these techniques will offer insight into this complex series of chemical reactions and help to better understand and optimize their efficiency.

The conversion of plant biomass into "cellulosic" ethanol is an alternative energy technology that has attracted significant research interest over the past decades and requires new tools to understand and optimize the conversion process. We demonstrate that stimulated Raman scattering (SRS) microscopy can be used to selectively map plant cell wall polymers, such as lignin and cellulose, simultaneously, at sub-micron spatial resolution and with linear concentration dependence and high speed. We then follow the acid chlorite delignification process, to further demonstrate the real-time imaging of lignin bleaching with a time resolution of a few seconds. SRS microscopy is a high sensitivity, label-free chemical imaging technique, and provides a new tool to improve our understanding of biomass conversion processes.

Systems Biology and Metabolic Engineering Approaches for Biological Hydrogen Production

97

Metabolomics and Fluxomics of *Clostridium acetobutylicum* Part 1: Systems-Level Kinetic Flux Profiling Elucidates a Complete TCA Cycle

Daniel Amador-Noguez* (damador@princeton.edu), Xiao-Jiang Feng (xfeng@Princeton.EDU), Herschel Rabitz, and **Joshua D. Rabinowitz** (joshr@princeton.edu)

Carl Icahn Laboratory, Dept. of Chemistry and Lewis-Sigler Institute for Integrative Genomics, Princeton University, N.J.

Project Goals (Abstracts 97-99): Microbial biofuel (i.e. hydrogen and butanol) production holds great promise as a source of renewable clean energy. A critical step towards more efficient biofuel production is improved understanding of the regulation of biofuel-related metabolism and the development of models that are sufficiently accurate to enable rational control of the network behavior. With the long term aim of enabling such control, we propose to develop integrated experimental-computational technologies for quantitative dissection of microbial biofuel-producing metabolism. These tools will be broadly applicable to many microbial biofuel producers. We plan to illustrate them with the organism *Clostridium acetobutylicum*.

Clostridium acetobutylicum is an organism with great potential for the commercial production of butanol and hydrogen. Much work has been done to elucidate the metabolic pathways by which solvents are produced in this anaerobic bacterium. However, there are still key pathways of primary metabolism that remain unresolved, including the TCA cycle and amino acid biosynthesis.

As is common for obligatory anaerobic organisms, *C. acetobutylicum* does not contain in its genome obvious homologues of many of the enzymes of the TCA cycle, including citrate synthase, fumarate reductase/succinate dehydrogenase, succinyl-CoA synthetase and α-ketoglutarate dehydrogenase. The apparent lack of these genes is inconsistent with the ability of *C. acetobutylicum* to synthesize α-ketoglutarate and the glutamate family of amino acids and to grow on minimal media. To address the inconsistency, prior metabolic modeling efforts proposed that an incomplete TCA cycle might function in the reductive (counterclockwise) direction to produce α-ketoglutarate. Alternatively, it was suggested that glutamate might be synthesized from ornithine by the arginine biosynthesis pathway running in reverse.

To elucidate the actual pathway that leads to α-ketoglutarate and glutamate production, and to investigate how the TCA cycle of *C. acetobutylicum* operates *in vivo,* we studied the dynamic incorporation of various isotope-labeled nutrients into metabolites in glycolysis, the TCA cycle, the pentose phosphate pathway and amino acid biosynthetic pathways using liquid chromatography-tandem mass spectrometry (LC-MS/MS). In contrast to the previously proposed hypotheses, our results demonstrate that this organism has a complete, albeit bifurcated, TCA cycle. Ketoglutarate is produced exclusively in the oxidative direction from oxaloacetate and acetyl-CoA via citrate. Succinate acts as a dead-end metabolite that can be produced in both the reductive direction from oxaloacetate via malate and fumarate and the oxidative direction via α-ketoglutarate. Our results therefore demonstrate the presence of the biochemical activity of all currently non-annotated enzymes of the TCA cycle including fumarate reductase, citrate synthase, α-ketoglutarate dehydrogenase and succinyl-CoA synthetase. The way in which the TCA cycle bifurcates in *C. acetobutylicum,* with its capacity to synthesize succinate both oxidatively and reductively, suggests that, in addition to its biosynthetic function, it may also play an important role in redox balance. This idea is supported by our observation that most of the succinate produced is excreted.

Our investigations also yielded important information about other unresolved primary metabolic pathways in *C. acetobutylicum.* We found that the Entner-Doudoroff pathway, an alternative pathway for glycolysis, is inactive. The oxidative pentose phosphate pathway is also inactive and this organism relies exclusively on the non-oxidative pentose pathway for the production of ribose-phosphate. Our investigation of the amino acid biosynthesis pathways revealed them to be complete and canonical with the exception of glycine. Glycine was formed from threonine instead of being synthesized by the canonical pathway via serine. Additionally, the one-carbon units required for the methionine, purine, and pyrimidine biosynthesis are not produced via the usual route from serine or glycine but are instead derived from the carboxyl group of pyruvate.

The observations obtained in this study are essential for the construction of an accurate genome-scale model of *C. acetobutylicum* metabolism and lay the groundwork for better understanding of integration of biosynthetic metabolism with solvent and hydrogen production.

98

Metabolomics and Fluxomics of *Clostridium acetobutylicum*, Part 2: Quantitative Flux Model Construction and Analysis

Xiao-Jiang Feng* (xfeng@Princeton.EDU), Daniel Amador-Noguez (damador@princeton.edu), Herschel Rabitz, and **Joshua D. Rabinowitz** (joshr@princeton.edu)

Carl Icahn Laboratory, Dept. of Chemistry and Lewis-Sigler Institute for Integrative Genomics, Princeton University, N.J.

Project Goals: See goals for abstract 97.

In the first part of the research, we employed liquid chromatography-tandem mass spectrometry (LC-MS/MS) to probe the dynamic incorporation of various isotope-labeled nutrients into metabolites of *Clostridium acetobutylicum* in glycolysis, the TCA cycle, the pentose phosphate pathway and amino acid biosynthetic pathways. The labeling patterns of the metabolites identified the metabolic network structure, including a complete and bifurcated TCA cycle, which was unavailable from genome sequence analysis.

To obtain a quantitative understanding of the metabolic fluxes, we formulated an ordinary differential equation (ODE) model of the metabolic network. The model equations represent the quantitative dynamics of the labeled and unlabeled metabolites during exponential growth phase following introduction of isotope-labeled glucose. A nonlinear global inversion algorithm was employed to identify the unknown model parameters, including metabolic fluxes and some metabolite concentrations, that quantitatively reproduced the dynamic labeling data and several experimentally measured steady state constraints. Analysis of the identified model parameters indicates that the main proportion of the glycolytic flux is directed towards production of acids (butyric and acetic acid) through acetyl-CoA and amino acid biosynthesis through aspartate, while the fluxes through the two branches of the TCA cycle are relatively low.

Additionally, we performed model discrimination studies to distinguish multiple network models that can result in the same qualitative isotope labeling patterns. Traditional flux balance analysis suggests that malate and oxaloacetate are produced from fumarate in the TCA cycle. However, model identification results indicate that this structure will not be able to reproduce the observed quantitative data, and malate should be upstream of fumarate. Moreover, the model identification results also show that production of succinate from α-ketoglutarate cannot be achieved via coupling with methionine and lysine biosynthesis alone. The canonical TCA reaction of succinyl-CoA to succinate is required to describe the quantitative dynamics of the relevant metabolites.

In summary, the integrated laboratory and computational investigation generated a genome-scale quantitative flux model of *Clostridium acetobutylicum* metabolism. Model-

based analyses also provided a valuable means for unraveling certain ambiguities in the network structure. The flux model and the advanced techniques developed in the studies will serve as the basis for metabolic engineering of *Clostridium acetobutylicum* in order to achieve optimal biohydrogen production.

Student Presentation

99

Metabolomics and Fluxomics of *Clostridium acetobutylicum*, Part 3: Analysis of the Acidogenic–Solventogenic Transition

Daniel Amador-Noguez, Ian Brasg,* Xiao-Jiang Feng, Herschel Rabitz, and **Joshua D. Rabinowitz** (joshr@princeton.edu)

Carl Icahn Laboratory, Dept. of Chemistry and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, N.J.

Project Goals: See goals for abstract 97.

The solvent and hydrogen-producing bacterium *Clostridium acetobutylicum* has two major metabolic modes. During exponential phase growth it produces acids (butyric and acetic acid), and during stationary phase it takes up the acids previously produced and convert them into solvents (butanol, acetone and ethanol). Controlling this transition and stabilizing the solventogenic state are critical aspects for the commercial production of solvents using this anaerobic bacterium. To this end, it would be useful to have a comprehensive understanding of the intracellular metabolic changes that are associated with the transition between acidogenesis and solventogenesis states.

A previous attempt to tackle this question used microarrays to identify the global gene expression patterns associated with the solventogenic transition. In addition to gene expression changes in solvent producing genes, significant changes were found in a large number of primary metabolic genes in glycolysis and amino acid biosynthesis pathways. Changes in gene expression, however, do not necessarily reflect changes in enzyme activity. Moreover, since complex transcriptional alterations occurred even among genes within pathways (e.g., some increased and some decreased), the transcriptional data alone were insufficient to determine overall metabolic changes.

Kinetic flux profiling is a method for probing cellular metabolic fluxes that is based on the dynamics of cellular incorporation of isotope-labeled nutrient into downstream metabolites. We have previously used this approach to elucidate the metabolic network structure of various unresolved pathways in *C. acetobutylicum* during exponential growth phase. In this ongoing study, we are now applying this approach to investigate the metabolic differences (pathway flux changes and intracellular metabolite concentrations) between the acidogenic and solventogenic states of this organism.

We found that the flux through glycolysis does not change markedly during solventogenesis. Also, flux into the non-oxidative pentose phosphate pathway remains relatively unaffected. There was, however, a large decrease in the synthesis of most glycolysis-derived amino acids, with the notable exception of increased serine biosynthesis. Most of the fluxes coming out of pyruvate (the last metabolite in glycolysis), including into alanine, valine and oxaloacetate production were greatly decreased. This caused an increased flux into Acetyl-CoA, which cascades into increased flux thorough the acidogenic/solventogenic pathways.

In a related poster, we show that *C. acetobutylicum* has a complete TCA cycle in which succinate can be synthesized in either the oxidative or reductive direction. During solventogenesis, the reductive TCA cycle is completely shutdown. Interestingly, however, the oxidative TCA cycle remains active, producing succinate that is mostly excreted. This observation suggests that the right part of the TCA cycle may play a key role in solventogenesis by producing additional reducing power for solvent production.

Our results highlight the complex metabolic reorganization that takes place in solventogenic *C. acetobutylicum* and provide insight into some possible metabolic regulation points that could be exploited to enhance solvent production. These observations also lay the groundwork for the construction of a genome-scale dynamic quantitative model of the transition from acidogenic to solventogenic metabolism.

100

Photobiological H_2 Production in *Cyanothece* ATCC 51142

Matthew R. Melnicki[1]* (matthew.melnicki@pnl.gov), Grigoriy E. Pinchuk,[1] Jennifer L. Reed,[2] Oleg V. Geydebrekht,[1] Eric A. Hill,[1] Jim K. Fredrickson,[1] Allan E. Konopka,[1] and **Alex Beliaev**[1]

[1]Pacific Northwest National Laboratory, Richland, Wash. and [2]University of Wisconsin, Madison

Project Goals (Abstracts 100-103): The PNNL Biofuels Scientific Focus Area (BSFA) will carry out fundamental research of microbial photoautotrophs with specific emphasis on photosynthetic energy conversion, reductant partitioning, and carbon metabolism in cyanobacteria focusing on: (i) functions of genes and proteins involved in photosynthetic metabolism; (ii) regulatory factors and networks governing the expression of photosynthetic machinery and the partitioning of reductant through central metabolic pathways; (iii) pathways related to photosynthetic growth and metabolism of cyanobacteria and subsystems (e.g. light-driven electron transfer, respiration, autotrophic carbon assimilation, macromolecule synthesis, nitrogen fixation) interactions; (iv) approaches to manipulate the metabolism of cyanobacteria to channel the reducing equivalents or photosynthetic intermediates to biofuels or biofuel precursors. Consistent with the goals

of the DOE BER Biological Systems Science Program, our long-term goal is to develop predictive systems-level understanding of photosynthetic metabolism through which one can identify and address key science issues that must be resolved to advance biofuel applications.

Biological H_2 production by bacteria and microalgae has been known for more than a century, and research directed at practical application of such microbial processes has been carried out for more than three decades. Although many biohydrogen production concepts have been described, fundamental technological challenges remain in making any such process a practical reality. Advances in microbial genome sequencing and functional genomics are greatly improving the ability to conduct system-level studies of microbial metabolism and to use the obtained knowledge to identify fundamental questions that must be resolved to advance biofuel applications. Genomics and metabolic engineering hold great promise for the rational design and manipulation of biological systems to make such systems efficient and economically attractive.

The research conducted as part of the PNNL Biofuels Scientific Focus Area (BSFA) focuses on elucidating the mechanisms of light-driven metabolism in a unicellular diazotrophic cyanobacterium *Cyanothece* sp. strain ATCC51142. Conditions promoting H_2 production by *Cyanothece* 51142 are being studied in order to develop a strategy for maximizing the output of H_2 using metabolic modeling approach. Initially, two-phase experiments have been employed to promote photosynthetically driven accumulation of glycogen that is subsequently converted to H_2. Specifically, during the first phase, strain 51142 was grown in continuous cultures under N-limitation in a photobioreactor sparged with CO_2-enriched Ar (0.3% v/v) and continuously illuminated at 150 μmol/m^2•s. Upon reaching steady-state (biomass concentration 80 mg/l of ash-free dry weight), the cultures were incubated in N-free medium in the absence of CO_2 using 100% Ar as sparging gas while measuring the off-gas composition by in-line mass-spectrometry. Phase two was initiated by placing cultures either under dark or the light (200 μmol/ m^2•s) conditions. Appropriate controls consisting of light and dark cultures amended with ammonium or 3- (3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), an inhibitor of photosystem II, were also included in the experiment. The data obtained revealed that, unlike non-nitrogen fixing cyanobacteria such as *Synechocystis* and *Synechococcus* spp., *Cyanothece* 51142, is not capable of generating significant amounts of H_2 from stored glycogen under dark fermentative conditions, nor does it produce H_2 under light conditions in the presence of ammonia or N_2. However, illuminated cultures exposed to an Ar atmosphere and deprived of N_2 and CO_2 produced significant amounts of H_2. O_2 was also produced along with H_2 at 1:2 ratio, whereas DCMU significantly (4.5-fold) decreased H_2 generation. It should be noted that analysis of cell-free culture supernatants did not revealed any accumulation of organic acids. These results suggest that PSII and therefore water photolysis played a significant role in H_2 evolution by strain 51142. The inhibition of H_2 production by ammonia or N_2 strongly *suggest that nitrogenase was the* enzyme primarily responsible for light-driven H_2 production, and whole cell assays revealed high nitrogenase activity in H_2-producing cells. Preliminary analysis of the material balance suggests the nitrogenase activity was supported by light-driven electron transfer. Within the scope of the proposed BSFA research, we will further elucidate and validate the pathways of light-driven two-step H_2 production by cyanobacteria and incorporate the experimental data into the metabolic model of *Cyanothece* 51142 to identify the means for maximization of H_2 production by this organism.

101

Constraint-Based Modeling for Maximizing the Metabolic Potential of Photoautotrophic Microorganisms

Jennifer L. Reed[1]* (reed@cae.wisc.edu), Trang Vu,[1] Jim K. Fredrickson,[2] Oleg V. Geydebrekht,[2] Eric A. Hill,[2] Allan E. Konopka,[2] Grigoriy E. Pinchuk,[2] Sergey M. Stolyar,[2] Jessica De Ingeniis,[3] Andrei L. Osterman,[3] Donald A. Bryant,[4] and **Alex S. Beliaev**[2]*

[1]University of Wisconsin, Madison; [2]Pacific Northwest National Laboratory, Richland, Washington; [3]Burnham Institute for Medical Research, La Jolla, California; and [4]Pennsylvania State University, University Park

Project Goals: See goals for abstract 100.

Photosynthetic microorganisms possess the unique ability to convert sunlight into chemical energy using water as the electron donor. Despite the wealth of information on the mechanistic aspects of bacterial photosynthesis and supramolecular complexes catalyzing the process of light conversion and CO_2 fixation, a system-level understanding of photosynthetic metabolism is yet to be achieved. Different phototrophic microorganisms display varying levels of light conversion efficiencies, which ultimately translate in different rates of electron transfer, ATP/NAD (P)H production, and growth. Understanding the origin of these properties will provide fundamental new insights that could be widely applied to the development of photosynthetic systems for biofuels development. Integral to that is the question, is the process of reductant partitioning in photoautotrophs which links energy-generating reactions with biosynthesis of biomass precursors and storage compounds. Within the scope of the PNNL Biofuels Scientific Focus Area (BSFA), we are exploring the mechanisms of energy conservation and carbon partitioning in cyanobacteria. One important outcome of the project will be development of a predictive tool, *i.e.* a genome-scale model, which provides a platform for integrating all knowledge and experimental data generated within the project. It will also have the ability to serve as an *in silico* tool for manipulating photosynthetic microorganisms to act as catalysts for solar energy conversion and will potentially allow development of a highly efficient biofuel production process.

As part of previous Genomics:GTL funding, we have built a genome-scale metabolic network for *Cyanothece* sp. ATCC 51142, a unicellular diazotrophic cyanobacterium that can temporally separate the process of light-dependent autotrophic growth and glycogen accumulation from N_2 fixation. The resulting model currently includes 798 genes, 682 proteins, 630 metabolites, and 656 reactions accounting for common pathways such as central metabolism, nucleotide and amino acid biosynthesis, and those that are more unique to cyanobacteria such as photosynthesis, carbon fixation, and cyanophycin production. Photosynthesis was modeled as sequential reactions that occur in each photosystems, in order to study the effect of different light wavelengths, and separate photosystem activities on cellular growth and hydrogen production rate. Predicted results from the metabolic model, based on growth simulations of the constraint-based model under different carbon and nitrogen sources for photoautotrophic, heterotrophic and mixotrophic conditions qualitatively agree with experimental data. Using a custom-built photobioreactor, which allows for the control and monitoring of incident and transmitted light, we have also studied the physiological response of *Cyanothece* sp. ATCC 51142 to nitrogen and light limitations imposed on photosystems I and II. Biomass composition and metabolite analyses were carried out to provide experimental validation for the model.

In addition, we have developed a draft metabolic network for *Synechococcus* sp. PCC 7002, a fast growing non-nitrogen-fixing cyanobacterium which exhibits the fastest growth rate of known cyanobacteria and is also remarkably tolerant to high light intensities. Understanding the origin of these properties could provide fundamental new insights that could be widely applied to the development of other biological systems for biofuels development. Initial comparisons between the reconstructed metabolic networks of *Cyanothece* 51142 and *Synechococcus* 7002 suggested that both networks share a significant number of pathways. However *Synechococcus* 7002 also displays notable differences, specifically in pathways involved in amino acid and folate metabolism. Once the reconstruction of *Synechococcus* 7002 network is complete, we will apply metabolic engineering algorithms to identify strategies for modulating the efficiencies of light conversion, carbon fixation, and photosynthate production.

102

Genome-Enabled Studies of Photosynthetic Microorganisms for Bioenergy Applications

Alex S. Beliaev[1]* (alex.beliaev@pnl.gov), **Allan E. Konopka**,[1] Harvey Bolton Jr.,[1] Mary S. Lipton,[1] Thomas O. Metz,[1] Grigoriy E. Pinchuk,[1] Aaron T. Wright,[1] Thomas C. Squier,[1] Sergey M. Stolyar,[1] Donald A. Bryant,[2] Matthew C. Posewitz,[3] and Jennifer L. Reed[4]

[1]Pacific Northwest National Laboratory, Richland, Wash.; [2]Pennsylvania State University, University Park; [3]Colorado School of Mines, Golden; [4] University of Wisconsin, Madison

Project Goals: See goals for abstract 100.

With the increasing concerns over the sustainability of a crop-based biofuel economy, there is a renewed interest in photosynthetic microorganisms, which use solar energy, H_2O, and CO_2, as effective alternatives for the production of biofuels and primary biomass. Cyanobacteria and microalgae have the potential to produce biofuels at a much higher productivity than vascular plants and they can be cultivated in freshwater and marine aquatic environments that do not compete for land resources with conventional agriculture. While structural and functional properties of protein complexes catalyzing the first steps of photosynthetic energy conversion reactions have been extensively explored, harnessing photosynthetic metabolism for biofuels production requires detailed knowledge of cellular subsystems and networks involved in electron transport, reductant partitioning, and energy storage pathways. The advances in microbial genome sequencing and functional genomics have greatly improved the ability to construct accurate systems-level models of microbial metabolism and to query the models for gene targets that enhance productivity by metabolic engineering.

