Meta-Proteomic Study of a Microbial Community Response to Cadmium Exposure

C.M.R. Lacerda¹, L. Choe², and K.F. Reardon* (reardon@engr.colostate.edu)
¹Colorado State University, Fort Collins, CO and ²Cornell University, Ithaca, NY

Microbial communities are the basis for engineered environmental bioprocesses. However, current methods, including 16S rDNA-based techniques, are incapable of providing information on the function of the community or its members, and thus the community remains a black box from a functional perspective. Microbial community proteomics has the potential to detect proteins expressed in an environment under different conditions. In this approach, a mixture of microorganisms can be viewed as a meta-organism, in which population shifts are a form of functional response. Despite its possible advantages, the use of meta-proteomics has been almost completely unexplored.

In this project, meta-proteomics was used as a tool to obtain functional information about the response of a microbial community to cadmium stress. Cadmium (10 mg/L) was added to a mixed culture and protein samples collected after 0.25, 1, 2 and 3 hours. Comparison of the two-dimensional gels from Cd-exposed and control cultures revealed that the community “reacts” by changing its protein profile, with both increased and decreased expression of substantial numbers of proteins that change with time. Within 15 minutes of exposure, nearly 20% of the proteins detected on the 2-D gels were found to change in level by three-fold or more. Mass spectrometric analysis was also performed to identify proteins that play central roles during the cadmium shock. Fifty proteins have been identified to date. Metabolic enzymes make up the largest functional group, followed by proteins involved with defense, protein synthesis and storage, and energy metabolism. A variety of temporal expression patterns was noted, with protein functional groups displaying one or more patterns.

This study demonstrated that proteins can be identified from a community of unsequenced organisms. Our proteomic analysis revealed significant shifts in the community physiology for both short and long term metal exposure, insights that could not have been obtained using traditional 16S rDNA methods. These results clearly demonstrate the ability of the meta-proteomic approach to detect changes in microbial communities and support its use for determining functional states of a microbial community.
Milestone 2

53

Investigating Novel Proteins in Acidophilic Biofilm Communities

Christopher Jeans¹, Clara Chan², Mona Hwang¹, Jason Raymond¹, Steven Singer¹, Nathan VerBerkmoes³, Manesh Shah³, Robert L. Hettich¹, Jill F. Banfield², and Michael P. Thelen* (mthelen@llnl.gov)

¹Lawrence Livermore National Laboratory, Livermore, CA; ²University of California, Berkeley, CA; and ³Oak Ridge National Laboratory, Oak Ridge, TN

A large fraction of proteins deduced from microbial genomes do not match any recognized sequences, and there is no systematic method at hand to analyze this collection known as “hypothetical proteins”. These have for the most part remained in the hypothetical category because authentication by protein detection or gene expression has been difficult; however, it is likely that a large portion are important, in some cases essential, for microbial fitness under natural conditions. The void in our knowledge regarding this dark side of genomics, representing nearly half of protein sequences, is a formidable obstacle to nearly every investigation enabled by full genome sequence information. In our study of acidophilic biofilm communities, we are focused directly on the problem of hypothetical proteins within a well-defined system in one of the most extreme limits of habitability. These microbes band together in very acidic streams of underground mine tunnels to form floating, matrix-bound, self-sustaining communities that derive electrons from iron (II) while in turn catalyzing the dissolution of pyrite (FeS₂). Such conditions result in the biological generation of acid mine drainage (AMD), ultimately causing many orders of magnitude greater metal dissolution and environmental acidification than abiotic forces alone. Since AMD is so prevalent and potentially disastrous to US inland waters associated with principle energy resources such as coal and uranium, understanding the mechanisms by which biofilm communities generate AMD is a priority with DOE’s Office of Biological and Environmental Research. Also relevant to the Genomics: Genomes to Life Program, our project will provide fundamental scientific contributions resulting from a systematic and thorough study of the hypothetical protein problem, including the development of methods to be used in these and other systems where a description of novel proteins will provide the keys to overall biological function.

