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

Anindita Bandyopadhyay[1], Jana Stöckel[1], Xueyang Feng[2], Yinjie Tang[2], Sujata Mishra[3], Xiaohui Zhang[3], Louis A. Sherman[3], Uma K. Aryal[4], Stephen J. Callister[4], Janani I. Shutthanandan[4], Carrie D. Nicora[4], Ljiljana Pasa-Tolic[4], Dave W. Koppenaal[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, IN; [4]Biological Sciences Division and [5]Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, WA.

Project number - DE-FC02-07ER64694

Project url: sysbio.wustl.edu/Pakrasi/projects/hydrogen.php

Project Goals: The aim of this project 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.

Comparative genomics and hydrogen production in Cyanothece: Analysis of the complete genome sequences of six *Cyanothece* strains (ATCC 51142, PCC 7424, PCC 7425, PCC 7822, PCC 8801, and PCC 8802) indicates that this genus maintains a plastic genome, incorporating new metabolic capabilities while simultaneously retaining archaic metabolic traits, a unique combination which provides the flexibility to adapt to various ecological and environmental conditions (Bandyopadhyay et al., 2011). Our study uncovered the presence of a large and contiguous nitrogenase gene cluster in four of the five newly sequenced strains. These strains were also analyzed for their ability to fix nitrogen and produce hydrogen. High rates of aerobic nitrogen fixation and hydrogen production were observed in five of the six strains, distinguishing *Cyanothece* as a genus of unicellular, aerobic nitrogen-fixing cyanobacteria. The contiguity of the nitrogenase gene cluster in *Cyanothece* sp. 7425 is interrupted by a 2.5Mbp insertion and this cluster exhibits a rearrangement of the *nif* genes as seen in some of the anaerobic nitrogen fixing *Synechococcus* strains. *Cyanothece* 7425 also appears to have lost its aerobic nitrogen-fixing ability and exhibits nitrogen fixation and hydrogen production only under anaerobic conditions. In addition, our study revealed the presence of 3 linear chromosomal elements in *Cyanothece* 7822. These linear elements, like the linear chromosome in *Cyanothece* 51142, harbor genes involved in various carbohydrate metabolism pathways and represent a unique feature of the genus *Cyanothece* compared to other sequenced cyanobacteria. A comparison of the genomes of the six *Cyanothece* strains revealed the presence of several pathways involved in fermentative metabolism, traits usually observed in non-oxygenic microbes. The ability of this genus to maintain a suboxic intracellular environment for a significant part of a diurnal cycle is unique and allows the *Cyanothece* cells to provide a platform for both aerobic and anaerobic metabolic processes. Our analysis suggests that members of the genus *Cyanothece* can be appealing as model organisms for studies pertaining to biohydrogen production.

Metabolomic studies: We have developed a web-based platform (MicrobesFlux) for generating and reconstructing metabolic models for annotated microorganisms. The MicrobesFlux is able to load the metabolic network (including enzymatic reactions and metabolites) of over 1,100 species from KEGG database (Kyoto Encyclopedia of Genes and Genomes) and then automatically converting it to a metabolic model. The platform also provides diverse customized

tools, such as gene knockouts and introduction of heterologous pathways, for users to redefine the model network. The reconstructed metabolic network can be formulated to a constraint-based flux model to predict and analyze the carbon fluxes in the metabolisms. The simulation results can be output in SBML format (The Systems Biology Markup Language). The MicrobesFlux is an installation-free and open-source platform that enables biologists with little programming knowledge to develop metabolic models for newly sequenced microorganisms. Our system allows users to construct metabolic networks of organisms directly from the KEGG database. It also provides users with predictions of microbial metabolism via flux balance analysis. This prototype platform can be a springboard to advanced and broad-scope metabolic modeling of complex biological systems by integrating other "omics" data or ^{13}C-assisted metabolic flux analysis results. This platform is being used for generating metabolic models of the *Cyanothece* strains and for predicting gene knockouts and pathway modifications. MicrobesFlux is available at http://tanglab.engineering.wustl.edu/static/MicrobesFlux.html

Proteomic Studies: A comparative analysis of the proteomes of the six *Cyanothece* strains was carried out to identify similarities and differences in the metabolism of the strains grown under nitrogen sufficient conditions. Among the five strains, the largest number of proteins (3967) could be detected in *Cyanothece* PCC7424 and the lowest (2728) in *Cyanothece* PCC7425. Our results reveal considerable similarities in proteins involved in conserved metabolic and biochemical pathways such as photosynthesis, respiration, N_2-fixation, H_2-production as well as proteins involved in housekeeping functions. Out of a total of 1369 predicted orthologs common to all *Cyanothece* strains, 644 have been detected in the comparative study. However, considerable differences were also evident, particularly in *Cyanothece* PCC7425 where significantly higher numbers of proteins (682 proteins) were uniquely expressed. This may suggest that proteins in *Cyanothece* PCC7425 diverged separately compared to the other strains, possibly as a result of specific selection pressure and subsequent adaptation to environment. Furthermore, in addition to commonly shared phycobilisome (PBS) proteins, we identified 3 phycobiliproteins and 2 PBS-linker polypeptides expressed specifically in PCC7822, PCC7424 and PCC8801 together with phycoerythrin (PE)-operon protein, suggesting their association with PE assembly. While homologs of PSI core proteins PsaAB were identified in all six strains, expression of another PsaB protein (Cyan7822_4989) was only detected in PCC7822.Our results elucidate proteomic footprints of six *Cyanothece* strains, and provide useful targets for future applications in biotechnology.

To determine cellular factors involved in H_2 production, we employed label-free quantitative proteomics in *Cyanothece* ATCC 51142 and PCC 7822 under eight culture conditions. Rates of H_2 production were measured for each condition and compared to changes of protein abundances. Our analysis provides systems level understanding about nitrogenase-mediated H_2 production under aerobic condition providing useful insights into how *Cyanothece* might be optimized for H_2 production.

Publications:

Bandyopadhyay A, Elvitigala T, Welsh EA, Stöckel J, Liberton M, Min H, Sherman LA, Pakrasi HB (2011). Novel metabolic attributes of *Cyanothece*, a group of unicellular nitrogen fixing cyanobacteria. *mBio* 2 (5): e00214-11. doi:10.1128/mBio.00214-11.

This work was supported by funding from DOE-BER (DE-FC02-07ER64694).