The PNNL Biofuels Scientific Focus Area (BSFA) conducts fundamental research of microbial photoautotrophs with specific emphasis on photosynthetic energy conversion, reductant partitioning, and central carbon metabolism focusing on: (i) functions of genes and proteins involved in photosynthetic metabolism; (ii) regulatory factors and networks governing the expression of photosynthetic machinery and the partitioning of reductant through central metabolic pathways; (iii) pathways related to photosynthetic growth and metabolism of cyanobacteria and subsystems (e.g., light-driven electron transfer, respiration, autotrophic carbon assimilation, macromolecule synthesis, nitrogen fixation) interactions; (iv) approaches to manipulate the metabolism of cyanobacteria to channel the reducing equivalents or photosynthetic intermediates to biofuels or biofuel precursors. Leveraging from the laboratory's cutting-edge technical capabilities, the research conducted under the PNNL BSFA will embody both scientific and technical tasks including development of continuous cultivation, biochemical, and genetic methods in conjunction with genomic, proteomic, metabolomic and modeling approaches for studying funda-

mental aspects of the phototrophic metabolism. Consistent with the goals of DOE BER Genomic Science Program, our long-term goal is to develop predictive systems-level understanding of photosynthetic metabolism through which one can identify and address key science issues that must be resolved to advance biofuel applications.

103

Phototroph-Heterotroph Co-Cultures for Studying Organism Interactions and Pathways of Solar Energy Conversion

Grigoriy E. Pinchuk[1]* (grigoriy.pinchuk@pnl.gov), Oleg V. Geydebrekht,[2] Jennifer L. Reed,[2] Donald A. Bryant,[3] Allan E. Konopka,[1] **Jim K. Fredrickson**,[1]* and **Alex S. Beliaev**[1]*

[1]Pacific Northwest National Laboratory, Richland, Wash.; [2]University of Wisconsin, Madison; and [3]Pennsylvania State University, University Park

Project Goals: See goals for abstract 100.

Much of terrestrial and aquatic photosynthesis occurs in complex microbial consortia, but little is presently known about interactions among microorganisms that contribute to their efficient solar energy capture and conversion. Under the PNNL Foundational Scientific Focus Area (FSFA), we are utilizing a combination of complementary laboratory-based research and exploration of naturally-occurring associations with combinations of defined photoautotroph-heterotroph cultures to understand the collective energy, carbon, and nutrient processing in microbial systems. The complex natural systems we have selected for study including the phototrophic mats of Yellowstone National Park (YNP) and surface waters of central Washington as of yet have not yielded cultivated members that can be used as model systems for detailed laboratory-based research. Due to the paucity of information on the heterotrophic population (s) in these natural systems, initial research and development of methodology for studying microorganism interactions is utilizing co-cultures of representative cyanobacteria and well-studied *Shewanella* species. While recognizing the opportunistic nature of this system, there is ample evidence that certain species of *Shewanella* live in association with autotrophic prokaryotes and examples of these associations are well documented. More importantly, we believe that a *Synechococcus-Shewanella* co-culture can be instrumental in gaining basic understanding of opportunistic interactions between photoautotrophic and heterotrophic bacteria. Our preliminary results using cyanobacteria-*Shewanella* co-cultures demonstrated that metabolic coupling and interactions between photoautotrophic and heterotrophic microorganisms may serve as a mechanism for controlling dissolved O_2 concentration, increasing Fe and Mn availability, and recycling nutrients in natural communities.

Complementary to the FSFA work, we are also exploring the potential of photoautotroph-heterotroph associations for bioenergy applications as part of the PNNL Biofuels Scientific Focus Area. Although phototroph-heterotroph associations are abundant in nature, the co-culture approach has been seriously under-appreciated. To date, engineering of microbes for biofuel production is being carried out using single strains by enhancing or deleting specific steps of a pathway or modulating activities of specific enzymes. However, synthesis of biofuel precursor molecules requires precise coordination and interactions of many proteins within various pathways, where any adjustments or increases in expression and/or activity levels can lead to substantial metabolic burden and suboptimal yields. In that regard, engineering of photosynthetic organisms, which carry out simultaneous light- and dark-phase reactions, is inherently challenging. Photosynthetic production of biofuels often requires optimization of two or more metabolic functions which can be mutually exclusive in a single microbial cell and therefore require either spatial and/or temporal separation (e.g. O_2 evolution and H_2 production; O_2 evolution and N_2 fixation, sensitivity of RuBisCo to O_2). To that end, co-culturing of photosynthetic and heterotrophic microorganisms offers efficient ways to optimally engineer the photosynthetic production of biofuels. By engineering photosynthetic strains which excrete organic carbon compounds (organic acids or sugars) and co-culturing them with a heterotrophic organism capable of utilizing the excreted compounds, one can physically separate the processes of photosynthesis and photosynthate conversion while allowing for net CO_2 consumption. The co-cultivation of phototrophs and aerobic heterotrophs also eliminates technical problems associated with oxygen-sensitivity and substrate delivery by creating favorable microaerobic CO_2-enriched environments for the photorophic microorganisms. Moreover, utilization of exogenously-added organic carbon by the heterotroph, decreases dissolved O_2 concentrations and induces the expression of O_2-sensitive enzymes (*e.g.* hydrogenase and nitrogenase) in the phototroph which in turn will generate reducing equivalents by light-driven water photolysis. Overall, we believe that the implementation of the co-culture approach will open new perspectives for designing efficient and cost-effective processes and will provide a novel platform for the development of consolidated bioprocessing methods leading to production of carbon-neutral energy at reduced economic and energetic costs.

104

Probing Metalloenzymes with Synchrotron Radiation – from Gamma Rays to Soft X-Rays

Stephen P. Cramer[1,2]* (spjcramer@ucdavis.edu)

[1]Dept. of Applied Science, University of California, Davis and [2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Project Goals: We are trying to develop a resource for biological and environmental x-ray spectroscopy at ALS beamline 4.0.2.

Nitrogenase is the enzyme responsible for the 'fixation' of nearly inert atmospheric dinitrogen to ammonia. It is ultimately responsible for half of the world's protein, while the other half depends on industrial fertilizer produced with hydrogen derived from fossil fuels. Nitrogenase uses a complex $MoFe_7S_9X$-homocitrate 'FeMo-cofactor' to convert N_2 to NH_3, but the detailed mechanism remains poorly understood.[1-3] Another type of enzyme, hydrogenase, catalyzes the interconversion of dihydrogen with protons and electrons. These enzymes use unusual forms of Fe-S clusters or Fe carbonyls, and their catalytic mechanisms are not understood.

One way to study Fe in biological systems is Nuclear Resonance Vibrational Spectroscopy (NRVS). In this synchrotron radiation technique, a sample is excited with a ~1 meV bandwidth beam near a Mössbauer resonance, and the delayed fluorescence is recorded as a function of excitation energy. When applied to Fe samples, NRVS is only sensitive to vibrations involving motion of ^{57}Fe. We will present results on model compounds, small Fe-S proteins, nitrogenase, and hydrogenase, and the needs and prospects for future improvements will be discussed.[4]

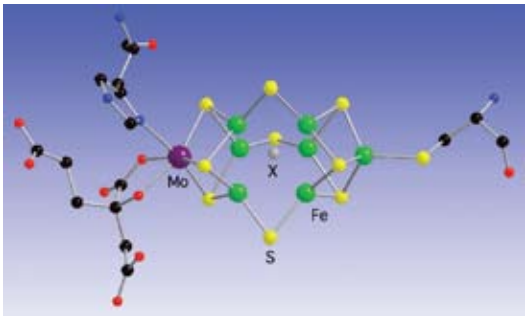


Figure. Structure of the nitrogenase FeMo-cofactor.

Although NRVS is a powerful tool for studying Fe, it is not applicable to the other metal centers of interest: Mo in nitrogenase or Ni in hydrogenase. We have thus resorted to soft x-ray spectroscopy as a probe for these sites. Some Ni L-edge and Mo M-edge spectra will be presented, and the issue of radiation damage with soft x-rays will be discussed.[5-6]

Finally, our efforts to develop a soft x-ray resource, ABEX, for biological and environmental science at the ALS will be summarized.

References

1. Peters, J.W.; Szilagyi, R.K. Curr. Opin. Chem. Biol., 2006, 10, 101-108.
2. Barney, B.M.; Lee, H.-I.; Santos, P.C.D.; Hoffman, B.M.; Dean, D. R.; Seefeldt, L.C. Dalton Trans., 2006, 2277–2284.
3. Dance, I. Chem. Asian J., 2007, 2, 936-946.
4. Xiao, Y.; Fischer, K.; Smith, M.C.; Newton, W.; Case, D.A.; George, S.J.; Wang, H.; Sturhahn, W.; Alp, E.E.; Zhao, J.; Yoda, Y.; Cramer, S.P. J. Am. Chem. Soc., 2006, 128, 7608-7612.
5. George, S.J.; Fu, J.; Guo, Y.; Drury, O.; Friedrich, S.; Rauchfuss, T.; Volkers, P.I.; Peters, J.C.; Scott, V.; Brown, S.D.; Thomas, C.M.; Cramer, S.P. Inorg. Chim. Acta, 2008, 361, 1157-1165.
6. George, S.J.; Drury, O.B.; Fu, J.; Friedrich, S.; Doonan, C.J.; George, G.N.; White, J.M.; Young, C.G.; Cramer, S.P. J. Inorg. Biochem., 2009, 103, 157-167.

Student Presentation

105

Stopped-Flow IR Spectroscopy of Hydrogenases

Celestine Grady,[1,2]* Simon George,[1,2] **Stephen P. Cramer**[1,2] (spjcramer@ucdavis.edu), Eric Shepard,[3] Shawn McGlynn,[3] Joan Broderick,[3] John Peters,[3] Jon M. Kuchenreuther,[4] James A. Stapleton,[4] and James R. Swartz[4]

[1]Dept. of Applied Science, University of California, Davis; [2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, Calif.; [3]Dept. of Chemistry and Biochemistry, Montana State University, Bozeman; and [4]Dept. of Chemical Engineering, Stanford University, Stanford, Calif.

Project Goals: Studying hydrogenases with IR spectroscopy can provide information on the creation of hydrogenase, the molecular and atomic makeup of the enzyme active site, the surrounding amino acid environment, and the mechanisms by which the enzyme reacts with hydrogen or inhibitors like oxygen. The understanding of these properties could lead to the development of improved mutant enzymes that can provide a substantial source of hydrogen.

As we move toward a future that depends less on hydrocarbons and more on a variety of sustainable energy sources, we recognize that hydrogen could be an important part of a new clean energy infrastructure. A possible effective and feasible solution for the mass production of hydrogen is through the manipulation of hydrogenases, enzymes that reversibly catalyze the evolution of molecular hydrogen from protons and electrons. Hydrogenases are found in organisms such as algae that could potentially be harvested for hydrogen production.

Studying hydrogenase with IR spectroscopy can provide information on the creation of the enzyme, the molecular makeup of the enzyme active site, the surrounding amino acid environment, and the mechanisms by which the enzyme reacts with hydrogen or inhibitors such as CO and oxygen. An IR spectrum is obtained by shining IR light through a sample and measuring the amount of light that is transmitted. There will be visible absorption lines for the frequencies of light that have excited some vibration within the molecule. Each vibration is unique to a functional group, and the IR spectrum can be interpreted to determine the composition of the sample. Here we present IR spectra of as-isolated hydrogenase samples, and hydrogenases combined with sodium dithionite, CO and oxygen. In addition,

we have used IR techniques such as Stopped Flow-FTIR and photolysis to obtain time-dependent observations of these hydrogenase reactions.

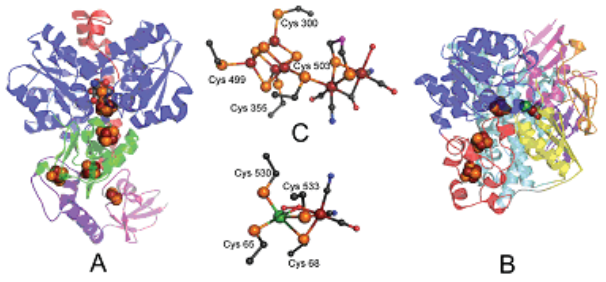


Figure. Structures of (A) CpI H_2ase, (B) [NiFe] H_2ase, (C) H-cluster, and (D) NiFe active site.

106

Novel Hydrogen Production Systems Operative at Thermodynamic Extremes

Michael J. McInerney,[1] Ralph S. Tanner,[1] Gregory Hurst,[2] Lars Rohlin,[3] and **Robert P. Gunsalus**[3]* (robg@microbio.ucla.edu)

[1]University of Oklahoma, Norman; [2]Oak Ridge National Laboratory, Oak Ridge, Tenn.; [3]University of California, Los Angeles

http://www.mimg.ucla.edu/faculty/gunsalus/

Project Goals: The goals of this collaborative project are to develop new research strategies to address the Genomics:GTL program needs in the area of bio-hydrogen production. This includes the delineation of the molecular machinery involved in hydrogen production from thermodynamically difficult substrates, as well as the characterization of new microbial model systems that generate high H_2 concentrations approaching 17% of the gas phase. We are performing systems-based studies of bio-hydrogen production in model anaerobic consortia as well as with pure culture model strains to identify key regulated steps. The results of these studies will greatly expand our ability to predict and model systems for H_2 production in novel anaerobes that are currently very poorly understood.

The goals of this collaborative project are to develop new research strategies to address the Genomics:GTL program needs in the area of bio-hydrogen production. This includes the delineation of the molecular machinery involved in hydrogen production from thermodynamically difficult substrates, as well as the characterization of new microbial model systems that generate high H_2 concentrations approaching 17% of the gas phase.

To identify the strategies used for hydrogen production in one model butyrate-degrading organism, we are performing genomic, proteomic, and transcript analysis on *Syntrophomonas wolfei*. This microbe is representative of an important but poorly understood class of hydrogen-producing organisms that are capable of syntrophic fatty and aromatic acid metabolism when co-cultured with suitable microbial partner(s). Their ability to produce H_2 requires reverse electron transport with energy input. Analysis of the *S. wolfei* genome reveals many genes with potential to accomplish this task. It possesses genes for three cytoplasmic and two externally located formate dehydrogenases plus two cytoplasmic and one externally located hydrogenase. By implication, either hydrogen or formate could be produced by *S. wolfei* during syntrophic interactions. Interestingly, the three cytoplasmic formate dehydrogenases plus one of the soluble-type hydrogenases appear to be NADH-linked since the respective gene clusters contain genes for NADH:quinone oxidoreductases chains E and F. This suggests that *S. wolfei*, like several other sequenced syntrophic metabolizers and anaerobes known to produce high molar ratios of hydrogen from glucose, may produce H_2 and/or formate from NADH by an electron bifurcation mechanism. To determine which of the above enzymes are utilized in pure culture, proteomic studies were performed initially using crotonate-grown cells. This was accomplished by analyzing whole cell-derived peptide mixtures with two-dimensional liquid chromatography/tandem mass spectrometry (2D LC-MS-MS) via the MudPIT approach. Two highly expressed hydrogenase enzymes were detected plus a novel, electron transfer flavoprotein-linked FeS-type reductase complex that is probably used to process electrons generated by the oxidation acyl-CoA intermediates. *S. wolfei* metabolizes fatty acids by the β-oxidation pathway and surprisingly, the genome reveals multiple homologues for many of the enzymatic steps even though it has an extremely restricted range of fatty acid substrates. Provisional protein assignments were also made by LC-MS-MS for each of the eight reactions leading to acetate formation. Energy is harvested by substrate level phosphorylation via acetate kinase to yield one ATP. Additional proteomic and transcript studies are in progress to further characterize the expression of genes/proteins involved in the reversed electron transfer process for hydrogen production. These studies will yield improved understanding of how *S. wolfei* and other syntrophic metabolizers thrive at low thermodynamic driving forces not possible for many other anaerobes.

In a companion project we are characterizing the genetic, biochemical, and physiological properties of a newly isolated anaerobic bacterium called *Anaerobaculum hydrogenoformans* strain OS1 that can generate H_2 at concentrations up to 17%. In one approach, genomic sequencing is being performed on this representative member of the Synergistetes group. The genome of approximately 2.4 MB in size has a GC content of 46.5%. The current assembly consists of 403 contigs with about 1.9 MB contained in the top six contigs. Machine annotation and manual curation is currently in progress to support a metabolic reconstruction of the cellular metabolism leading to hydrogen formation in strain OS1 when grown on hexoses and pentoses. Since strain OS1can also grow syntrophically in the presence of a H_2-consuming methanogen, it suggests an ability for a more complex alternative lifestyle.

* Presenting author

107

Hydrogen Production Comes at the Expense of Calvin Cycle CO_2 Fixation During Photoheterotrophic Growth by *Rhodopseudomonas palustris*

James B. McKinlay* (mckinla1@uw.edu), Yasuhiro Oda, Colin R. Lappala, and **Caroline S. Harwood**

Dept. of Microbiology, University of Washington, Seattle

Project Goals: The goals of this project are (i) to use ^{13}C-metabolic flux analysis and other approaches to understand the central metabolic changes involved in H_2 production by *R. palustris* and (ii) to use the resulting information to guide the metabolic engineering for improving H_2 production characteristics.

There is currently a pressing need for renewable fuels to negate the adverse social, economic, and environmental impacts of burning fossil fuels. H_2 is a promising biofuel, having about three-times the energy content of gasoline. Although most manufactured H_2 comes from fossil fuels H_2 can also be produced biologically. *Rhodopseudomonas palustris* uses energy from sunlight and electrons from organic waste to produce H_2 via nitrogenase. In order to understand and improve this process we used ^{13}C-acetate to track and compare central metabolic fluxes in non-H_2 producing wild-type *R. palustris* and an H_2-producing mutant. Wild-type cells metabolized 22% of the acetate to CO_2 and then fixed 68% of this CO_2 into cell material using the Calvin cycle. This Calvin cycle flux enabled *R. palustris* to re-oxidize nearly half of the reduced cofactors generated during acetate oxidation. The H_2-producing mutant produced a similar amount of CO_2 but the Calvin cycle flux was much lower, re-assimilating only 12% of the CO_2. In this mutant, H_2 production assumed much of the redox balance burden as about 90% of the electrons for H_2 production were diverted away from the Calvin cycle. Microarray and Q-PCR analyses showed that the shift of electrons from the Calvin cycle to H_2 production involved transcriptional control of Calvin cycle operons. However, this transcriptional control did not require the redox-sensing two-component regulatory system RegSR. When Calvin cycle flux was disrupted completely by deleting the genes encoding ribulose 1,5-bisphosphate carboxylase, *R. palustris* was forced to use H_2 production alone to maintain redox balance. This mutant exhibited a 1.5-fold increase in H_2 yield but at a cost to the growth rate. These results demonstrate how systems level approaches, such as ^{13}C-metabolic flux analysis, can lead to effective strategies to improve H_2 yield. Furthermore, our results underscore that the Calvin cycle and nitrogenase have important electron-accepting roles separate from their better known roles in ammonia production and biomass generation.

108

Optimization of NSR-seq for Transcriptome Analysis in *Rhodopseudomonas palustris*

Christopher D. Armour,[1,2] Yasuhiro Oda[3]* (yasuhiro@u.washington.edu), John C. Castle,[1,4] Somsak Phattarasukol,[3] Colin R. Lappala,[3] Matthew C. Biery,[1,2] **Caroline S. Harwood**,[3] and Christopher K. Raymond[1,2]

[1]Rosetta Inpharmatics LLC, Merck and Co., Inc., Seattle, Wash.; [2]NuGEN Technologies, Inc., San Carlos, Calif.; [4]Institute for Translational Oncology and Immunology, Mainz, Germany; and [3]Dept. of Microbiology, University of Washington, Seattle

Project Goals: The goal of this project is to exploit next generation sequencing technology to characterize the transcriptional networks underlying hydrogen production in the photosynthetic bacterium *Rhodopseudomonas palustris*.

The emergence of next generation sequencing (NGS) technology has opened up new opportunities for the optimization of biological systems for alternative energy production. The open query data format of NGS digital readouts is particularly useful for characterizing the transcriptional networks underlying biofuel production in genetically diverse bacterial strains. The success of sequence-based expression profiling, however, depends on the availability of efficient methods for the construction of high complexity cDNA libraries that are compatible with NGS platforms. Conventional random-priming techniques produce libraries that are largely composed of ribosomal RNA (rRNA) transcripts, so affinity purification schemes are commonly applied to reduce rRNA content prior to reverse transcription. While this approach has been moderately effective at removing rRNA in some systems, it has generally been ineffective for organisms with genomes of high G + C content. Moreover, affinity-based rRNA depletion requires high RNA inputs to obtain ample purified material for cDNA synthesis. To overcome these limitations, we have developed an alternative strategy, called Not-So-Random (NSR) priming, which utilizes computationally designed hexamers to synthesize cDNA selectively from non-rRNA template molecules (Armour et al., 2009). In addition to reducing rRNA load, NSR library construction preserves transcript strand polarity and requires only 1 μg of total RNA input. We have adapted NSR-seq methodology, which was originally developed in mammalian systems (Armour et al., 2009), to the photosynthetic bacterium *Rhodopseudomonas palustris* to dissect the biochemical pathways involved in hydrogen production using the Illumina Sequencing-By Synthesis platform.

With the ultimate goal of profiling diverse strain backgrounds, we designed NSR hexamers against rRNA transcripts obtained from six *R. palustris* strains for which complete genome sequences were available. Alignment of all possible hexamer sequences (4,096) to the 5S, 16S, and 23S rRNA sequences from each strain resulted in the identification of 1,203 NSR primers that had no perfect match

complementarity to any of the rRNA filter transcripts. Oligonucleotides containing each NSR hexamer were synthesized individually with a 10 nt universal tail sequence at the 5' terminus and pooled prior to library construction. An NSR hexamer pool synthesized in the forward strand orientation was used for second strand synthesis (sense), whereas a pool containing the reverse complements was used in the antisense cDNA reaction. Distinct 5' tail sequences were used for the first and second strand synthesis primer pools, so that transcript strand orientation would be maintained through the library construction process.

Sequence analysis of a test library generated with total RNA isolated from *R. palustris* strain CGA009 indicated that NSR-priming did result in mRNA enrichment compared to conventional random-priming, but the effect was modest relative to a random-primed library built with RNA that had been pre-treated with the MICROB*Express*™ Bacterial mRNA Enrichment Kit (Ambion, Inc.). Only 3% of the reads mapping to the genome aligned unambiguously in the random primed control library, compared to 7% and 11% for the NSR and MICROB*Express*™ treated libraries, respectively. Moreover, we found that NSR-priming had a differential effect on rRNA species; 23S abundance was reduced by 50%, whereas 16S levels increased slightly relative to the control. A closer inspection of the distribution of NSR reads across rRNA loci revealed that only a few template sites accounted for the majority of rRNA priming events. This allowed us to remove problem hexamer sequences from the original NSR primer pool without significantly diminishing sequence complexity. Constructing libraries with a refined set of NSR primers that included 925 of the original 1,203 hexamers, the so-called 'cut300' primer set, increased the number of unambiguous alignments to 22% of all mapped reads. We also observed that specific bases in the universal primer sequence upstream of the NSR hexamer site contributed to rRNA priming, thus offering another opportunity to enhance primer selectivity. Re-engineering the 3' end of the tail sequence (CGA>TTA) further increased mRNA enrichment. With these improvements, 42% of NSR reads that map to the genome align to unique sites. Further testing indicated that the resulting NSR-seq mRNA expression profiles were highly reproducible and strand-specific.

Reference

1. Armour, C. D., J.C. Castle, R. Chen, T. Babak, P. Loerch, S. Jackson, J.K. Shah, J. Dey, C.A. Rohl, J.M. Johnson and C.K. Raymond (2009) Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. Nat Methods 6:647-649.

109

Strand-Specific NSR RNA-seq Analysis of *Rhodopseudomonas palustris* Reveals Additional Features of its Transcriptome that May Influence Hydrogen Production

Yasuhiro Oda[1]* (yasuhiro@u.washington.edu), Colin R. Lappala,[1] Somsak Phattarasukol,[1] Christopher D. Armour,[2,3] John C. Castle,[2,4] Matthew C. Biery,[2,3] Christopher K. Raymond,[2,3] and **Caroline S. Harwood**[1]

[1]Dept. of Microbiology, University of Washington, Seattle; [2]Rosetta Inpharmatics LLC, Merck and Co., Inc., Seattle, Wash.; [3]NuGEN Technologies, Inc., San Carlos, Calif.; and [4]Institute for Translational Oncology and Immunology, Mainz, Germany

Project Goals: The overall goal of this project is to use a systems level approach to dissect metabolic and regulatory networks necessary for nitrogenase-catalyzed hydrogen production by a phototrophic bacterium *Rhodopseudomonas palustris*.

Hydrogen gas has good potential for as a transportation fuel because it is clean burning and has a high energy content. Bacteria can produce hydrogen by several different enzymatic routes. The photosynthetic bacterium *Rhodopseudomonas palustris* produces copious amounts of hydrogen via the enzyme nitrogenase, which generates both ammonia and hydrogen as products of dinitrogen gas reduction. When nitrogen gas is not available, nitrogenase uses only protons and electrons as substrates and produces pure hydrogen gas. This reaction requires large amounts of ATP and electrons, which *R. palustris* can obtain from sunlight and biomass, respectively, by fairly complex metabolic routes. Thus hydrogen production involves the appropriate integration of dozens of metabolic reactions. Our long-term goal is to integrate transcriptomic and phenotypic data from up to 100 *R. palustris* stains using Bayesian network analysis to identify all genes involved in hydrogen production. This would include genes; such as central carbon metabolism genes that may change very little in expression and are therefore are not easily recognized when just a few strains are analyzed by conventional transcriptomic techniques.

As a start, we have been working to develop improved techniques of transcriptome analysis by deep cDNA sequencing. This is necessary because strain-to-strain variation precludes the use of traditional microarrays. We modified for use in bacteria a strand-specific cDNA sequencing method called Not-So-Random (NSR) RNA-seq (Armour et al., 2009). This method uses a collection of computationally selected oligonucleotides to selectively enrich non-rRNA cDNAs. Also the cDNA libraries are prepared in such a way as to preserve strand specificity and therefore reveal the overarching themes of sense and antisense strand transcription across the genome (Armour et al., 2009). The cDNA libraries are sequenced by Illumina sequencing technology (25 bp sequence reads).

* Presenting author

When we applied this method to *R. palustris* strains CGA009 and TIE-1, on the order of 60% of the total sequencing reads were non-rRNA reads starting from 1 μg of total RNA. Of the remaining reads (on the order of 2 million) approximately 70% mapped to genes with no base pair mismatches. We tested three growth conditions: nitrogen-fixing (hydrogen-producing) – high light, nitrogen-fixing (hydrogen-producing) – low light, and ammonia – high light. In each condition over 90% of the genes in each genome were expressed. The most highly expressed genes in the genomes were light harvesting 2 (LH2) and light harvesting 4 (LH4) genes. Interestingly the LH4 operon was expressed at higher levels under nitrogen-fixing as compared to ammonia-grown conditions at high light, perhaps reflecting the increased need for cellular ATP to supply to the nitrogenase enzyme. The LH4 operon was expressed at its highest levels under low light conditions.

We developed software to visually map cDNA reads onto the chromosomal map. Color-coding allows us to quickly visualize open-reading frame transcripts (the largest class of reads), anti-sense reads within genes, and the 5' untranslated regions of transcripts, allowing estimation of transcription start sites. Transcripts that map to intergenic regions in the opposite orientation from flanking genes are candidates for small trans-acting RNAs. Such a candidate sRNA was found in the nitrogenase gene cluster. Anti-sense reads in the 5' untranslated regions of genes may represent antisense cis-acting sRNAs. A candidate for this type of regulatory RNA is found in the region 5' of the LH4 operon.

Our results show that the NSR approach is an effective way to circumvent the problem of excessive rRNA reads, which - because rRNA is present in such overwhelming amounts in cells relative to other RNAs - will dominate sequence-based transcriptional analysis if permitted to do so. In addition the strand-specific sequencing information obtained using the NSR protocol has allowed us to identify potential cis- and trans-acting sRNAs that could constitute a previously unrecognized layer of regulatory control over hydrogen gas production.

Reference

1. Armour, C. D., J.C. Castle, R. Chen, T. Babak, P. Loerch, S. Jackson, J.K. Shah, J. Dey, C.A. Rohl, J.M. Johnson and C.K. Raymond (2009) Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. Nat Methods 6:647-649.

110

Genetic Manipulation of the Hyperthermophilic Hydrogen Producer, *Thermotoga maritima*

D. White,[1]* R. Singh,[1] K. Noll,[2] R. Kelly,[3] and **P. Blum**[1]

[1]Beadle Center for Genetics, University of Nebraska, Lincoln; [2]Dept. of Molecular and Cell Biology, University of Connecticut, Storrs; and [3]Dept. of Chemistry and Biomolecular Engineering, North Carolina State University, Raleigh

Project Goals: The anaerobic hyperthermophile, *Thermotoga maritima* (Tma), ferments carbohydrates to from molecular hydrogen (H_2) as one of its by-products. A high overall H_2 yield makes *T. maritima* a preferable model to evaluate hydrogen production. Our focus is to develop and use genetic methods to manipulate the *Tma* genome to investigate the metabolic flux of carbon and hydrogen synthesis.

The anaerobic hyperthermophile, *Thermotoga maritima (Tma),* ferments carbohydrates to from molecular hydrogen (H_2) as one of its by-products. A high overall H_2 yield makes *T. maritima* a preferable model to evaluate hydrogen production. Our focus is to develop and use genetic methods to manipulate the *Tma* genome to investigate the metabolic flux of carbon and hydrogen synthesis. A new *Tma* genetic marker has been developed by screening for uracil auxotrophs among spontaneous mutants resistant to the pyrmidine analog, 5-fluroroorotic acid (5-FOA). The *pyrE-64* mutant (strain PBL3001) arose by a two nt deletion (-TG) at chromosomal positions 351,539 (-T) and 351,538 (-G), 155 nt from the end of *pyrE*. This mutation results in a premature stop codon (TGA) 64 nt before the natural stop and therefore reduces protein length by 21 AA (from an original 187 AA). Auxotrophy was confirmed by demonstrating growth in a defined medium was dependent upon uracil supplementation. Stability of *pyrE-64* was evaluated by enrichment and characterization of gain-of-function prototrophic suppressors. The *pyrE-100* mutant (strain PBL3021) restored the *pyrE* reading frame by deletion of an additional one nt flanking the primary lesion. Additional genetic markers may also arise from studies on spontaneous novobiocin resistant isolates that target gyrase. Current efforts are underway to repair *pyrE-64* by directed recombination using a suicide vector and, by complementation using a *groESp*::*pyr*E promoter fusion fragment carried on a replicating shuttle vector based on a synthetic copy of pRQ7 fused to pUC19. Recombination strategies will target carbon catabolic pathways and components of the multiple hydrogenases while complementation strategies will be used to import new traits into *Tma*. Both genetic approaches will support collaborative efforts on transcriptomics, studies of the *Tma* toga and metabolic modeling.

Student Presentation

111

Carbohydrate Fermentation to Hydrogen by Hyperthermophilic *Thermotoga* Communities

A.D. Frock,[1]* S.R. Gray,[1] J.S. Notey,[1] K.M. Noll,[2] P.H. Blum,[3] and R.M. Kelly[1] (rmkelly@ncsu.edu)

[1]Dept. of Chemistry and Biomolecular Engineering, North Carolina State University, Raleigh; [2]Dept. of Molecular and Cell Biology, University of Connecticut, Storrs; and [3]Beadle Center for Genetics, University of Nebraska, Lincoln

Project Goals (Abstracts 111-112): Objective 1—Examine the regulation of substrate catabolic proteins and pathways as this relates to carbon partitioning, disposition of reducing power, and H_2 generation in *Thermotoga maritima*. Objective 2—Dissect catabolic and regulatory pathways using genetic approaches based on past success with other hyperthermophiles. Objective 3—Thermotogales biodiversity arises from adaptive specialization that expands on a conserved minimal genome; physiological characterization of selected novel traits will be done to expand understanding of biohydrogenesis.

Members of the hyperthermophilic bacterial genus *Thermotoga* are of special interest for biological hydrogen production due to high yields and the ability to use a broad range of complex carbohydrates. These high hydrogen yields are related to a more narrow range of fermentation products characteristic of hyperthermophiles compared to mesophiles. High temperature bioconversion also benefits from reduced risk of contamination and less recalcitrant biomass.

Comparison of several hyperthermophilic *Thermotoga* species (*T. maritima*, *T. neapolitana*, *T. petrophila*, and *T.* sp. RQ2) revealed that these species share a "core genome" of about 1500 genes, which generally includes most genes and pathways involved in central metabolism. Among genes involved in converting carbohydrates to hydrogen, those involved in carbohydrate degradation and transport appear to be more divergent among these species. In order to study species-specific characteristics and inter-species interactions in mixed culture, we have created a multi-species genus level cDNA microarray. This array is an expansion of a pre-existing *T. maritima* whole genome array, to which unique genes from *T. neapolitana, T. petrophila,* and *T.* sp. RQ2 were added.

Hydrogen production rate was measured for these four species in pure culture, as well as four-species mixed culture, during growth on a simple substrate (glucose) and a complex substrate (mix of seven different polysaccharides). Under these conditions there was no significant variation in growth rate, final cell density, or hydrogen production, although it should be noted that substrate consumption and hydrogen yield were not measured in these experiments.

The multi-species microarray also presents the possibility of following the evolution of the community structure for the mixed culture. Differentiation of these closely-related species by visual inspection was not possible nor were 16S rRNA approaches appropriate, given the 99%+ identity among these species. However, the multi-species microarray could be used to "count" each member species in the mixed culture. The first step for this method is identification of cDNA probes which are unique for one of the four species. Unique probes were predicted by BLAST against the four genomes, and the probes were experimentally tested by isolating genomic DNA from each of the species and hybridizing to the multi-species array. Initial results indicate that approximate enumeration of these species is possible provided that "unique" probes have been confirmed by microarray hybridization and that cDNA hybridized to the array is not present in spot-saturating quantities. This array is now being used to follow the evolution of the mixed community in batch and chemostat culture.

112

Functional Genomic Analysis of the Microbial Ecology of Hyperthermophilic *Thermotoga* Species

J.S. Notey,[1] S.R. Gray,[1] A.D. Frock,[1] K.M. Noll,[2] P.H. Blum,[3] and R.M. Kelly[1] (rmkelly@eos.ncsu.edu)

[1]Dept. of Chemistry and Biomolecular Engineering, North Carolina State University, Raleigh; [2]Dept. of Molecular and Cell Biology, University of Connecticut, Storrs; and [3]Beadle Center for Genetics, University of Nebraska, Lincoln

Project Goals: See goals for abstract 111.

The completed genome sequences of several hyperthermophilic *Thermotoga* species have given rise to functional genomic-based studies of the microbial ecology of these bacteria in pure and mixed cultures. Transcriptional response analysis of *T. maritima* growing synthrophically with the hyperthermophilic archaeon *Methanocaldococcus jannaschii* triggered quorum sensing-behavior and led to identification of a putative signaling peptide responsible for exopolysaccharide production (Johnson et al., 2006; Montero et al., 2006). Co-culture of *T. maritima* with another hyperthermophilic archaeon *Pyrococcus furiosus* led to identification to a genome locus containing putative bacteriocins and toxin-antitoxin loci (Montero, 2005; Gray et al., in preparation). Recent work focusing on multispecies cultures of hyperthermophilic *Thermotoga* species induced transcription of ORFs in this same locus, suggesting that this segment of the genome encoded genes important for ecological interactions. The genomes of *T. maritima*, *T. petrophila*, *T. sp.* RQ2, and *T. neapolitana*, as well as their transcriptomes in pure and mixed cultures, were analyzed to examine the similarities and differences among these bacteria with respect to microbial ecology. Also, experiments underway describe our current efforts to understand interspecies interactions in high temperature biotopes.

References

1. Johnson, M.R., S.B. Conners, C.I. Montero, C.J. Chou, K.R. Shockley, and R.M. Kelly. 2006. "The *Thermotoga maritima* Phenotype Is Impacted by Syntrophic Interaction with Methanococcus jannaschii in Hyperthermophilic Coculture. Appl Environ Microbiol 72 (1), 811–18.
2. Montero, C.I., D.L. Lewis, M.R. Johnson, S.B. Conners, E.A. Nance, J.D. Nichols, and R.M. Kelly. 2006. "Colocation of Genes Encoding a tRNA-mRNA Hybrid and a Putative Signaling Peptide on Complementary Strands in the Genome of the Hyperthermophilic Bacterium *Thermotoga maritima*. J Bacteriol 188 (19), 6802–07.

113

Systems-Level Understanding of *Thermotoga maritima* – The Transcriptional Architecture

DaeHee Lee[1]* (dlee@bioeng.ucsd.edu), Dmitry Rodionov,[2] Vasiliy Portnoy,[1] **Andrei Osterman**,[2] **Bernhard O. Palsson**,[1] and **Karsten Zengler**[1]

[1]Dept. of Bioengineering, University of California, San Diego; and [2]Burnham Institute for Medical Research, La Jolla, Calif.

Project Goals: This project is focused on a systems-level understanding of biological hydrogen production using *Thermotoga maritima* as a model organism. The project will address the basic science required to improve our understanding of hydrogen production from various carbon sources including glucose, cellulose, starch and xylan by this thermophilic microorganism. The overall goal is 1) to reconstruct the regulatory and metabolic network in *T. maritima* using various sets of "omics" data, 2) to integrate regulatory and metabolic networks into one "integrated" genome-scale model, 3) to confirm and validate the ability of the integrated model to predict processing of various environmental signals.

This project is focused on a systems-level understanding of biological hydrogen production using *Thermotoga maritima* as a model organism. We therefore have integrated a metabolic reconstruction of *T. maritima* that contains 479 metabolic genes, 565 metabolites (non-unique) and 646 internal and external metabolic reactions with 478 protein structures to generate the first three-dimensional reconstruction of the central metabolic network of a bacterium. To understand how the flow of information from the genome to the different states of the metabolic network is archived we studied the transcriptome architecture of *T. maritima.* A protocol for chromatin immunoprecipitation (ChIP) has been adapted to work for *T. maritima*. Furthermore, we developed a method that allows for the genome-wide determination of transcription start sites (TSSs) with a single base-pair resolution. Genome-wide transcription profiles using high-density tilled arrays were integrated with binding regions of RNA polymerase and TSS information to generate an experimentally verified map of the transcriptional landscape, laying the foundation of the experimentally elucidation of the operon structure in *T. maritima.* In addition we used an integrated approach to systematic annotation and reconstruction of transcriptional regulons in the available genomes of the Thermotogales. Two major components of this analysis are (i) annotation and propagation of previously known regulons from model organisms to others (e.g., arabinose regulon AraR, arginine regulon ArgR), and (ii) *ab initio* prediction of novel regulons (e.g., inositol regulon InoR, mannose regulon ManQ). In addition to playing a key role in regulon reconstruction, regulons provide an additional layer of genome context, helping to significantly improve the accuracy of functional annotations and metabolic reconstruction. Comparative analysis of the InoR regulon led to a discovery of a new pathway of inositol catabolism, which is currently under experimental investigation by in vitro enzymatic assays.

114

A Report on the Investigation of the Cell Envelope Proteins of *Thermotoga maritima*

Amanda K. Petrus[1]* (amanda.petrus@uconn.edu), Paul H. Blum,[2] Robert M. Kelly,[3] and **Kenneth M. Noll**[1]

[1]Dept. of Molecular and Cell Biology, University of Connecticut, Storrs; [2]Beadle Center for Genetics, University of Nebraska, Lincoln; and [3]Dept. of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh

Project Goals: We have undertaken an effort to purify the toga fraction of *T. maritima* cells and to identify the proteins that constitute the toga when cells grow on different substrates.

The outer envelope of a bacterium creates the interface that controls that organism's interactions with its environment and its utilization of carbon and energy sources. This cell envelope is a very important cell component when those sources are large and insoluble polymers, such as plant-derived polysaccharides. These bulky food sources require enzymatic degradation prior to being taken up into the cell. Thermophilic bacteria of the order Thermotogales have outer envelopes formed by a structure called the "toga" (1). The toga balloons over the cytoplasmic aspect of these cells forming a pronounced periplasmic space. The goal of this study is to characterize the proteins that compose this toga and to determine their roles in allowing cells to utilize complex, insoluble polysaccharides. These compounds can serve as a renewable energy source for the biohydrogenesis carried out by these cells.

Thermotoga maritima is the most extensively studied species of the Thermotogales. It has an optimal growth temperature of 77° C and can grow on a variety of simple and complex sugars, which lead to the production of carbon dioxide, hydrogen and acetic acid as the major products of fermentation. A complete genome sequence is available for *T. maritima* and several of its sugar hydrolases have been isolated and characterized (2). By contrast, there are only two major toga structural proteins currently identified, Ompα and

Ompβ, and little is known about their functions (3). No evidence has been reported of lipids in the outer envelope.

Ompβ is a porin protein that constitutes a large fraction of the toga (4). Ompα is a rod- shaped spacer protein that connects the outer envelope to the cell. Its carboxy terminus is hydrophobic and most likely anchored into the Ompβ layer. It also remains associated with that layer in the parts of the toga that dissociate from the cytoplasmic membrane (3). It is not clear if Ompα is attached by its amino terminus and, if so, whether it is attached to the cytoplasmic membrane or the peptidoglycan layer. In addition to these structural proteins, at least two sugar hydrolases, a xylanase (1) and an amylase (5), have been identified in the toga fraction.

We have undertaken an effort to purify the toga fraction of *T. maritima* cells and to identify the proteins that constitute the toga when cells grow on different substrates. We have successfully utilized a freeze-thaw/homogenization mechanical shearing method to selectively release toga proteins without major contamination by cytoplasmic proteins. Our efforts have resolved the fraction using 2-D electrophoresis and the resulting protein spots will be subjected to analysis by mass spectrometry, with the goal of identifying those proteins by referencing the annotated genome sequence. Though the genome sequence is available, the gene encoding Ompβ has yet to be identified. Ompα is encoded by ORF TM1729. The toga fraction has been resolved to fewer than twenty proteins on 2-D gels and a major spot migrating to a position consistent with a putative Ompβ ORF has been found. The toga fraction obtained in our study will be analyzed by mass spectrometry to identify the sequence of the Ompβ protein as well as other proteins associated with the toga.

References

1. Liebl, W., Winterhalter, C., Baumeister, W., Armbrecht, M., Valdez, M. 2008. Xylanase attachment to the cell wall of the hyperthermophilic bacterium Thermotoga maritima. J. Bacteriol. 190:1350-1358.
2. Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. L., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., Fraser, C.M. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature. 399:323-329.
3. Engel, A. M., Brunen, M., Baumeister, W. 1993. The functional properties of Ompβ, the regularly arrayed porrin of the hyperthermophilic bacterium *Thermotoga maritima.* FEMS Microbiol. Lett. 109:231-236.
4. Rachel, R., Engel, A. M., Huber, R., Stetter, K. O., Baumeister, W. 1990. A porin-type protein is the main constituent of the cell envelope of the ancestral bacterium *Thermotoga maritima.* FEBS Lett. 262:64-68.
5. Schumann, J., Wrba, A., Jaenicke, R., Stetter, K.O. 1991. Topographical and enzymatic characterization of amylases from the extremely thermophilic eubacterium *Thermotoga maritima*. FEBS Lett. 282:122-126.

115

Structure-Assisted Modeling of a Metabolic Interactome in *Thermotoga maritima*

Ying Zhang,[1] Marat Kazanov,[1] Xiaoqing Li,[1] Dmitry Rodionov,[1] **Karsten Zengler**,[2] **Bernhard O. Palsson**,[2] Kenneth Noll,[4] Adam Godzik,[1] and **Andrei Osterman**[1,3]* (osterman@burnham.org)

[1]Burnham Institute for Medical Research, La Jolla, Calif.; [2]Dept. of Bioengineering, University of California, San Diego, La Jolla; [3]Fellowship for Interpretation of Genomes, Burr Ridge, Ill.; and [4]Dept. of Molecular and Cell Biology, University of Connecticut, Storrs

Project Goals: This project is focused on a systems-level understanding of biological hydrogen production using *Thermotoga maritima* as a model organism. The project will address the basic science required to improve our understanding of hydrogen production from various carbon sources including glucose, cellulose, starch and xylan by this thermophilic microorganism. The overall goal is 1) to reconstruct the regulatory and metabolic network in *T. maritima* using various sets of "omics" data, 2) to integrate regulatory and metabolic networks into one "integrated" genome-scale model, 3) to confirm and validate the ability of the integrated model to predict processing of various environmental signals.

Genomics-based metabolic reconstruction technology strongly impacts fundamental understanding of cellular organisms and drives multiple applications in bioengineering. However, a shortcoming of a traditional metabolic reconstruction is its "structure-blindness" with respect to atomic-level interactions between metabolic enzymes and metabolites. A recent progress in high-throughput structure determination led to a nearly complete coverage of a relatively limited fold space of protein families that comprise metabolic networks of model bacteria, opening an opportunity to bridge a gap in structural understanding of *metabolic interactome*. This study extends our previously reported integration of genomics-based metabolic reconstruction of *T. maritima* with experimentally determined and computationally modeled 3D structures towards mapping of molecular interactions between metabolic proteins (enzymes, transporters, transcriptional regulators) and their cognate ligands (substrates, products, cofactors, effectors). Metabolite cross-docking appears to be one of the promising approaches to modeling such interactions, identification of ligand-binding sites and even prediction of specific ligands for proteins of unknown function. We have assessed publically available docking tools for their ability of accurate enzyme-substrate recognition. Cross-docking against a comprehensive set of small-molecule metabolites and comparative analysis of score distribution was performed for a panel of ~ 50 enzymes conserved and essential in the metabolic network of *T. maritima* (and most other bacteria). While showing some encouraging trends, this analysis revealed many limitations of a brute-force global cross-docking approach. Some of these limitations may be partially resolved by narrowing

down a set of compared enzymes and ligands as illustrated by the analysis of six groups of FGGY sugar kinase family with distinct substrate specificities. Representatives of these groups were identified in *T. maritima* and experimentally characterized in our previous study (reported at DOE-GTL, 2009). A comparative analysis of substrate (sugars) and product (sugar-phosphates) ranking allowed us to identify conformations that improved or impaired "dockability" in a panel of experimental 3D structures and homology-based 3D models. Combining limited structural data with massive comparative sequence analysis allows us to accurately map Specificity Determining Residues (SDR) in large protein families with variations in substrate specificity. This is also illustrated by the example of FGGY sugar kinase family where >800 proteins from hundreds microbial genomes were analyzed by a modified mutual information-based method of Gelfand and Mirny. This analysis revealed a combination of divergent and convergent scenarios in the evolution of substrate specificity within this large and functionally versatile family. A similar approach was successfully applied for the analysis of specificity evolution in a family of transcriptional regulators from ROK family represented by six proteins from *T. mariitima* that were identified and characterized in this project. Structure-assisted identification of functional sites in enzymes and other metabolic proteins is expected to improve accuracy of assignment, cross-genome projection and prediction of previously unknown gene functions. To capture and provide access to this and other types of information about genes and proteins supporting evolutionary and systems-level analysis of *T. maritima,* we initiated a development of the WIKITOGA web site. This web site will integrate several types of automated annotations (structural, functional, regulatory) and modeling tools with Wikipedia-style community contribution. Its content will be gradually extended from the initial focus on metabolic and regulatory networks of *T. maritima* towards whole-genome analysis of all Thermotogales species.

116

Using Hydrogen Isotopes to Assess Proton Flux during Biological Hydrogen Production 1. Determining Fractionation Factors and the Proton Transfer Pathway in Hydrogenases

Helen W. Kreuzer[1]* (Helen.Kreuzer@pnl.gov), Hui Yang,[2] Liang Shi,[1] Nathaniel E. Ostrom,[3] Hasand Gandhi,[3] and **Eric L. Hegg**[2] (EricHegg@msu.edu)

[1]Pacific Northwest National Laboratory, Richland, Wash.; [2]Dept. of Biochemistry and Molecular Biology and [3]Dept. of Zoology, Michigan State University, East Lansing

Project Goals: To improve our understanding of biological H_2 production using H/D isotope ratios.

Biological H_2 production by hydrogenase enzymes has enormous potential as an environmentally sustainable source of energy. Hydrogenases, found throughout nature in many diverse organisms, are among the most efficient H_2-producing catalysts known. Although considerable progress has been made in elucidating the metabolic pathways involved in H_2 metabolism, many uncertainties remain. One major impediment to improving our understanding of H_2 metabolism is our inability to adequately define the regulation of and the flux through key pathways involved in H_2 production. Thus far, few attempts have been made to utilize hydrogen isotopes to improve our knowledge of the H_2 metabolic pathways, perhaps because the source of protons for hydrogenase enzymes is intracellular water. Until recently, intracellular water was generally assumed to be isotopically equivalent to extracellular water, and therefore it was perhaps thought that hydrogen isotopes would not be informative. We are exploiting our recent discoveries that intracellular water can be isotopically distinct from extracellular water, and that the contribution of protons from metabolic substrates to intracellular water can be quantified, to develop the use of hydrogen isotopes for studying intracellular proton trafficking.

We predicted that the isotope ratio of H_2 produced by various hydrogenases would differ because of slight differences in the active sites and proton transfer pathways. We further predicted that we can measure this difference via isotope-ratio mass spectrometry, and that the H/D isotope ratios would allow us to address fundamental questions concerning biological H_2 production including the source of the H_2. To test this predictions, we purified five different hydrogenases (three [FeFe]-H_2ases and two [NiFe]-H_2ases) and established conditions that allowed us to quantify the specific activity of the purified H_2ases and the amount of H_2. In addition, we built a custom chromatographic system for the analysis of H_2 and interfaced this system with an isotope ratio mass spectrometer (IRMS). We obtained a reproducibility of better than 3‰ for δD, and this precision is maintained down to a lower sensitivity limit of 0.2 μmol H_2 in 1 mL of headspace volume. In addition, because H_2 diffuses so readily through most materials (resulting in isotopic fractionation), it was necessary to develop and validate a robust protocol for capturing biologically-produced H_2. Subsequently, reaction conditions were developed that allow for the reproducible formation and capture of the optimal concentration of H_2.

Using the enzymes and optimized protocols established above, we determined the isotope ratio of the H_2 produced by three different [FeFe]-H_2ases (*Clostridium pasteurianum, Shewanella oneidensis,* and *Chlamydomonas reinhardtii*) and two [NiFe]-H_2ases (*Shewanella oneidensis* and *Desulfovibrio fructosovorans*). Significantly, the data indicate that all 5 hydrogenases produce H_2 with a unique isotopic signature. This proves our initial hypothesis, that different H_2-producing enzymes have different fractionation factors, and that these differences are reflected in the isotope ratio of the H_2.

Building on these results, we are using this data to help elucidate the proton transfer pathway in [FeFe]-H_2ases. To accomplish this task, we mutated a number of residues proposed to be critical for proton transport in the *C. pasteurianum* enzyme. Mutations that affect the proton pathway will change the fractionation factor by changing the ener-

getics of proton vs. deuteron migration, while mutations that affect H_2ase activity by other mechanisms will not change the fractionation factor. We have generated nine variants and all have been shown to alter H_2ase activity. Future studies will ascertain if these mutations also alter the H/D isotope ratio of the H_2.

In the second phase of the project, we are also utilizing the experimentally-determined H_2ase fractionation factors for in vivo studies to identify the major H_2-forming pathway in *S. oneidensis* under a variety of growth conditions (see accompanying poster).

117

Using Hydrogen Isotopes to Assess Proton Flux during Biological Hydrogen Production 2. In Vivo Studies with *Shewanella oneidensis*

Eric A. Hill,[1] James J. Moran[1]* (James.Moran@pnl.gov), **Eric L. Hegg**,[2] and **Helen W. Kreuzer**[1]

[1]Pacific Northwest National Laboratory, Richland, Wash. and [2]Dept. of Biochemistry, Michigan State University, East Lansing

Project Goals: To improve our understanding of biological H_2 production using H/D isotope ratios.

Biological hydrogen production by hydrogenase enzymes represents a potentially sustainable, non-polluting source of energy. In order to fully exploit the capabilities of biological systems for hydrogen production, a deeper understanding of regulation of and fluxes through key pathways involved in hydrogen production is needed. We have developed methods using stable isotope measurements that allow us to trace protons from extracellular water and metabolic substrates into intracellular water and cellular metabolites, including hydrogen gas. In this part of our project, we are applying those methods to elucidate hydrogen production pathways in vivo, using *S. oneidensis* as a model organism.

Shewanella oneidensis MR-1 is a facultative anaerobe capable of transferring electrons to a variety of terminal acceptors including iron, manganese, and other metals. *S. oneidensis* encodes two hydrogenase enzymes, the [FeFe]-hydrogenase HydA and the [NiFe]-hydrogenase HyaB. Hydrogenases catalyze the reversible reaction of protons plus electrons to form hydrogen gas. If electron acceptors in its growth environment are limited, *S. oneidensis* reduces protons, producing hydrogen gas.

The objectives of this portion of our project are to determine:

- **What is the contribution of each hydrogenase enzyme to hydrogen production?** When the organism is perturbed, does the manipulation differentially affect the flux through each hydrogenase?
- **Is there channeling of protons between organic substrates and hydrogenases?** Although it is well established that the addition of organic substrates increases H_2 production under certain conditions (e.g. acetate to the growth media of green algae or glucose to the growth media of cyanobacteria), the precise mechanism by which this occurs is not entirely clear. Some have suggested that protons are directly channeled from specific substrates into hydrogen production. We will test this hypothesis.

We have characterized hydrogen production by wild-type and electron transfer-deficient *S. oneidensis* strains. When the wild-type organism is cultured in sealed headspace vials with limited electron acceptors, the headspace hydrogen concentration initially increases, then decreases, then steadily increases. In strains deficient in metal reduction, we do not observe the decrease in headspace H_2 concentration, demonstrating that the electron transport deficiency impacts hydrogenase activity.

We hypothesize that we will be able to use stable isotopes to dissect proton fluxes through the two hydrogenase enzymes under these and other culture conditions. We have determined that HydA and HyaB evolve isotopically distinct hydrogen from the same substrate water (presented in a companion poster). In this, the second phase of our project, we are measuring the stable isotope content of the intracellular water that presumably is the substrate for hydrogenase activity, as well as that of the hydrogen gas produced in vivo, in the wild-type and mutant strains. Similar measurements will enable us to dissect proton trafficking in the presence of different organic substrates.

118

Systems Biology of Hydrogen Regulation in *Methanococcus maripaludis*

Qiangwei Xia,[1] Erik L. Hendrickson,[1] Sung Ho Yoon,[2] Thomas J. Lie,[3] **William B. Whitman**,[4] **Murray Hackett**,[1] **Nitin Baliga**,[2] and **John A. Leigh**[3]* (leighj@u.washington.edu)

[1]Dept. of Chemical Engineering, University of Washington, Seattle; [2]Institute for Systems Biology, Seattle, Wash.; [3]Dept. of Microbiology, University of Washington, Seattle; and [4]Dept. of Microbiology, University of Georgia, Athens

http://maggie.systemsbiology.net
http://faculty.washington.edu/leighj/

Project Goals: 1. Use transcriptomics, proteomics, and metabolomics to study the systems biology of H_2 metabolism, formate metabolism, nitrogen fixation, and carbon assimilation in *Methanococcus maripaludis*. 2. Determine the mechanism of H_2 sensing and transcriptional regulation by H_2.

Background

We are engaged in a long-term effort to understand regulatory networks in hydrogenotrophic methanogens, members

of the Archaea whose energy metabolism specializes in the use of H_2 to reduce CO_2 to methane. (Many hydrogenotrophic methanogens can use formate as an alternative to H_2 and CO_2). Our studies focus on *Methanococcus maripaludis*, a model species with good laboratory growth characteristics, facile genetic tools, and a tractable genome of 1722 annotated ORFs. Much of our work to date has focused on the response that occurs when supplies of essential nutrients are decreased to growth-limiting levels. Thus, we have studied the responses to H_2 limitation, nitrogen limitation, phosphate limitation, and leucine limitation (using a leucine auxotroph) (1-3). A key aspect of our approach is the use of continuous culture for maintaining defined nutrient conditions (4).

Proteomics

We used high-coverage quantitative proteomics to determine the response of *M. maripaludis* to growth-limiting levels of H_2, nitrogen, and phosphate (1). Six to ten percent of the proteome changed significantly with each nutrient limitation. H_2 limitation increased the abundance of a wide variety of proteins involved in methanogenesis. However, one protein involved in methanogenesis decreased: a low-affinity [Fe] hydrogenase, which may dominate over a higher-affinity mechanism when H_2 is abundant. Nitrogen limitation increased known nitrogen assimilation proteins. In addition, the increased abundance of molybdate transport proteins suggested they function for nitrogen fixation. An apparent regulon governed by the euryarchaeal nitrogen regulator NrpR was identified. Phosphate limitation increased the abundance of three different sets of proteins, suggesting that all three function in phosphate transport. The global proteomic response of *M. maripaludis* to each nutrient limitation suggests a wider response than previously appreciated. The results give new insight into the function of several proteins, as well as providing information that should contribute to the formulation of a regulatory network model.

Five different approaches were compared for measuring protein abundance ratios (5). The results suggest that at the limit of deep sampling, frequency based measurements are competitive with metabolic stable isotope labeling in terms of power to detect abundance change. In addition, false discovery rates and local false discovery rates were compared as complementary approaches to multiple hypothesis testing for quantitative significance. These findings will be discussed in detail in a poster by M. Hackett.

Transcriptomics and metabolomics

A tiling array was designed and used to measure gene expression changes along a growth curve. An initial transcriptome map has been constructed and is currently being hand-annotated. The results will be discussed in detail in a poster by S.H. Yoon and N. Baliga. In addition, methods are being worked out for the measurement of key metabolites.

Hydrogen regulation and metabolism

mRNA levels for key enzymes of methanogenesis are regulated by H_2 availability (3). Results to be presented suggest that this regulation relies on sensing of some intracellular redox indicator, rather than on the external H_2 concentration.

We are also investigating whether H_2 is a necessary intermediate during growth on formate (6). This is of interest in the context of the potential for H_2 production by nitrogenase during growth on formate. If H_2 is not a necessary intermediate, then we should be able to eliminate enzymes that would deplete H_2 that is produced by nitrogenase. Experiments are underway to test the essentiality of genes whose products could produce H_2 during growth on formate.

This research was supported by the Office of Science (BER), U.S. Department of Energy, Award No. DE-FG02-08ER64685.

References

1. Xia Q, *et al.* (2009) Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen *Methanococcus maripaludis*. *BMC Microbiol* 9:149.
2. Hendrickson EL, *et al.* (2008) Global responses of *Methanococcus maripaludis* to specific nutrient limitations and growth rate. *J Bacteriol* 190:2198-2205.
3. Hendrickson EL, Haydock AK, Moore BC, Whitman WB, and Leigh JA (2007) Functionally distinct genes regulated by hydrogen limitation and growth rate in methanogenic Archaea. *Proc Natl Acad Sci U S A* 104:8930-8934.
4. Haydock AK, Porat I, Whitman WB, and Leigh JA (2004) Continuous culture of *Methanococcus maripaludis* under defined nutrient conditions. *FEMS Microbiol Lett* 238:85-91.
5. Xia Q, Wang T, Beck DAC, Taub F, Leigh JA, and Hackett M (2010) Quantitative local false discovery rates, deep sampling and detecting protein abundance change for the model organism *Methanococcus maripaludis*. *Proteomics* (in review).
6. Lupa B, Hendrickson EL, Leigh JA, and Whitman WB (2008) Formate-dependent H_2 production by the mesophilic methanogen *Methanococcus maripaludis*. *Appl Environ Microbiol* 74:6584-6590.

119

Quantitative Local False Discovery Rates, Deep Sampling and Protein Abundance Change for *Methanococcus maripaludis*

Qiangwei Xia,[1,2,4] Tiansong Wang,[2] David A.C. Beck,[3] Fred Taub,[2] **John A. Leigh**,[2] and **Murray Hackett**[1]* (mhackett@u.washington.edu)

[1]Dept. of Chemical Engineering, [2]Dept. of Microbiology, and [3]eScience Institute, University of Washington, Seattle; and [4]Dept. of Chemistry, University of Wisconsin, Madison

Project Goals: 1. To find the best balance between statistical power to detect protein abundance change and the need to decrease the time required for a complete proteome analysis for Methanococcus maripaludis. 2. To establish just how much sampling is required for spectral counting to become as efficient or perhaps even more efficient for generating protein abundance ratios relative to traditional metabolic stable isotope labeling. The answer to this question is heavily dependent on the specific mass

spectrometry instrumentation and methods used in the investigation, both of which have experienced several upgrades and improvements since we last addressed this question in the peer reviewed literature in 2006. 3. To establish the most efficient data reduction and transformation procedures for the use of local false discovery rates (lfdr) with spectral counting data. Lfdr is one approach among several to multiple hypothesis testing, based on the more general concept of false discovery rate. Although proposed initially in the context of microarray-based transcriptome data, lfdr is equally applicable to large-scale quantitative proteomic experiments and RNA-Seq.

Protein abundance ratios were measured using five different approaches for the Archaeon *Methanococcus maripaludis*, a model organism of interest for analytical studies because of the tractable nature of its proteome in terms of size, proteome extraction efficiency and other positive features. Multidimensional capillary HPLC coupled with tandem mass spectrometry was used for analysis of heavy (^{15}N) and natural abundance (^{14}N) tryptic digests of *M. maripaludis* grown in chemostats. Here we report our comparison of abundance ratios based on heavy and light proteomes mixed prior to mass spectrometry; spectral counting of heavy and light proteomes mixed; spectral counting of heavy and light proteomes analyzed separately; summed signal intensities for mixed heavy and light proteomes; and summed signal intensities for heavy and light proteomes analyzed separately. Protein identifications were saturated and proteome penetration maximized at ~ 91% of the predicted protein-encoding open reading frames. False discovery rates and local false discovery rates were compared as complementary approaches to multiple hypothesis testing for quantitative significance. Power calculations, dynamic range and other observations reported suggest that at the limit of deep sampling frequency based measurements are competitive with metabolic stable isotope labeling in terms of power to detect abundance change.

This research was supported by the Office of Science (BER), U.S. Department of Energy, Award No. DE-FG02-08ER64685.

120

Development of Metabolic Network Models of *Rhodobacter sphaeroides* for the Prediction of Quantitative Contributors to H_2 Production

L. Safak Yilmaz[1]* (yilmaz@cae.wisc.edu), Timothy J. Donohue[2] (tdonohue@bact.wisc.edu), and **Daniel R. Noguera**[1] (noguera@engr.wisc.edu)

[1]Dept. of Civil and Environmental Engineering and [2]Dept. of Bacteriology, University of Wisconsin, Madison

Project Goals: 1) Develop a metabolic network model of *Rhodobacter sphaeroides*. 2) Use this model to (a) understand the bioenergetics of growth conditions that produce hydrogen from different carbon sources, so that optimal conditions of hydrogen production can be predicted and (b) determine reactions and pathways that detract electrons from hydrogen production, so that hydrogen production can be potentially maximized by genetic manipulations.

Introduction and Objective

The metabolically versatile organism *Rhodobacter sphaeroides* produces H_2 while using light as an energy source and organic substrates as an electron donor. The overall goal of this project is to elucidate gene networks that contribute to and detract from H_2 production and genetically modify *R. sphaeroides* to divert as high a fraction of substrate electrons as possible to H_2 generation. As a tool to achieve this goal, we are developing metabolic network models of *R. sphaeroides* which can integrate genomic knowledge and experimental data, thereby allowing the prediction of optimization strategies.

Methodology

Stoichiometric reaction network models (Palsson, 2006) were reconstructed using the KEGG database (http://www.genome.jp/kegg/pathway.html), the genome of *R. sphaeroides* strain 2.4.1 (http://www.rhodobacter.org/), and other metabolic network models available in the literature (http://gcrg.ucsd.edu/In_Silico_Organisms). Model development is divided into three stages based on the scale of metabolic pathways represented, as explained below.

Stage 1 activities built a core network model that represented central carbon metabolism. Carbon pathways including glycolysis, citric acid cycle, and the pentose-phosphate pathway are connected to an electron transport chain previously proposed for purple non-sulfur bacteria (Klamt et al., 2008). The energetic contribution of light is part of the electron transport chain, H_2 production is added as a nitrogenase reaction in the absence of N_2 based on available data (see poster by Kontur et al.). The synthesis of the known electron sink polyhydroxybutyrate (PHB) is part of the network as a pathway starting from acetyl coenzyme A. To simulate cell synthesis, precursor metabolites are assembled in a biomass reaction taken from *Escherichia coli* core network models (Palsson, 2006).

Ongoing activities in *Stage 2* expands the Stage 1 model to include amino acid and CO_2 fixation pathways. The existing biomass equation (Stage 1) will also be modified based on the utilization of precursor metabolites in the new pathways and on the average protein composition of purple non-sulfur bacteria (Kobayashi and Kobayashi, 1995) to include amino acids as part of biomass.

A planned *Stage 3* model will add other major metabolic pathways (lipid metabolism, nucleic acid metabolism, glycan biosynthesis, and metabolism of cofactors) to the Stage 2 model. Biomass assembly will be based on building blocks using average cell composition of purple non-sulfur bacteria (Kobayashi and Kobayashi, 1995).

Once reaction networks are constructed, the models are established in MATLAB (MathWorks, Natick, MA) and GAMS (GAMS Development Corporation, Washington, DC) to carry out in silico flux balance analysis (FBA) with

linear programming. The models are trained and tested with experimental results from a systematic analysis of electron flow in *R. sphaeroides* growing photosynthetically on a variety of carbon sources (Yilmaz et al., in review). The data include H_2 generation, PHB synthesis, biomass formation, and production of soluble microbial products (SMP) during exponential and stationary phases of batch cultures, and similar data is being generated from chemostats in ongoing work. All cultures are fed with glutamate as the sole nitrogen source to maximize H_2 production, while single organic acids (succinate, lactate, pyruvate, and fumarate) or sugars (mostly glucose) are used as the carbon source.

Results

The Stage 1 model qualitatively captures aerobic growth without H_2 production and anoxygenic photosynthetic growth with H_2 production. The model predicts that H_2 production is a necessary electron accepting pathway in photosynthetic growth, without which excess electrons cannot be recycled. Another surprising prediction is that, SMP production, which was a significant electron sink in experimental cultures (Yilmaz et al., in review), occurs during dimmer light conditions (i.e. with a limited flux of the light reaction), possibly representing shadowing in dense cultures. The FBA results always predict a larger CO_2/H_2 ratio (≥1.0 in partial pressure) than what is found in the headspace of batch cultures (~0.2). This difference is attributed to the lack of CO_2 fixation pathways in our Stage 1 model. Although unconstrained simulations do not quantitatively match experimental data, constrained FBA fits to experimental results for the production of key electron sinks.

The Stage 1 model lacks pathways to simulate glutamate consumption as a carbon, nitrogen and electron source, but the Stage 2 can account for this due to addition of amino acid pathways. This addition and the introduction of CO_2 fixation pathways to the Stage 2 model are expected to provide quantitative and qualitative predictions for the mechanisms of growth and H_2 production as a function of carbon sources. The databases used were checked for the absence of missing reactions in amino acid and carbon fixation pathways before we began reconstruction of the Stage 2 model. A major improvement in the Stage 3 model will be the ability to include a global analysis of metabolism with FBA and existing microarray data. To this end, we developed a prototype visualization tool that overlays flux and gene expression profiles to KEGG pathway maps. The poster will present results from the Stage 1 model, the current Stage 2 model, and illustrations from the visualization tool for integrated microarray and flux data.

References

1. Klamt, S, Grammel, H, Straube, R, Ghosh, R, and Gilles, E D (2008) Modeling the electron transport chain of purple non-sulfur bacteria. Molecular Systems Biology. 4:156.
2. Kobayashi, M, and Kobayashi, M (1995) Waste Remediation and Treatment Using Anoxygenic Phototrophic Bacteria. In: Blankenship, R E, Madigan, M T, and Bauer, C E (ed), Anoxygenic Photosynthetic Bacteria. Kluwer Academic Publishers Dordrecht, Netherlands, pp 1269-1282
3. Palsson, B O. 2006. Systems Biology. Cambridge University Press, New York, NY.
4. Yilmaz, L S, Kontur, W, Sanders, A P, Sohmen, U, and Donohue, T J. Electron Partitioning During Light- and Nutrient-Powered Hydrogen Production by *Rhodobacter sphaeroides*. Bioenergy Research. In review.

121

Cellular Redox Balance and the Integrative Control of Carbon Assimilation and Hydrogen Production in *Rhodobacter sphaeroides* and *Rhodopseudomonas palustris*

Rick A. Laguna* (laguna.2@osu.edu) and **F. Robert Tabita**

Dept. of Microbiology, Ohio State University, Columbus

Project Goals: Elucidate molecular mechanisms by which evolved strains of NSP bacteria derepress the synthesis of nitrogenase, allowing this enzyme system to catalyze the production of high levels of hydrogen gas. Perform computational and molecular modeling studies to identify further control circuits that might be manipulated to improve hydrogen production. Determine how these different regulatory networks are integrated in the cell. Determine how various regulatory protein complexes we have implicated contribute towards regulating hydrogen production. Examine the intracellular organization of these complexes and determine how small effector metabolites influence their function. Integrate these studies to engineer the most efficient strain for maximal hydrogen production.

Nonsulfur purple (NSP) photosynthetic bacteria are characterized by their metabolic versatility. These organisms are capable of growth during photosynthetic and non-photosynthetic conditions, in the absence or presence of oxygen, respectively, and they can synthesize cell mass via the assimilation of either organic or inorganic carbon sources, with needed energy obtained via photochemical or dark chemical processes. During aerobic chemolithoautotrophic and anaerobic photolithoautotrophic growth conditions, CO_2 serves as the sole carbon source and is reduced into cellular carbon by the Calvin-Benson-Bassham (CBB) reductive pentose phosphate pathway. Under photoheterotrophic growth conditions, organic carbon compounds, usually organic acids, are oxidized into cell mass. In NSP bacteria such as *Rhodobacter sphaeroides* and *Rhodopseudomonas palustris*, when malate is used as the carbon source during photoheterotrophic growth, the reduction of metabolically produced CO_2 through the CBB cycle is the preferred means by which these organisms consume excess reductant produced from the assimilation of the organic electron donor (1-3). However, when the CBB cycle is inactivated, these organisms face the problem of either not growing or adapting by using other redox balancing mechanisms. In one interesting adaptive scenario, excess reductant is consumed by the derepression of the nitrogenase complex, resulting in the production of hydrogen gas in both *Rb. sphaeroides*

and *Rps. palustris*, as well as *Rhodospirillum rubrum* and *Rb. capsulatus* (4,5). This is of interest considering that malate is more oxidized than cell mass. Therefore the need to consume excess reductant through the CBB cycle should not be required. However upon examining the assimilation of malate in *Rb. sphaeroides* and *Rps. palustris* (both assimilate malate through the tricarboxylic acid pathway) and by using various CBB mutant strains, it has been shown that excess reductant is produced and must be consumed by the CBB cycle in order to maintain cellular redox balance. The reason for this is that as malate is assimilated into cell mass in both organisms, excess reductant is generated. Acetate can also be used as a carbon source for photoheterotrophic growth in *Rb. sphaeroides* and *Rps. palustris*. Unlike malate, acetate is at the same oxidation state as cell mass. Therefore excess reductant should not be generated by the assimilation of acetate and the CBB cycle should not be required. In *Rb. sphaeroides* this is the case; however in *Rps. palustris* there is a need for the CBB cycle. The reason for this is that the two organisms use different metabolic pathways to assimilate acetate. *Rb. sphaeroides* uses the ethylmalonyl-CoA pathway to assimilate acetate (6), which does not generate excess reductant from this process. Therefore CBB mutant strains of *Rb. sphaeroides* are capable of photoheterotrophic growth using acetate as the carbon source. However *Rps. palustris* uses the glyoxylate pathway to assimilate acetate (7), which does produce excess reductant. Therefore CBB mutant strains of *Rps. palustris* cannot grow photoheterotrophically with acetate as the carbon source. We are currently examining how the assimilation of different carbon sources affects redox balance and hydrogen production in *Rb. sphaeroides* and *Rps. palustris*, as well as *R. rubrum* and *Rb. capsulatus*, during photoheterotrophic growth conditions. Indications are that there are differences that might be eventually exploited to maximize hydrogen production.

References

1. Falcone, D. L. and F. R. Tabita. 1991. Expression of endogenous and foreign ribulose 1,5-bisphosphate carboxylase-oxygenase (RubisCO) genes in a RubisCO deletion mutant of Rhodobacter sphaeroides. J. Bacteriol. 173:2099-2108.
2. Romagnoli, S. and F. R. Tabita. 2006. A novel three-protein two-component system provides a regulatory twist on an established circuit to modulate expression of the cbbI region of Rhodopseudomonas palustris CGA010. J. Bacteriol. 188:2780-2791.
3. Joshi, G. S., Romagnoli, S., VerBerkmoes, N. C., Hettich, R. L., Pelletier, D., and Tabita, F. R. 2009. Differential accumulation of form I RubisCO in Rhodopseudomonas palustris CGA010 under photoheterotrophic growth conditions with reduced carbon sources. J. Bacteriol. 191:4243-4250.
4. Joshi, H. M. and F. R. Tabita. 1996. A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. Proc. Natl. Acad. Sci. U. S. A. 93:14515-14520.
5. Tichi, M. A., and Tabita, F.R. 2000. Maintenance and control of redox poise in Rhodobacter capsulatus strains deficient in the Calvin-Benson-Bassham pathway. Arch. Microbiol. 174:322-333.
6. Alber, B.E., Spanheimer, R., Ebenau-Jehle, C., and Fuchs, G. 2006. Study of an alternate glyoxylate cycle for acetate assimilation by Rhodobacter sphaeroides. Mol. Microbiol. 61:297-309.
7. Larimer, F.W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Land, M.L., Pelletier, D.A., Beatty, J.T., Lang, A.S., Tabita, F.R., Gibson, J.L., Hanson, T.E., Bobst, C., Torres, J.L., Peres, C., Harrison, F.H., Gibson, J., Harwood, C.S. 2004. Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris. Nature Biotechnol. 22:55-61.

122

Development of *Cyanothece* as a New Model Organism for Photobiological Hydrogen Production

Anindita Banerjee,[1] Jana Stöckel,[1] Lawrence Page,[1] Xueyang Feng,[2] Bing Wu,[2] Yinjie Tang,[2] Hongtao Min,[3] Sujata Mishra,[3] Xiaohui Zhang,[3] Louis A. Sherman[3]* (lsherman@purdue.edu), Stephen J. Callister,[4] Richard D. Smith,[4] and **Himadri B. Pakrasi**[1,2]

[1]Depts. of Biology and [2]Energy, Environmental and Chemical Engineering, Washington University, St. Louis, Mo.; [3]Dept. of Biological Sciences, Purdue University, West Lafayette, Ind.; and [4]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Wash.

sysbio.wustl.edu/Pakrasi/projects/hydrogen.php

Project Goals: The objective of this proposal is to develop the cyanobacterium *Cyanothece* as a model organism for photobiological hydrogen production. Members of the genus *Cyanothece* are unicellular oxygenic prokaryotes with the ability to fix atmospheric nitrogen. Our long-term goal is to develop a deep understanding of the metabolism of these microbes as it pertains to H_2 evolution. Specifically, we are using genome sequencing, transcriptomics, proteomics, metabolomics, mutagenesis, biochemical analysis and physiological approaches, all of which are encased in a systems biology framework.

Hydrogen production by *Cyanothece*: *Cyanothece* are unicellular, nitrogen-fixing cyanobacteria. They possess both the nitrogenase and hydrogenase enzyme systems that can catalyze biological hydrogen production. Hydrogen production by different *Cyanothece* strains was assessed by employing appropriate physiological perturbations, based on knowledge from prior transcriptomic and proteomic studies. *Cyanothece* sp. ATCC 51142 exhibited exceptionally high rates of hydrogen production under aerobic conditions. This is in striking contrast to other known unicellular photosynthetic hydrogen producing strains, which produce hydrogen under largely anaerobic conditions. The rates of hydrogen production were significantly enhanced by growing *Cyanothece* cells in the presence of glycerol or high levels of CO_2. Both of these carbon sources enhanced the level of glycogen, an intracellular energy reserve in the cells. Our studies revealed that the high rate of hydrogen production by this strain is

* Presenting author

largely mediated by an efficient nitrogenase enzyme system. Programmable photobioreactors equipped with sensors to monitor several critical parameters are now being used to further characterize these strains to achieve higher rates of hydrogen production.

Comparative genomics: Complete genome sequences of five *Cyanothece* strains (ATCC 51142, PCC 7424, PCC 7425, PCC 8801, PCC 8802) are currently available and two more (PCC 7822, ATCC 51472) are in the process of completion at the DOE Joint Genome Institute. The sequences reveal significant metabolic diversity within this group of cyanobacteria. The genome sequence information is being used to generate a *Cyanothece* pan-genome (in collaboration with JGI), comprising of the "core genome" (containing all of the genes common to each genus member) and the "dispensable genome" (containing unique genes or genes shared between two or more strains). The unique genes are likely to confer strain-specific attributes and will be analyzed for their role in hydrogen production.

Genetics: *Cyanothece* 7822 was successfully transformed by electroporation using a modification of the asymmetric PCR technique originally developed for eukaryotic algae. The first target was *nifK*; the gene encoding one of the two subunits of the nitrogenase MoFe protein and a spectinomycin-resistance cassette was inserted in the middle of the *nifK* gene. This strain has been stable for more than 9 months and is resistant to spectinomycin and is incapable of growing on plates or in liquid medium that is lacking combined nitrogen (*i.e.*, it cannot fix atmospheric N_2). The strain is also completely unable to produce hydrogen under any growth condition, although it can still reduce acetylene at rates that are particularly high when the cultures are incubated under argon prior to measurement. This indicates that the strain can still assemble a MoFe complex, although it cannot produce the main products of N_2 fixation, H_2 and NH_3.

Metabolomic Studies: Over the past year, we have developed new metabolite extraction methods and mass spectrometry techniques. Using such new approaches, we characterized the central metabolic pathways in *Cyanothece* 51142, and discovered a novel isoleucine biosynthesis pathway that involves citramalate synthase. We have studied impacts of carbon and nitrogen sources on the central metabolism and hydrogen production by *Cyanothece* 51142, using both biochemical methods and ^{13}C isotopomer approaches. We quantitatively determined CO_2 fixation and carbon substrate utilization under mixotrophic growth conditions in *Cyanothece* 51142. We are currently developing ^{13}C-assisted dynamic flux models to study autotrophic metabolism in cyanobacteria.

Proteomic Studies: In preparation for conducting relative quantitative proteomics analyses, using the AMT tag proteomics approach, construction of reference peptide databases for 6 of 7 *Cyanothece* strains has been completed. Over 460 LC-MS/MS datasets have been generated and analyzed using the high-throughput proteomics capabilities at PNNL. Complete databases correspond to strains *Cyanothece* sp. PCC 8801, PCC 8802, PCC 7424, PCC 7425, PCC 7822, and ATCC 51142. LC-MS/MS datasets for *Cyanothece* sp. ATCC 51472 have also been generated with identification of peptide sequences waiting upon the draft completion of the genome sequence (at the Joint Genome Institute) of this strain. Percent observed coverage of predicted proteins from unique peptides (10% false discovery rate) ranges from roughly 40% to 70%, which inversely correlates to the size range of genome sequences; *Cyanothece* 7424 has the largest genome sequence (~6.5 Mb) and smallest percent observed coverage among the strains.

Project number - DE-FC02-07ER64694

123

Genome-Wide Network Analysis of Metabolism in *Chlamydomonas reinhardtii*

Roger L. Chang[1]* (rlchang@ucsd.edu), Ani Manichaikul,[2] Balaji Santhanam,[3] Lila Ghamsari,[3] Erik F.Y. Hom,[4] **Kourosh Salehi-Ashtiani**,[3] and **Jason A. Papin**[2]

[1]Dept. of Bioengineering, University of California, San Diego; [2]Dept. of Biomedical Engineering, University of Virginia, Charlottesville; [3]Center for Cancer Systems Biology and Dept. of Cancer Biology, Dana-Farber Cancer Institute, and Dept. of Genetics, Harvard Medical School, Boston, Mass.; and [4]Dept. of Molecular and Cellular Biology, Harvard University, Cambridge, Mass.

http://bme.virginia.edu/csbl
http://ccsb.dfci.harvard.edu

Project Goals: Our goal in this study is to carry out genome-wide network reconstruction of *C. reinhardtii* metabolism integrated with structural and functional annotation, to provide a framework to integrate these metabolic pathways for investigation of global properties of metabolism.

Algae have garnered significant interest in recent years, both for their potential as a source of biofuel, as well as for their use in the production of nutritional supplements. Among eukaryotic microalgae, *Chlamydomonas reinhardtii* has established its position as an ideal model organism, popular for its relatively fast doubling time and its ability to grow under standardized conditions in the laboratory. Characterization of metabolic functions in *C. reinhardtii* provides a framework for developing engineering strategies towards generation of strains with improved production of commercial targets, as well as studying diverse cellular processes such as photosynthesis and cell motility. Extensive literature on *C. reinhardtii* metabolic function and many mutants with metabolic phenotypes provide a solid foundation toward detailed characterization of individual metabolic pathways in this organism.

Our goal in this study is to carry out genome-wide network reconstruction of *C. reinhardtii* metabolism integrated with structural and functional annotation, to provide a framework to integrate these metabolic pathways for investigation of global properties of metabolism. The availability of complete

genome sequence data (*JGI* v4.0), recently released for *C. reinhardtii*, allowed us to perform in-house annotation of the metabolic genes encoded within this genome (*please see our accompanying abstract by Ghamsari et al. for details*). Briefly, functional annotations enable the identification of the presence of enzymes encoded within the genome of the organism to define the scaffolding of the reconstruction. Comprehensive literature searches are used as the primary form of evidence to establish the structure of all metabolic pathways of interest. Finally, this information is supplemented with more general knowledge of metabolic pathways, as provided in classical biochemistry textbooks and also available in online databases.

We report the first genome-scale reconstruction of *C. reinhardtii*, accounting for all pathways and metabolic functions indicated by the latest release of the genome (JGI v4.0) combined with our in-house generated functional annotation. The reconstruction accounts for 978 genes, associated with 1671 reactions, and includes 1029 unique metabolites. As the most comprehensive metabolic network reconstruction of *C. reinhardtii* to date, ours is the first to account for multiple wavelengths of light involved in metabolism and includes considerable expansion of fatty acid metabolism over previous reconstructions, with pathway details accounting for metabolism of individual R-groups. Further, the metabolic network reconstruction presented here provides a greater level of compartmentalization than existing reconstructions of *C. reinhardtii*, with the inclusion of the lumen as a distinct component of the chloroplast for photosynthetic functionality, and the eyespot used to guide the flagella in phototaxis.

We present simulations under a variety of growth conditions (e.g. acetate/no acetate, light/no light, aerobic/anaerobic), and physiological validation of *in silico* gene knockout against known mutant data for a variety of phenotypes (e.g. increased use of acetate; light; CO_2; nitrogen; and other media components, amino acid requiring, altered color). We further present detailed simulations demonstrating how photon absorption and different wavelengths of light affect downstream metabolic processes, elucidating the benefits of sunlight versus artificial light conditions. Our well-validated and comprehensive genome-scale reconstruction of *C. reinhardtii* metabolism provides a valuable quantitative and predictive resource for metabolic engineering toward improved production of biofuels and other commercial targets.

124

Identifying the Metabolic Potential of *Chlamydomonas reinhardtii* by Large-Scale Annotation of its Encoded Open Reading Frames

Lila Ghamsari[1]* (Lila_Ghamsari@dfci.harvard.edu), Balaji Santhanam,[1] Roger L. Chang,[2] Ani Manichaikul,[3] Erik F.Y. Hom,[4] Marc Vidal,[1] **Jason A. Papin**,[3] and **Kourosh Salehi-Ashtiani**[1]

[1]Center for Cancer Systems Biology and Dept. of Cancer Biology, Dana-Farber Cancer Institute, and Dept. of Genetics, Harvard Medical School, Boston, Mass.; [2]Dept. of Bioengineering, University of California, San Diego; [3]Dept. of Biomedical Engineering, University of Virginia, Charlottesville; and [4]Dept. of Molecular and Cellular Biology, Harvard University, Cambridge, Mass.

http://ccsb.dfci.harvard.edu
http://bme.virginia.edu/csbl

Project Goals: The release of the complete genome sequence of *C. reinhardtii* has made this unicellular alga an ideal model for metabolic engineering; however, the annotation of the relevant genes has not been verified yet and the much-needed metabolic network model is currently unavailable. Using the integrated annotation and metabolic network modeling that we recently established (*Manichaikul et al.*, *Nature Methods 2009*), we are engaged in efforts to: 1) assign enzymatic functions to the annotated proteome of *C. reinhardtii*, 2) experimentally verify or refine the structure of the annotated open reading frames (ORFs), and 3) build a genome-wide metabolic network model for the organism based on the assigned metabolic functions (*please also see our accompanying abstract by R.L. Chang et al.*).

Results: We used the new *JGI* "filtered transcript models" (Chlre4_best_transcripts and Chlre4_best_proteins), and the *Augustus 5* models released through the *JGI* portal (http://genome.jgi-psf.org/Chlre4/Chlre4.home.html) for both functional assignments and structural annotation verifications. Enzymatic functional assignments were made by associating Enzyme Classification (EC) numbers through reciprocal blast searches against Uniprot enzyme database (with over 100,000 protein entries). The best match for each translated ORF was identified (with an e-value threshold of 10^{-3}) and the EC number from the Uniprot best match was transferred on to the ORF. We extended the EC assignments to the respective paralogs of the ORFs by clustering ORFs using BLASTCLUST (sequence identity cut-off of 35% and sequence length cut-off of 70%) within each annotation group (i.e., *Augustus* 5 and *JGI filtered models*). Altogether, we were able to assign 970 EC numbers to 1,448 *JGI* and to 1,877 *Augustus* models. Over 93% of the EC terms were assigned to both *JGI* and *Augustus* models (**Fig. 1A**). We then carried out all possible pairwise alignments between the *JGI* and *Augustus* transcripts that had been assigned the same EC numbers by the above-mentioned

* Presenting author

procedure. In contrast to the high overlap between the two models in terms of EC assignments, less than half of each set were found to be 100% identical in sequence (**Fig. 1B**), indicating that the structural annotation of many of the two sets differ from one another.

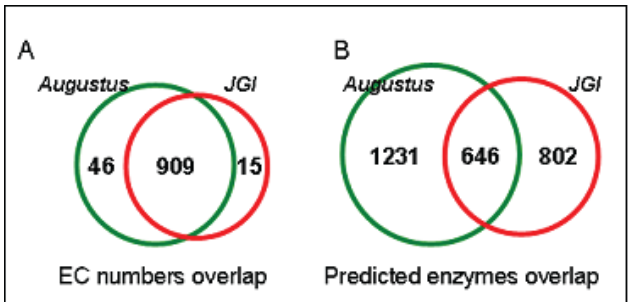


Figure 1. Annotation and overlap of *JGI* and *Augustus* ORF models.

To experimentally verify the structure of both *JGI* and *Augustus* ORF models, we carried out open reading frame (ORF) verification by RT-PCR on all ORFs that we had assigned EC numbers to (as well as a set of positive control ORFs). Following optimization of the RT-PCR procedure for high GC content of the *C. reinhardtii* transcriptome, we were able to observe positive RT-PCR products for approximately 70% of the transcripts. Following cloning, we carried out 454FLX sequencing of the ORFs and aligned the 454FLX reads to the ORF reference sequences (**Fig. 2**). We obtained 95-100% coverage of the ORF length for 940 of the ORF models; 215 ORFs with 50 to 95% coverage, 207 with 10 to 50%, and less than 10% coverage for the remainder. These results indicate that at least half of the transcript models are accurately annotated. Cloning and parallel sequencing of the remaining *JGI* metabolic ORFs are in progress.

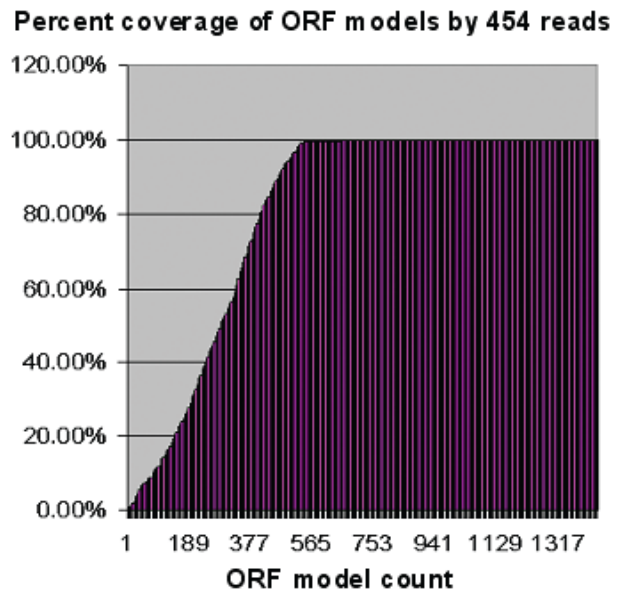


Figure 2. Sequence verification of metabolic ORF models. 1,877 ORF models were tested by RT-PCR, recombinational cloning and 454 FLX sequencing.

The verified metabolic ORF clone resource that we have generated will be made available without restrictions to the research community.

125

Pathway of Fermentative Hydrogen Production by Sulfate-Reducing Bacteria

Kimberly L. Keller[1,2]* (kellerkl@missouri.edu), Steve D. Brown,[3,2] Iris Porat,[3] and **Judy D. Wall**[1,2] (WallJ@missouri.edu)

[1]University of Missouri, Columbia; [2]VIMSS (Virtual Institute of Microbial Stress and Survival) http://vimss.lbl.gov/, Berkeley, Calif.; and [3]Oak Ridge National Laboratory, Oak Ridge, Tenn.

Project Goals: The production of hydrogen by fermentative pathways of the anaerobic sulfate-educing bacteria (SRB) of the genus *Desulfovibrio* is the focus of the project. The limitations to hydrogen production identified in these model organisms may be informative for those microbes chosen for industrial hydrogen generation. We propose to determine the contribution of substrate-level phosphorylation to respiratory growth on sulfate and the contribution of respiration to fermentation of pyruvate. The enzymes for pyruvate oxidation will be established in two strains of *Desulfovibrio* to identify the reduced product available for hydrogen generation. Electron sinks potentially competing with protons will be eliminated individually and together to determine the plasticity of electron flow to hydrogen.

We are exploring the production of hydrogen by fermentative pathways of the anaerobic sulfate-reducing bacteria (SRB) of the genus *Desulfovibrio*. The SRB have not been considered model organisms for hydrogen production, yet they have a multifaceted hydrogen metabolism. Strains of *Desulfovibrio* can ferment organic acids in the absence of terminal electron acceptors and produce rather large amounts of hydrogen. We believe a study of the limitations to hydrogen production in these model organisms (in particular, *Desulfovibrio* G20 and *Desulfovibrio vulgaris* Hildenborough) may be informative to decipher the flow of electrons in those organisms chosen for industrial application for hydrogen production. We have proposed to:

- Determine the contribution of fermentation to the respiratory energy budget as well as the contribution of respiration processes to the fermentation of pyruvate. Ultimately, we seek to separate these pathways and deduce the dependence of the bacterium on these combined processes.
- Identify the enzyme (s) responsible for the oxidation of pyruvate to elucidate the reduced substrate initially available for hydrogen generation during fermentation.
- Eliminate alternative electron sinks that are likely to compete for protons, reducing the overall yield of hydrogen.

An analysis of metabolic end products of *Desulfovibrio* G20 cultures grown on pyruvate as the sole electron donor and acceptor has indicated that respiration and fermentation occur simultaneously. Conversely, cells respiring sulfate with lactate also apparently gain energy by substrate level phosphorylation. We are attempting to delete the gene encoding acetate kinase to confirm that this enzyme is essential for substrate-level phosphorylation during pyruvate fermentation and determine the effect of this deletion on the efficiency of pyruvate and lactate respiration.

To determine what pathways are present for energy generation during respiration or fermentation, we have initiated an analysis of the proteomics in each condition. Cultures from wild-type *Desulfovibrio* G20 grown by pyruvate fermentation and by sulfate respiration with lactate are being used to identify enzymes which appear differentially expressed between growth modes, as well as those are found in high abundance in both growth modes. These results provide candidates for further exploration

The construction of a Tn5 transposon mutant library in *Desulfovibrio* G20 is allowing us to establish the importance of the candidate enzymes identified by the proteomics. The mutants are grown on different substrates and their growth and metabolites are compared to those of wild-type *Desulfovibrio* G20. We have already begun an analysis of hydrogen production and metabolite changes of a number of G20 Tn*5* mutants, including those lacking a fumarate reductase, molybdopterin oxidoreductase, formate dehydrogenases, and malic enzyme. No growth differences where found between wild-type and mutants lacking formate dehydrogenases or malic enzyme mutants, probably because there are apparently multiple isozymes encoding these enzymes that could compensate for the single enzyme loss. As expected, there was no succinate measured for the mutant with the Tn5 inserted in the fumarate reductase. In addition, the fumarate reductase mutant was unable to grow when fumarate only was provided as the sole electron donor and acceptor. This lack of fumarate dismutation has also been observed with a *Desulfovibrio* G20 plasmid insertion mutation in the cycA gene which encodes the type-1 tetraheme cytochrome c3.

Advances continue to be made in the genetic manipulation *D. vulgaris*, including a markerless deletion system allowing the sequential deletion of multiple genes. The system, which uses the *upp*-encoded uracil phosphoribosyltransferase as an element for counterselection, is being refined to increase efficiency of mutant selection. Preliminary experiments show the wild-type G20 strain is quite sensitive to 5-fluorouracil and attempts to make *Desulfovibrio* G20 more amenable to gene specific mutations continues. We are currently in the process of generating a barcoded transposon library in *Desulfovibrio vulgaris* Hildenborough.

126

Transcriptomic Analyses of the Sulfate-Reducing Bacterium *Desulfovibrio* G20 and a Type-1 Tetraheme Cytochrome c3 Mutant during Their Transitions into Stationary Phase

Iris Porat[1]* (porati@ornl.gov), Kimberly L. Keller,[2,3] Barbara J. Giles,[2] Timothy J. Tschaplinski,[1] Steve D. Brown,[1] and **Judy D. Wall**[2,3] (WallJ@missouri.edu)

[1]Oak Ridge National Laboratory, Oak Ridge, Tenn.; [2]University of Missouri, Columbia; and [3]VIMSS (Virtual Institute of Microbial Stress and Survival) http://vimss.lbl.gov/, Berkeley, Calif.

Project Goals: The production of hydrogen by fermentative pathways of the anaerobic sulfate-reducing bacteria (SRB) of the genus *Desulfovibrio*. is the focus of the project. The limitations to hydrogen production identified in these model organisms may be informative for those microbes chosen for industrial hydrogen generation. We propose to determine the contribution of substrate-level phosphorylation to respiratory growth on sulfate and the contribution of respiration to fermentation of pyruvate. The enzymes for pyruvate oxidation will be established in two strains of *Desulfovibrio* to identify the reduced product available for hydrogen generation. Electron sinks potentially competing with protons will be eliminated individually and together to determine the plasticity of electron flow to hydrogen.

Fermentative hydrogen generation provides a mechanism for anaerobic microbes to release electrons in a neutral fashion during oxidation of vast quantities of organic matter. We are examining the capacity of a soil anaerobe, the sulfate-reducing bacterium *Desulfovibrio* G20, to generate hydrogen from organic acids. These apparently simple pathways have yet to be clearly established. Previously, an insertion muta-

tion in *cycA,* encoding the type-1 tetraheme cytochrome c_3, was constructed in *Desulfovibrio* G20 (Rapp-Giles et al. 2000). The growth rate of the CycA$^-$ mutant was similar to that of the wild type during lactate-supported sulfate respiration or pyruvate fermentation. However, CycA$^-$ appeared to be unable to respire sulfate with pyruvate as the electron donor and it was unable to dismutate fumarate. Interestingly, CycA$^-$ growing with lactate/sulfate generated more hydrogen gas than G20, suggesting that electron flow was rerouted.

To determine any changes in gene expression that might account for these differences, both *Desulfovibrio* G20 and the CycA mutant were grown in pH-controlled fermentors in define medium by lactate-supported sulfate respiration or by pyruvate fermentation. RNA was prepared and microarray analysis performed to determine differential gene expression between exponential and stationary phase G20 cells grown by respiration or by fermentation and then determine any differences in gene expression in the CycA mutant grown under comparable conditions. The transcriptomic results demonstrated gene expression profile changes of G20 in response to carbon and energy depletion upon entering stationary phase as well as those occurring in response to pyruvate fermentation versus lactate respiration with sulfate. Hierarchical cluster analysis of transcriptomic profiles for G20 and the CycA mutant revealed 12 distinct clusters. Also, the arrays were clustered by growth phases. As would be predicted, in both strains genes involved in translation and transcription were expressed at high levels during the log phase. In addition, some genes for biosynthesis of amino acids, co-enzymes, lipid and carbohydrates were up-regulated during logarithmic growth. During the stationary phase, large numbers of genes encoding proteins for energy generation were up-regulated, including hydrogenases, dehydrogenases and ATP synthase. Also, once the electron donors were exhausted, genes involved in the flagella biosynthesis and in stress responses were increased in expression.

Interestingly, two well defined clusters with 14 ORFs each showed significant transcript differences between the two strains and the two media. In the CycA mutant, one of the clusters that included a large operon encoding the enzymes for conversion of pyruvate to succinate was down regulated. The second cluster showed decreased expression in pyruvate medium compared to lactate sulfate medium in both the log and stationary phases for G20 as well as for the CycA mutant. That cluster encodes two operons, one containing the type II c_3 transmembrane complex genes and a second with other energy related protein genes.

Metabolic profiles were obtained for G20 and CycA cultures in respiring and fermenting cultures. Comparing the metabolites from the CycA mutant to those from G20 in lactate/sulfate medium, higher concentrations of glycerol-1-phosphate, stearate, citramalate, aspartate, glycine and nucleotides were detected regardless of growth phase. Lower concentrations of lactate and succinate were found in the mutant. In contrast, in pyruvate fermenting cultures, the concentrations of glutamate, lactate, aspartate, N-acetylaspartate, glycine, and 5-oxo-proline were higher in the CycA mutant than in G20 in log phase and opposite in stationary phase. The reverse was found for concentrations of succinate and trehalose in the mutant, lower than those from G20 in log phase and higher in stationary. These intracellular changes may also reflect overall redirection of substrates and electron flow in the two strains.

Reference

1. Rapp-Giles, B. J., L. Casalot, R.S. English, J.A. Ringbauer, Jr., A. Dolla, and J.D. Wall. 2000. Cytochrome c_3 mutants of *Desulfovibrio desulfuricans*. Appl. Environ. Microbiol. 66:671-677.

Student Presentation

127

Genome Resequencing of Thermoacidophilic Archaeal Carbon Flux Mutants

Y. Maezato* and **P. Blum** (pblum1@unl.edu)

Beadle Center for Genetics, University of Nebraska, Lincoln

Project Goals: The project (DE-FG36-08G088055) supports research on the ability of extremely thermoacidophilic microbes and their proteins to accelerate lignocellulose processing from switchgrass and to maximize sugar release. Delineation of carbon metabolic pathways in *Sulfolobus solfataricus* (*Sso*) is an essential step towards engineering traits associated with lignocellulosic bioprocessing. In *Sso*, carbon catabolism varies between two phenotypic states. In car mutants, growth on hexoses is blocked while pentose catabolism remains unaffected. In light of this characteristic, identification of car was undertaken.

The project (DE-FG36-08G088055) supports research on the ability of extremely thermoacidophilic microbes and their proteins to accelerate lignocellulose processing from switchgrass and to maximize sugar release. Delineation of carbon metabolic pathways in *Sulfolobus solfataricus* (*Sso*) is an essential step towards engineering traits associated with lignocellulosic bioprocessing. In *Sso*, carbon catabolism varies between two phenotypic states. In *car* mutants, growth on hexoses is blocked while pentose catabolism remains unaffected. In light of this characteristic, identification of *car* was undertaken. To genetically map *car* and its regulated targets, a genomic large insert library (BACs) was sorted by selection for clones that restored hexose metabolism in the *car* mutant. Shotgun subclone libraries led to the identification of a glycolyltic gene, *kdgK*, and a chromatin modification gene, *hdac-1*, as components of the *car* mutant glycolytic defect. Genome resequencing was conducted to further characterize the genetic basis for *car*. Nearly 50 differences were identified distinguishing genomes of the *Sso* wild type and *car* mutant derivative. These included the CRISPR associated RAMP module Cas gene cmr2. Additional three-way whole genome comparisons further narrowed strain differences. Detailed genotyping combined with genomic

reconstruction is underway to further assess these whole genome sequencing results.

128

Metabolic Fluxes: Quantifying Competition Between Nitrate and Proton Reduction during Fermentative NAD (P)H Formation in Real-Time

Gennady Ananyev,* Nicholas J. Skizim, and **G. Charles Dismukes** (dismukes@rci.rutgers.edu)

Waksman Institute of Microbiology and Dept. of Chemistry and Chemical Biology, Rutgers University, Piscataway, N.J.

Project Goals (Abstracts 128-129): The goal of this project is the design and utilization of tools which will allow a quantitative understanding of the mechanism and kinetics of hydrogen production in cyanobacteria.

The ability to continuously monitor metabolite concentrations *in vivo* as they undergo transient changes in response to environmental perturbations is among the most sought goals of metabolomics. Here we provide direct measures of the quantitative flux of intracellular reductant and its temporal dynamics between competing pathways. We have developed a fluorometric method for monitoring the concentration of intracellular reduced pyridine nucleotides, NAD (P)H, and combined it with a real-time electrochemical assay of dissolved H_2 concentration for simultaneous measurements on whole cells. Here we describe an application to quantitative kinetics in the cyanobacterium *Arthrospira maxima*, the most prolific fermentative H_2 producer in this group. *A maxima* produces H_2 via a bidirectional NiFe-hydrogenase, the sole H_2-metabolizing enzyme in this organism. We demonstrate that the two temporal phases of H_2 production induced following the onset of anaerobiosis arise from distinct metabolic processes responsible for (a) anabolic production of NADPH (Phase-1) and (b) catabolic production of NADH (Phase-2). Phase-1 starts within minutes of anaerobiosis and may extend up to 2-3 h. Phase-1 H_2 is shown to correlate with the residual pool of photosynthetically produced NADPH; it decreases with increasing aerobic dark time prior to anaerobiosis. The second phase starts within 3-20 h and follows the rise of intracellular NAD (P)H with a short lag time (Δt = 24 min) indicative of the time to achieve a redox poise ([NAD (P)H//[NAD $(P)^+$]>100) sufficient for H_2 formation. The major Phase-2 H_2 is produced by autofermentation of carbohydrate reserves. The yield of Phase-2 H_2 is shown to be inversely related to the concentration of NO_3^- in the medium. A positive linear relationship is observed between the NO_3^- concentration and the delay time for onset of Phase-2 H_2 production over more than two decades change in concentration. Experiments carried out at various extracellular concentrations of inhibitors of nitrate reductase (N_3^-, CN^-) provide solid evidence for a direct competition for consumption of cellular reductant by hydrogenase and nitrate reductase (*nar*). This evidence provides a firm basis for proposed metabolic engineering of pathways.

Supported by DOE-GTL DE-FG02-07ER64488 and AFOSR-MURI FA9550-05-1-0365.

129

Real-Time Co-Detection of Dissolved H_2 and Intracellular NAD (P)H Concentrations in Microbes

Gennady Ananyev, Nicholas J. Skizim,* and **G. Charles Dismukes** (dismukes@rci.rutgers.edu)

Waksman Institute of Microbiology and Dept. of Chemistry and Chemical Biology, Rutgers University, Piscataway, N.J.

Project Goals: See goals for abstract 128.

In an effort to probe the kinetics and quantify the concentration of hydrogen produced by living microorganisms, here we describe a home-built combination cell for simultaneous real-time detection of dissolved H_2 and reduced pyridine nucleotide levels, [NAD (P)H]. Electrochemical H_2 detection (LOD: 2 nC = 1 nM H_2) from live cells with sample volumes as small as 5 µl is accomplished with a custom made Clark-type cell (or reverse fuel cell) comprised of a membrane covered Pt/Ir electrode poised at a bias of +220 mV . Total reduced pyridine nucleotide = [NADPH] + [NADH] is assayed by selectively exciting NAD^+ and $NADP^+$ through an ultraviolet light emitting diode (365 nm) operating in pulse mode, and subsequently measuring fluorescence emission at 470± 20 nm with a photodiode. Calibration curves with standards reveal linear responses for both H_2 and NAD (P)H detection using inactivated microbes as background, from which quantitative concentrations could be assessed. Using these tools we have been able to observe for the first time multiple temporal phases of H_2 production in anaerobically poised cyanobacteria which possess a [NiFe]-hydrogenase, as well as those containing a nitrogenase. We show that two main phases of hydrogen production correlate with the availability of residual NADPH produced via photosynthesis prior to the onset of anaerobiosis (phase 1), and NADH produced by anaerobic autofermentation of glycogen reserves (phase 2) in strains containing a [NiFe]-hydrogenase as their only hydrogen metabolizing enzyme. Application of these tools to N_2 fixing cyanobacteria has similarly revealed at least three kinetic phases of dark hydrogen production coupled to fermentative metabolism. Experiments are planned on two diazotrophic *Synechococcus* thermophiles which lack all types of hydrogenase. Together these data shall provide a basis for discriminating between and quantifying the H_2 evolution pathways involving hydrogenase and nitrogenase.

Supported by DOE-GTL DE-FG02-07ER64488 and AFOSR-MURI FA9550-05-1-0365

130

Systems Level Approaches to Understanding and Manipulating Heterocyst Differentiation in *Nostoc punctiforme*: Sites of Hydrogenase and Nitrogenase Synthesis and Activity

Harry Christman,[1] Elsie Campbell,[1] Douglas Rissser,[1] Brett Phinney,[2] Wan-Ling Chiu,[3] and **John C. Meeks**[1]* (jcmeeks@ucdavis.edu)

[1]Dept. of Microbiology and [2]Genome Center, University of California, Davis; and [3]Dept. of Biology, Virginia Commonwealth University, Richmond

Project Goals: Heterocyst-forming cyanobacteria, such as *Nostoc punctiforme*, are applicable for cost-effective photo-biohydrogen production, providing the frequency of heterocysts in the filaments can be increased by genetic manipulation and the metabolic end product, H_2, is uncoupled from growth.

We hypothesize that if the frequency of heterocysts, sites of nitrogen fixation and hydrogen evolution, can be increased approximately 3-fold above the less than 10% normally found in filaments, then heterocyst-forming cyanobacteria would be applicable for cost effective photo-biohydrogen production. In *Nostoc punctiforme*, the heterocyst frequency of 8% in the free-living state is increased to 30-35% when in symbiotic association with terrestrial plants, such as the hornwort *Anthoceros* spp. and the angiosperm *Gunnera* spp. We are applying transctriptomic and proteomic analyses with wild-type and mutant strains to identify the regulatory circuits of free-living heterocyst differentiation and how those circuits have been co-opted during symbiotic growth.

Transcriptomics. HetR and HetF are two positive regulatory elements in the differentiation of heterocysts. Mutations in either structural gene in *N. punctiforme* lead to the inability to differentiate heterocysts in free-living cultures or to support N_2-dependent growth of the hornwort plant partner, although the mutants do infect the plant. Conversely, over expression of the genes from a multicopy plasmid in trans, results in the differentiation of multiple heterocysts that are localized in distinct clusters with a nonrandom spacing pattern in the filaments. The heterocyst frequency can reach 40% of the cells, but the heterocysts do not fix nitrogen in support of vegetative cell growth. A question is whether the gene products operate at the same pathway in the regulatory cascade of heterocyst differentiation. Time course (0.5 to 24 h) DNA microarray experiments, using a Nimblegen platform, with the data analyzed in the R statistical environment and subsequently clustered using the Genesis program, yielded two interesting results. A total of 124 genes were up-regulated in both mutants in responsive to combined nitrogen limitation. Only one known up-stream gene for heterocyst differentiation gene (*nrrA*) was up-regulated. The up-regulated genes include those for assimilation of the alternative inorganic nitrogen sources nitrate and urea. Only one amino acid transport gene was up-regulated, implying *N. punctiforme* does not search for organic nitrogen sources during nitrogen starvation. An additional common cluster of 330 genes were down-regulated in the two mutants and these included many encoding proteins of central metabolism, protein synthesis and photosynthetic energy metabolism. Approximately 53% of the up- and down-regulated transcripts encode proteins of unassigned function. The second result of interest is that 130 and 775 genes were uniquely up-regulated, and 29 and 1063 were uniquely down-regulated in the *hetF* and *hetR* mutants, respectively. These results imply that HetF and HetR have physiological roles in addition to heterocyst differentiation, which is consistent with the presence of the genes in filamentous cyanobacteria that do not differentiate heterocysts. We are now examining the transcriptional profiles of wild type strains over expressing HetR and HetF in order to identify genetic targets and any differences in the targets in the two constructs for comparisons to the loss of function mutant data. We have also initiated analysis of a unique pattern mutant in which a high heterocyst frequency (~ 35%) is manifest as multiple singular heterocysts with a 3-4 vegetative cell spacing between heterocysts. This is the pattern we observe in the symbiotic growth state. However, this mutant also does not grow with N_2 as the sole nitrogen source.

Proteomics. We have completed a single Mudpit-based MS/MS run of the total proteome of a N_2-grown culture of *N. punctiforme*. The cell extract was processed into three fractions; an initial 14,000 x g pellet of membrane proteins, a 150,000 x g supernatant of soluble proteins and a 150,000 x g pellet of primarily carboxysomes and phycobilisomes. The fractions were sub-fractionated by PAGE, sliced sections digested with trypsin and the eluted products further fractionated by 2 dimensional LC followed by electrospray injection. This 600 member proteome is comparable to one we previously defined of the supernatant proteins of an ammonium-grown culture, except for an enrichment in membrane proteins, plus proteins associated with heterocysts and nitrogenase function. Notable is the constitutive presence of DNA photolyases and enzymes for metabolism of reactive oxygen species, stress factors to which cyanobacteria are constantly exposed. The value of this proteome will be enhanced by comparison to a N_2 plus fructose grown mixotrophic culture now in progress. In the symbiotic growth state *N. punctiforme* grows as a photomixotroph in *Anthoceros* and most likely as a heterotroph in *Gunnera*. Thus, these proteomes are essential for comparisons to symbiotic growth with the accompanying high heterocyst frequency and high rates of nitrogenase activity.

Student Presentation

131

Contributors to Light- and Feedstock-Powered Hydrogen Production in *Rhodobacter sphaeroides*

Melanie A. Spero[1]* (spero@wisc.edu), Ugur Sohmen[2]* (sohmen@wisc.edu), Wayne S. Kontur,[1] L. Safak Yilmaz,[2] **Daniel R. Noguera**[2] (noguera@engr.wisc.edu), and **Timothy J. Donohue**[1] (tdonohue@bact.wisc.edu)

[1]Dept. of Bacteriology; and [2]Dept. of Civil and Environmental Engineering, University of Wisconsin, Madison

Project Goals: Our research seeks to generate a quantitative understanding of light- and feedstock-powered hydrogen (H_2) production in the photosynthetic bacterium *Rhodobacter sphaeroides*. We will determine the source (s) of H_2 production in *R. sphaeroides* and determine the effect on fuel production of eliminating or minimizing processes predicted to siphon reducing power from H_2 production. This will be accomplished by deleting genes in pathways predicted to impact H_2 production and assaying for the amount of reductant that is partitioned to H_2, biomass, as well as other electron sinks. Mutations tested to date include those in genes encoding enzymes predicted to generate H_2 (nitrogenase), detract from H_2 production (hydrogenase), and in pathways predicted to compete for reducing power that could be utilized for H_2 production (synthesis of polyhydroxybutyrate, carbon dioxide fixation). This work also compares H_2 production in log-phase batch cultures to late stationary phase cultures, where the amount of reducing power shuttled into biomass production is minimized. Each mutant will be assayed for H_2 production under active or non-growing conditions on a variety of carbon sources to determine the impact of these cellular processes on H_2 production.

These studies are performed in the α-proteobacterium *R. sphaeroides*, the most-studied photosynthetic bacterium, since it is known to produce relatively large amounts of H_2 under photoheterotrophic growth conditions. A working model of the processes predicted to impact H_2 production was developed based on prior knowledge of metabolic pathways, the genome sequence of this bacterium, and global gene expression data (see posters by Yilmaz et. al., and Kontur et al.).

H_2 production in *R. sphaeroides* is proposed to be mainly or completely associated with the nitrogenase enzyme. To determine the contribution of nitrogenase to H_2 production, a mutant strain of *R. sphaeroides* was generated that lacks the structural genes of nitrogenase (*nifHDK*). The growth properties of the ΔNif mutant support the hypothesis that nitrogenase-mediated H_2 production is an important electron sink with some, but not all, carbon feedstocks. Experiments are ongoing to use the ΔNif mutant to determine if other *R. sphaeroides* enzymes can contribute to H_2 production under defined conditions.

One enzyme predicted to detract from H_2 production is uptake hydrogenase (Hup), which can oxidize H_2 into protons and electrons. A mutant strain of *R. sphaeroides* lacking the Hup structural genes (*hupSL*) is being used to determine the impact of Hup activity on H_2 production, especially since global gene expression data predict a wide range of Hup activities present when cells grow in media containing different carbon sources as feedstocks.

We are also testing the effects of other electron sinks on H_2 production, by blocking pathways known or predicted to compete for reducing power. For example, mutant strains with gene deletions in polyhydroxybutyrate synthesis or carbon dioxide sequestration (via the Calvin cycle) are being analyzed to determine effects on H_2 production. Each of these mutants is being assayed for H_2 production on various carbon sources in order to evaluate how altering the flow of reducing power into other electron sinks can alter the distribution of reducing power into various products.

Because biomass is as a significant electron sink in exponentially growing cells, the distribution of electrons into different products is also being analyzed in stationary phase cultures, when electron flow towards biomass production is predicted to be minimized. Preliminary results indicate that partitioning of reducing power to H_2 production varies between growing and stationary phase cultures on some carbon feedstocks. We will report on experiments with wild type and mutant strains that seek to monitor electron partitioning during stationary phase to better understand how feedstock supply impacts the flow of electrons to H_2 and other electron sinks.

132

SurR Regulates Hydrogen Production in *Pyrococcus furiosus* by a Sulfur-Dependent Redox Switch

Gina L. Lipscomb,[1]* Hua Yang,[1,4] Annette Keese,[2] Gerrit J. Schut,[1] B.C. Wang,[1] Michael Thomm,[2] Janet Westpheling,[3] Michael W.W. Adams,[1] and **Robert A. Scott**[1] (rscott@uga.edu)

[1]Dept. of Biochemistry and Molecular Biology, University of Georgia, Athens; [2]Dept. of Microbiology, University of Regensburg, Germany; [3]Dept. of Genetics, University of Georgia, Athens; and [4]Centers for Disease Control and Prevention, Atlanta, Georgia

Project Goals: The goal of SAPHyRe (Systems Approach to Probing Hydrogen Regulation) is to develop a detailed systems-level description of the regulatory and metabolic networks controlling hydrogen production in the hyperthermophilic archaeon *Pyrococcus furiosus* (*Pf*). *Pf* will be used as the model organism to investigate its response to various environmental conditions relevant to *all* hydrogen-producing microorganisms, such as carbon and nitrogen sources, metal availability, and oxidative and reductive stresses. The outcome of this project will

serve two purposes: 1) it will bring us one step closer to utilizing *P. furiosus* in development of H_2 as an alternative energy source and 2) it will serve as a model methodology for investigating the regulatory pathways of hydrogen production in other organisms.

We present structural and biochemical evidence for a redox switch in the archaeal transcriptional regulator SurR of *Pyrococcus furiosus*, a hyperthermophilic anaerobe. *P. furiosus* produces H_2 during fermentation, but undergoes a metabolic shift to produce H_2S when elemental sulfur (S^0) becomes available. Changes in gene expression occur within minutes of S^0 addition, and the majority of these S^0-responsive genes are regulatory targets of SurR, a key regulator involved in primary S^0 response. SurR was shown *in vitro* to have dual functionality, activating transcription of some of these genes, notably the hydrogenase operons, and repressing others, including a gene encoding sulfur reductase. This work demonstrates that the activity of SurR is modulated by cysteine residues in a CxxC motif that constitute a disulfide switch. Oxidation of the switch with S^0 inhibits sequence-specific DNA binding by SurR, leading to *de*activation of genes related to H_2 production and *de*repression of genes involved in S^0 metabolism.

133

Pathways and Regulatory Network of Hydrogen Production from Cellulose by *Clostridium thermocellum*

Jonathan Millen[1]* (jonathan_millen@hotmail.com), Ji Hye Jo,[2] Shiv Thammannagowda,[2] Grant Pennington,[2] Matthew Heckman,[1] Pin-Ching Maness,[2] Mary Lipton,[3] **J.H. David Wu**[1] (davidwu@che.rochester.edu)

[1]University of Rochester, N.Y.; [2]National Renewable Energy Laboratory, Golden, Colo.; and [3]Pacific Northwest National Laboratory, Richland, Wash.

Project Goals: The overall objective of this research is to understand H_2 metabolic pathways in *Clostridium thermocellum* and the underlying regulatory network at the molecular and systems levels. We hypothesize that: multiple hydrogenases work concertedly to contribute to hydrogen evolution; transcription factors control hydrogenase expression at the transcription level; and hydrogen metabolism is coupled to cellulolysis and fermentative pathways in this bacterium. We are determining the hydrogenase expression and metabolic network nodes on cells subjected to different culture conditions and metabolic pathway inhibitors, to probe differential expression of the various hydrogenases and their interrelationship with other cellular metabolic pathways. We are also mapping connections in the transcription factor network controlling linked metabolic pathways. Finally, we will purify FeFe-hydrogenases from its native producer and heterologously expressed *E. coli* to determine their subunit compositions, endogenous redox partners, and the direction of reaction (hydrogen production vs. uptake) to shed light on their roles in hydrogen metabolism.

Clostridium thermocellum, a thermophilic, ethanogenic, and cellulolytic anaerobe, produces a complex extracellular cellulolytic organelle called the cellulosome. The cellulosome contains various depolymerizing enzymes that are arrayed on a protein scaffold and effectively degrades complex cellulosic substrates. During cellulose fermentation, the bacterium evolves hydrogen at a high rate. Analysis of its genome sequence reveals the existence of at least three putative hydrogenase genes (CtHydA1, CtHydA2 and CtHydA3) central to hydrogen metabolism. Furthermore, at least 20 genes are potentially related to hydrogen metabolism. The bacterium is thus remarkably versatile in employing various enzymes, some of which are potentially novel, for hydrogen metabolism, indicating the significance of this biological process. Yet little is known concerning the pathway for hydrogen production and the underlying regulatory mechanism/network that control these hydrogenase and the related genes as well as cellulolytic process and other metabolic pathways in the organism.

Due to the novelty of CtHydA3, it was selected for the initial study. Genes encoding CtHydA3, a ferredoxin-like protein (Ct_3004), and three FeFe-hydrogenase maturation proteins (CtHydE, CtHydF, and CtHydG) have been cloned into three plasmids and co-transformed into *E. coli* strain Rosetta (DE3) and BL21 (DE3) for heterologous expression. In addition, a 6X His-tag sequence was fused to either the C- or N-termini of CtHydA3. Protein immunoblots confirmed the expression of the C-terminus His-tagged CtHydA3 (73 kDa band) in *E. coli* Rosetta (DE3), but not in *E. coli* BL21 (DE3), likely due to differences in codon usage between *C. thermocellum* and *E. coli*. The expression is further corroborated by a five-fold increase in *in vitro* hydrogenase activity in the soluble cell extract of the recombinant Rosetta strain, mediated by reduced methyl viologen. However, no difference in *in vivo* hydrogen production was detected in the recombinant Rosetta strain, suggesting an inability of the recombinant hydrogenase to contribute to the host's hydrogen metabolism. The C-terminal His-tagged protein failed to bind to a TALON metal affinity column to facilitate its purification. Work is underway to express the N-terminal tagged His-CtHydA3 and explore its affinity purification and characterization.

To identify transcription factors controlling metabolic pathways, we developed an affinity purification method by immobilizing promoter DNA sequences to a solid support. DNA-binding proteins from the *C. thermocellum* cell lysate, obtained by growing on cellobiose or crystalline cellulose and eluted from the affinity columns, were identified by the MALDI-TOF or LC-MS-MS techniques. Several transcription factor candidates were identified. In a reversed approach, we expressed in *E. coli* various putative DNA-binding proteins found in the genome of *C. thermocellum*. Their target DNA-binding sites will be screened by using a DNA microarray we designed and confirmed by EMSA (electrophoretic mobility shift assay).

The studies will provide important insights into the pathway and regulatory mechanism/network controlling hydrogen metabolism and cellulolysis as well as other pertinent metabolic pathways in this very intriguing cellulolytic and thermophilic bacterium, which catalyzes the rate-limiting cellulose-degradation reaction in a single-step process of biomass conversion (or CBP, Consolidated Bioprocessing). Detailed understanding of the pathway and regulatory mechanism/network will ultimately provide rationales for engineering, alternating, or deregulating the organism for biomass conversion to liquid and hydrogen fuels.

134

Filling Knowledge Gaps in Biological Networks: Integrated Global Approaches to Understand Biofuel Metabolism in *Chlamydomonas reinhardtii*

Venkataramanan Subramanian,[1,2] Alexandra Dubini,[1] Claudia Catalanotti,[3] Wenqiang Yang,[3] Christopher H. Chang,[1] Ambarish Nag,[1] David M. Alber,[1] Peter A. Graf,[1] Glenn Murray,[2] Kwiseon Kim,[1] Arthur R. Grossman,[3] Matthew C. Posewitz,[2] and **Michael Seibert**[1]* (mike.seibert@nrel.gov)

[1]National Renewable Energy Laboratory, Golden, Colo.; [2]Colorado School of Mines, Golden, Colo.; and [3]Carnegie Institution of Washington, Stanford, Calif.

Project Goals: Development of photobiological, biofuels-production processes, a key component of DOE's renewable energy mission, would be accelerated by improved understanding of cellular metabolism and its regulation. Currently, a lack of comprehensive experimental data hinders the development of reliable metabolic models that have robust predictive capabilities. Therefore, we are employing high-performance computing to (1) estimate parameters that delimit the space of stable solutions for experimentally constrained metabolic models, (2) explore network capabilities in silico, and (3) integrate experimental, systems biology data to verify and refine metabolic models. Biochemical reactions are modeled with the fundamental independent variables being enzyme concentrations and Michaelis-Menten parameters. Through iterative model building and an understanding of cellular metabolism obtained from high throughput "omics" data, we are constructing metabolic models that link individual enzyme reactions and the activities of specific metabolic pathways with the production of biofuels in the green alga, *Chlamydomonas reinhardtii*. High-performance simulation and optimization will predict metabolic outputs based on the kinetic parameters governing individual reactions, identify metabolic limitations, and predict specific manipulations that are likely to improve biofuel outputs. The work is envisioned as an important contribution toward strengthening our knowledge of energy-related biosystems.

The goal of this project, jointly funded by the DOE Computational Biology and SciDAC Programs, is to develop a means to globally map, *in silico*, all biological pathways in *Chlamydomonas reinhardtii* (*Chlamydomonas* throughout) that can impact the production of H_2 and other biofuels. *Chlamydomonas* is the first alga for which a fully sequenced genome was available (thanks to the JGI), and this organism has recently emerged as the prototype for investigating the regulation of basic metabolism, particularly fermentative processes. *Chlamydomonas* has a complex anaerobic metabolic network that can be induced in the dark and produces H_2 along with other fermentation products such as formate, acetate, ethanol, and CO_2. Previous studies have focused on determining genes (by microarray and RT-PCR analysis) that were differentially regulated as the result of shifting cultures of the CC-425 parental strain from aerobic growth to dark anaerobiosis [1]. Anoxia led to differential expression of genes involved in fermentation metabolism, and specifically caused accumulation of transcripts encoding pyruvate formate lyase (*PFL1*) and pyruvate:ferredoxin oxidoreductase (*PFR1*). Moreover, *Chlamydomonas* synthesized the metabolites formate, acetate, and ethanol in the ratio 1:1:0.5. Recent experiments have shown that in a strain lacking proteins involved in assembling an active hydrogenase enzyme (*hydEF* mutant), the metabolite ratio shifted to 2:1:1, and succinate accumulated instead of H_2. Interestingly, levels of transcripts encoding proteins involved in fermentation also changed in the mutant relative to parental cells following the imposition of anoxic conditions; the *PFL1* mRNA accumulated to a higher level and the *PFR1* mRNA to a lower level in the mutant than in parental cells [2]. These results have allowed us to generate a physiological model that highlights the flexibility of fermentation metabolism and H_2 production in *Chlamydomonas*. Additional mutants (e.g., *ΔPFL*, *ΔFMR*, *ΔHYDA2*, *ΔHYDA2:ΔPFL*, *ΔHYDA2:ΔMME4*) have been identified this year using a HTP screening procedure, and the analyses of these mutants are enabling us to further refine our knowledge of fermentation metabolism and its regulation in *Chlamydomonas*. Furthermore, transcript, protein and metabolite analyses of various single and double mutants (e.g., *ΔHYDA1, ΔADH1, ΔPDC3, ΔHYDA1: ΔHYDA2, ΔPFL1: ΔADH1)* are already underway.

To integrate the results described above with various experimental results discussed below, we have developed metabolic models which, when simulated using our recently developed software toolkit, enable us to make informed decisions about the best ways to link experimentally measurable parameters to the biochemical reactions and their regulation in *Chlamydomonas*. The multi-compartment model includes carbon metabolism (glycolysis, tricarboxylic acid cycle, starch metabolism, Calvin-Benson Cycle), oxidative phosphorylation, and fermentative metabolism. Steady-state kinetic relationships are expressed as thermodynamically consistent ordinary differential equations. Merging of component pathway models, incorporation of ionization reactions, transformation to C++, compilation of the data, and linking the information to high-performance executable programs is largely automated using our *High-Performance Systems Biology Toolkit* (HiPer SBTK) [3]. The programs being used

are for parameter sampling, data fitting, and local or global optimization within a defined space of kinetic parameters and/or enzyme concentrations. Job configuration is possible through an auxiliary graphical interface or by direct editing of simple text. The code has been modularized, permitting facile incorporation of new techniques for high-dimensional sampling, optimization, and model integration. Key future developments will include investigation of model phase space with respect to initial conditions, and model expansion to include the photosynthetic light reactions and the effects of changes in light intensity on cellular metabolism.

Our current biological research, using global trancriptomic- and proteomics-based approaches, is supporting extensive computational analyses; this research is focused on elucidating the regulation of algal fermentative pathways and identifying various pathways that either directly or indirectly impact H_2 and biofuels production. We recently used 2D-differential gel electrophoresis (DIGE) and shotgun mass spectrometry to identify, quantify and compare proteins present under anoxic, H_2-producing conditions [dark and sulfur-deprived] to those present under oxic, non-H_2-producing conditions. Preliminary DIGE results indicate that the levels of 189 proteins are similar under oxic and anoxic conditions, 41 proteins are higher during anoxic growth and 69 proteins are higher during oxic growth. Furthermore, mass-spectrometry-based, shotgun-proteomic experiments (using the LTQ-Orbitrap system) with whole cell protein extracts detected 1485 proteins under anoxic conditions, 853 proteins under sulfur-deprivation conditions, and 1664 proteins under oxic conditions; analyses of this dataset is underway. Finer resolution will be achieved in the future by proteomic dissection of specific subcellular compartments. In-depth comparative analyses of the specific proteins identified under the various conditions will be discussed.

In summary, high-throughput 'omics' techniques are being used to input transcript, protein, and metabolism knowledge (from both parental and mutant strains) into computational models that explore metabolism (and its flexibility) in the green alga, *Chlamydomonas*. The coupling of experimental results, mutant analyses, and *in silico* modeling is expected to improve our understanding of the complexity of *Chlamydomonas* metabolic networks, including ways in which the cells adjust to changing environmental conditions (i.e., by modulating metabolite fluxes), and how blocking steps in specific metabolic pathways alter metabolite flow. This information as a whole will suggest critical strategies for engineering *Chlamydomonas* metabolism for more efficient production of H_2 and/or other algal biofuels.

References

1. Mus, F., Dubini, A., Seibert, M., Posewitz M.C., and Grossman A.R. (2007) "Anaerobic acclimation in Chlamydomonas reinhardtii: Anoxic gene expression, hydrogenase induction and metabolic pathways", J. Biol. Chem. 282 (35), 25475-25486.
2. Dubini, A., Mus, F., Seibert, M., Grossman, A.R., and Posewitz M.C., (2009) "Flexibility in anaerobic metabolism as revealed in a mutant of Chlamydomonas reinhardtii lacking hydrogenase activity ", J. Biol. Chem. 284, 7201-7213.
3. Chang, C.H., Graf, P., Alber, D.M., Kim, K., Murray, G., Posewitz, M., and Seibert M., (2008) "Photons, Photosynthesis, and High-Performance Computing: Challenges, Progress, and Promise of Modeling Metabolism in Green Algae", J. Phys. Conf. Ser. 125: 012048.

135

Development of Biologically-Based Assays to Study Rate-Limiting Factors in Algal Hydrogen Photoproduction

Alexandra Dubini,[1] Matthew Wecker,[2] and **Maria L. Ghirardi**[1]* (maria_ghirardi@nrel.gov)

[1]National Renewable Energy Laboratory, Golden, Colo. and [2]GeneBiologics, LLC, Boulder, Colo.

Project Goals: See below.

Photobiological H_2 production from water is a clean, non-polluting and renewable technology. Although the potential light conversion efficiency to H_2 by biological organisms is theoretically high (about 10%), the system is currently limited by biochemical and engineering constraints. These limitations include (but are not restricted to) the extreme O_2 sensitivity of the biological hydrogenases and the low availability of reductants to the hydrogenase due to the existence of competing metabolic pathways.

To addresses the O_2 sensitivity issue, our research is developing a new, biologically-based assay to screen large microbial populations for improved H_2-production properties. This novel assay is based on the H_2-sensing system of the photosynthetic bacteria *Rhodobacter capsulatus*. The H_2-sensing system is being optimized as a green fluorescence protein-based reporter-assay for heterologous hydrogen production by the bacteria. It will then be used to screen for O_2 tolerant [FeFe]-hydrogenases generated through directed-evolution techniques. The hydrogenases of *Clostridium acetobutylicum, Chlamydomonas reinhardtii* and *Bacteroides thetaiotaomicron*, along with their respective assembly proteins have been introduced into broad host range vectors and are being shuttled into *R. capsulatus*.

To address the issue of competitive metabolic pathways with H_2 production, we have started using a yeast two hybrid. A complex library of 10^7 preys has been constructed and 8 different baits are being used to screen the library. Those baits include two hydrogenases and the 6 potential partner ferredoxins. We wish to deconvolute the hydrogenase interactors *in vivo* by fishing preys from the library and hopefully discover which of one the ferredoxins is the direct electron donor. We also aim at understanding the ferredoxin interaction network as they are key players in the hydrogenase pathway and are electron donor to many other competitive pathways. The screens are in progress and will help us having a better idea of the metabolic network involved in the hydrogen production pathway.

136

Transcriptome and DNA Methylome Analysis of Algae Using Ultra-High-Throughput Sequencing

David Casero Díaz-Cano,[1] Shawn Cokus,[1] Madeli Castruita,[2] Steven Karpowicz,[2] Sabeeha Merchant,[2] and **Matteo Pellegrini**[1] (matteop@mcdb.ucla.edu)

[1]Department of Molecular, Cell and Developmental Biology and [2]Department of Chemistry and Biochemistry, University of California, Los Angeles

Project Goals: Develop algal biofuels.

High-throughput sequencing has undergone remarkable increases in efficiency over the past few years. Some of the most efficient machines are those produced by Illumina, which sequence tens of billions of bases per week. We are currently using these sequencers in to study the transcriptome of algae such as *chlamydomonas*. *Chlamydomonas reinhardtii*, a unicellular eukaryote in the plant lineage, has been exploited in the laboratory over the last 50 years as a model organism for the study of eukaryotic photosynthesis. The advent of massively parallel short read sequencing technology opens the door to (near) full coverage of the *Chlamydomonas* transcript map via deep sequencing of mRNAs. To evaluate the potential of Illumina's Solexa technology for a) generating a whole transcriptome for *Chlamydomonas*, b) identifying differentially expressed genes, and c) reconstructing gene models *de novo*, we analyzed RNAs isolated from a variety of conditions. We have verified that these libraries may be used to quantitatively estimate transcript fold changes in different conditions using existing gene models. We are also developing a new annotation pipeline using only the short read sequencing data, and have shown that these approaches allow us to accurately reconstruct gene models.

137

Cell-Free Synthetic Pathway Biotransformations (SyPaB) for Producing Hydrogen and Even Fixing Carbon Dioxide

Y.-H. Percival Zhang[1,2*] (ypzhang@vt.edu), Jonathan Mielenz,[2,3] and Mike Adams[2,4]

[1]Biological Systems Engineering Dept., Virginia Tech, Blacksburg; [2]DOE BioEnergy Science Center (BESC), Oak Ridge, Tenn.; [3]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tenn.; and [4]Dept. of Biochemistry and Molecular Biology, University of Georgia, Athens

Project Goals: Demonstrate a new direction for synthetic biology — cell free synthetic pathway biotransformation (SyPaB) Produce 12 moles of hydrogen from low-cost cellulosic materials for the first time Solve several challenges for cell-free SyPaB Propose a novel pathway for CO_2 fixation by using electricty

Cell-free synthetic pathway biotransformations (SyPaB), a new direction of synthetic biology, is to implement complicated biochemical reaction network by assembling a number of purified enzymes and coenzymes.[3] As compared to microbial fermentation, SyPaB has several advantages: (i) high product yields without synthesis of cell mass or formation of by-products; (ii) great engineering flexibility (i.e., easy assembly and control when the building blocks are available); (iii) high product titer; (iv) fast reaction rate; and (v) broad reaction condition. But SyPaB suffers from several obvious weaknesses, such as no ability for self-duplication, costly enzymes, enzyme deactivation, costly and labile coenzymes, and so on.

We have designed the non-natural pathways that completely oxidize starch or cellodextrins by using water as an oxidant for generation of **12 mol of hydrogen per mol of glucose equivalent**[4,5] (Fig. 1). These catabolic pathways comprise substrate phosphorylation mediated by phosphorylases, pentose phosphate pathway, and hydrogenesis mediated by hydrogenase. Also, we have increased overall hydrogen production rate by 10 fold through pathway optimization, high temperature and high substrate concentrations.[4] The above reaction is a unique entropy-driven chemical reaction, i.e., low-temperature heat energy is absorbed and is converted to chemical energy – hydrogen that we can utilize for the first time. Also, we propose to use renewable carbohydrate as a hydrogen carrier to solve hydrogen storage challenge (Fig. 2).[6]

The opinion that SyPaB is too costly for producing low-value biocommodities are mainly attributed to the lack of stable standardized building blocks (e.g., enzymes or their complexes), costly labile co-enzymes, and replenishment of enzymes and co-enzymes. The economical analyses clearly suggest that developments in stable enzymes or their complexes as standardized parts, efficient coenzyme recycling, and use of low-cost and more stable biomimetic coenzyme analogues, would result in much lower production costs than do microbial fermentations because the stabilized enzymes have more than three orders of magnitude higher weight-based total turn-over numbers than microbial biocatalysts, although extra costs for enzyme purification and stabilization are spent[3] (Fig.3).

Fig. 4 clearly suggest that the ultimate hydrogen production costs would be as low as $1.50 per kg, where carbohydrate accounts for 80% of the final product price[2]. Developing thermostable enzymes with TTN_W of > 100,000 are easily reached based on our experiences.[7,8] For example, we have obtained three recombinant thermophilic building blocks — #2 *Clostridium thermocellum* phosphoglucomutase[7], #4 *Thermotoga maritima* 6-phosphogluconate dehydrogenase,[8] and #11 *T. maritima* fructose bisphosphatase expressed in *E. coli*, all of which have TTN_W of > 200,000 at ~60 °C. The above results suggest that discovery and utilization of highly stable thermophilic enzymes from extremophiles that have known genomic sequences are highly operative. Recycling NAD with TTN of 1,000,000 has been reported and the

use of less costly biomimetic NAD will be more economically promising, as shown in Fig. 4.[6,9]

In addition to high-yield generation of hydrogen from biomass sugars, we have designed a novel artificial photosynthesis pathway that can utilize electricity to fix CO_2 for producing ethanol and amylose (Fig. 5). When this process is implemented on large scales, it would solve several challenges for sustainability, such as CO_2 fixation, electricity storage, food production, transportation fuel production, water conservation or maintaining an ecosystem for space travel (the concept paper is under review for publication).

References

1. Y.-H.P. Zhang, Microbe 4 (2009) In press.
2. Y.-H.P. Zhang, J.-B. Sun, A.-P. Zeng, J.-J. Zhong, Curr .Opin. Microbiol. (2010) Invited/In preparation.
3. Y.-H.P. Zhang, Biotechnol. Bioeng. Accepted (2010) http://dx.doi.org/10.1002/bit.22630.
4. X. Ye, Y. Wang, R.C. Hopkins, M.W.W. Adams, B.R. Evans, J.R. Mielenz, Y.-H.P. Zhang, ChemSusChem 2 (2009) 149-152.
5. Y.-H.P. Zhang, B.R. Evans, J.R. Mielenz, R.C. Hopkins, M.W.W. Adams, PLoS One 2 (2007) e456.
6. Y.-H.P. Zhang, Energy Environ. Sci. 2 (2009) 272-282.
7. Y. Wang, Y.-H.P. Zhang, J. Appl. Microbiol. 108 (2010) 39-46.
8. Y. Wang, Y.-H.P. Zhang, Microb. Cell Fact. 8 (2009) 30.
9. J.D. Ryan, R.H. Fish, D.S. Clark, ChemBioChem 9 (2008) 2579-2582.

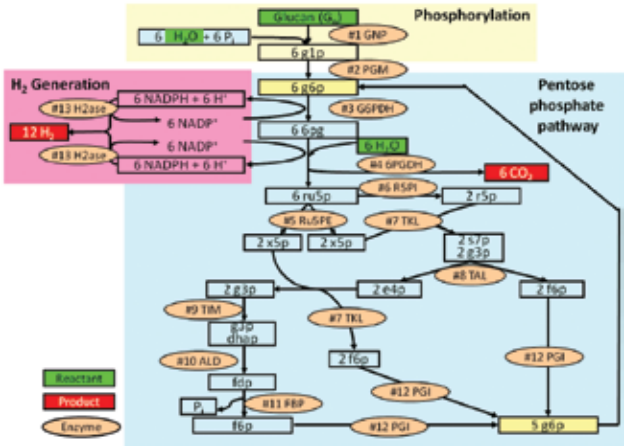


Fig. 1. The non-natural SyPaBs for high-yield hydrogen generation from starch or cellulosic materials [4, 5].

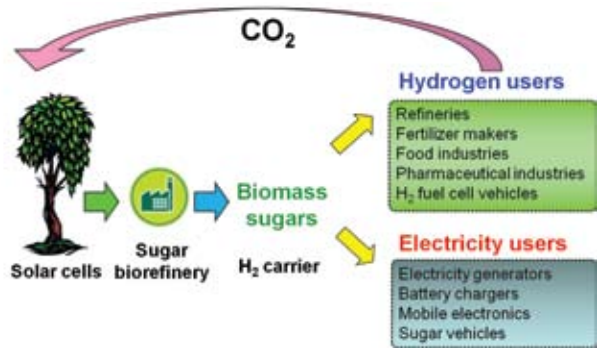


Fig. 2. Scheme of the hydrogen economy based on renewable carbohydrates [1].

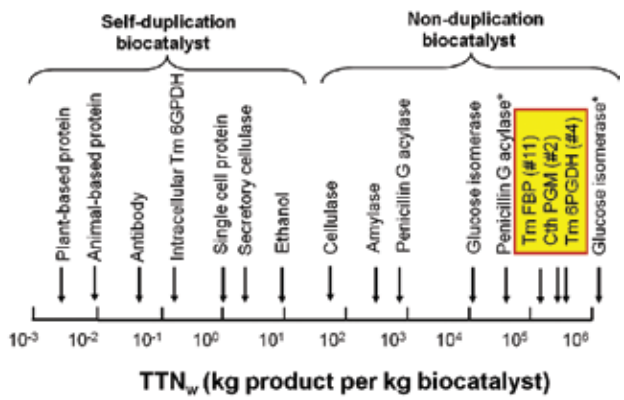
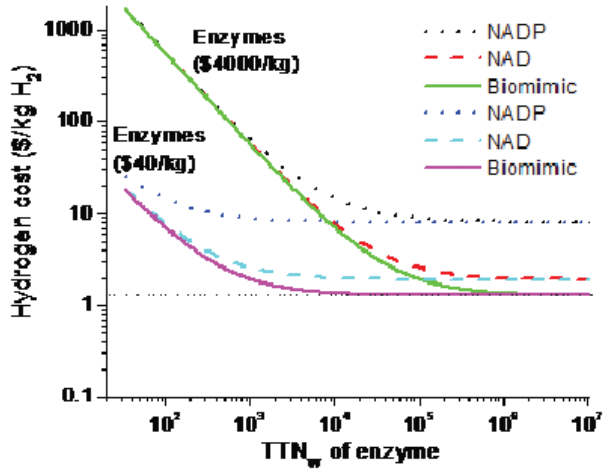


Fig. 3. Comparison of weight-based total turn-over number (TTNW) of self-duplication living biocatalysts and non-duplication enzymes [1, 3].



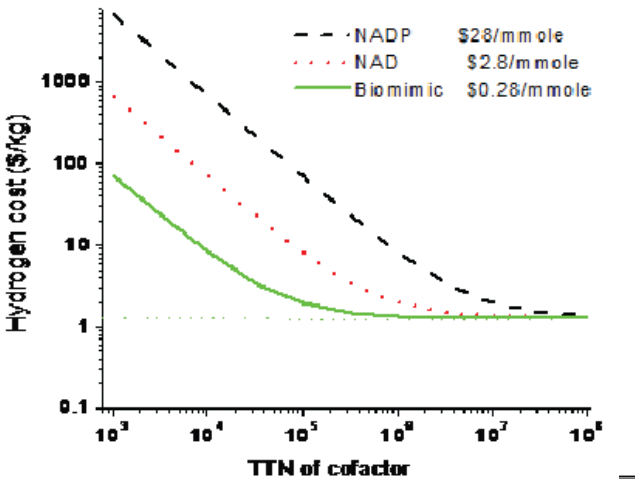


Fig. 4. The estimated hydrogen production costs in terms of turnover number of enzymes and co-enzymes. Carbohydrate prices are $0.18 per kg [2].

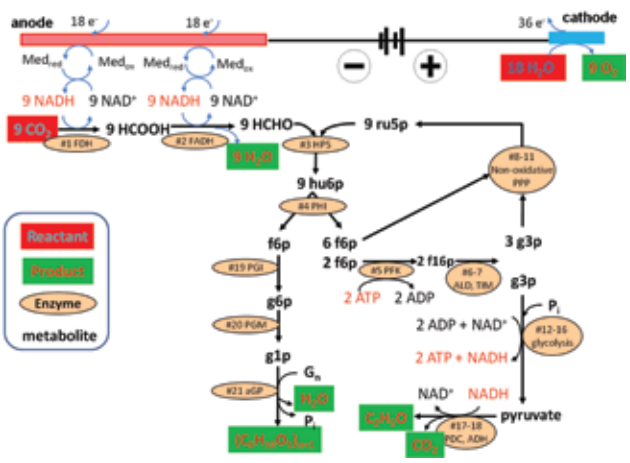


Fig. 5. Artificial photosynthesis that can fix CO_2 by using electricity for producing biofuels and food (under review).

Computing for Bioenergy

138

Bayesian Computational Approaches for Gene Regulation Studies of Bioethanol and Biohydrogen Production

Charles E. Lawrence[1] (charles_lawrence@brown.edu), Lee Newberg,[2,3] William Thompson,[1] and Lee Ann McCue[4]*

[1]Center for Computational Molecular Biology and the Division of Applied Mathematics, Brown University, Providence, R.I.; [2]The Wadsworth Center, New York State Dept. of Health, Albany; [3]Dept. of Computer Science, Rensselaer Polytechnic Institute, Troy, N.Y.; [4]Pacific Northwest National Laboratory, Richland, Wash.

http://www.brown.edu/Research/CCMB/

Project Goals: We are extending our phylogenetic Gibbs sampling algorithms to reconstruct the joint posterior space of the ancestral states of regulatory motifs and developing point estimates and confidence limits for these discrete high-dimensional objects. We are also applying our existing models and technologies, along with the above modifications, to clades of alpha-proteobacterial species, to identify regulatory mechanisms and reconstruct the ancestral states of the regulatory networks for the efficient fermentation of sugars to ethanol and the production of biohydrogen.

Decreasing America's dependence on foreign energy sources and reducing the emission of greenhouse gases through the development of biofuels are important national priorities. These priorities have catalyzed research on cellulosic ethanol as a clean, renewable energy source to replace fossil fuels, and biohydrogen as a carbon-free energy carrier. Turning these biofuels into viable alternative energy sources requires further research into the degradation of cellulose and fermentation of the resulting sugars, and the metabolic and regulatory networks of biohydrogen production. The genomes of many of the microbial species capable of these processes have been sequenced by the GTL and other programs, and many more are expected soon. These sequence data provide a wealth of information to explore nature's solutions for the production of biofuels. In particular, among the over 170 α-proteobacterial species with genome sequence data available are several species with metabolic capabilities of interest, including efficient fermentation of sugars to ethanol and the ability to produce hydrogen. Understanding the regulatory mechanisms and complex interplay of metabolic processes in these species is key to realizing the promise of biofuels. Thus, our research goal is to identify the ensemble of solutions that have been explored by the α-proteobacteria to regulate the metabolic processes key to biofuel production.

* Presenting author