In addition to the detailed geochemical and microbiological description of AMD biofilms by Jill Banfield and colleagues at UC Berkeley, the foundation of our current investigation is genomic and proteomic datasets that are directly linked for each of several dominant organisms. Characterization of abundant species in one such biofilm and the associated genome analyses resulted in reconstruction of near complete or partial genomes for two different Leptospirillum bacteria and three archaea¹, and further work is now extending the sequences within this biofilm and also to ones with differing population structures. In our attempt to understand essential biofilm metabolic activities and the partitioning of functions between individual organisms, proteins native to an environmental biofilm were analyzed by shotgun MS proteomics, carried out by Robert Hettich and associates at ORNL. This approach coupled with protein biochemistry confirmed that many hypothetical proteins are expressed at detectable levels in several biofilm organisms, and also indicated that a majority of abundant extracellular proteins are novel². This groundwork has enabled us to examine the novel proteins expressed in biofilms harvested from several neighboring locations, in carefully measured geochemical conditions, using a combination of computational analyses/predictions, and protein biochemistry coupled with MS identification.

To reveal any functional features of proteins overlooked by the first level of bioinformatics, we have drilled further into protein sequence homology, initially using a database containing the completed
genome sequences of 250 prokaryotes. We interrogated this data with several hundred novel protein sequences that were detected by MS proteomics in one of the dominant organisms in the AMD biofilms, *Leptospirillum* group II. Using parameters that can identify distant homology, we were able to classify more accurately “unique” and “conserved” protein sequences. Proteins found only in *Lepto*. II, in addition to those similar to (but distinctly different from) proteins in the other abundant biofilm bacterium, *Lepto*. group III (e.g., *L. ferrodiazotrophum* sp. nov.), are considered unique. Genes encoding these proteins may have arisen in direct response to environmental selection, and are likely to be vital to AMD biofilm functions. Also notable are the similarities found between the novel proteins of *Lepto*. II and III to those in “functionally related” organisms, such as *Geobacter metallireducens* and *Ralstonia metallidurans*, both of which can be found in toxic metal-rich niches. Other unknown proteins that are homologous to 20 – 90 other microbial sequences indicates conservation of core functions that remain unidentified. Descriptions of proteins in this latter category will have perhaps the most impact on genome-enabled microbiology. In addition to individual protein sequence homology, analysis of gene synteny in sequences flanking novel genes is a measure of evolutionary conservation. In some cases clues to functions were found from annotations to neighboring genes within homologous operons, indicating a biochemical pathway or protein complex in which a novel protein resides. These analyses are being used to indicate interesting novel protein targets to pursue, using the abundance inferred from proteomics datasets as a guide.

In addition to these predictive criteria, we also use biochemical properties such as protein mass, isoelectric point and some sequence information (e.g., predicted or experimentally determined signal/transit peptides) to facilitate the isolation and identification of proteins for experimental characterization. This is being pursued mostly in biofilm extracts that have been fractionated into extracellular and membrane-associated proteins, both enriched with respect to novel proteins, with a major goal to characterize the functional components of protein complexes. There are several examples to illustrate this approach. From biofilm extracts, we purified one of the most abundant extracellular proteins, identified its *Lepto*. II gene and found sequence similarity with only one unknown protein sequence (from *Ralstonia metallidurans*) in the 250 fully sequenced microbes. This soluble red cytochrome of 16 kDa contains an unusual heme, has a reduction potential of 640 mV – close to that of soluble iron at pH 2.0 – and rapidly oxidizes iron (II to III), yielding a unique absorption peak at 579 nm (ref. 2). We used antibodies to localize this ‘cytochrome 579’ in biofilms, and found that the protein is highly concentrated near the surface of *Lepto*. II cells. Complementing these results, we have isolated another abundant, novel protein from *Lepto*. II that is primarily membrane associated and has similar spectral characteristics to cytochrome 579, but is a much larger protein. This 57kDa cytochrome forms a complex with cytochrome 579, indicated by the co-electrophoresis of the two proteins in nondenaturing gels. Further testing will establish the role of the larger cytochrome in iron redox reactions, and the function of the two proteins in this complex, towards a more accurate description of the initial steps in electron transfer from iron sulfate in mine solutions to the cellular electron transport system. We have also purified two other protein complexes and several individual proteins from both extracellular and membrane fractions using similar methods, and these are targets for detailed characterization. These several results indicate that our approach can be used to identify novel gene products and biochemical functions.

To process many more proteins in future work, we have started to separate biofilm proteins using a variety of selective chromatographic media and identify the principle proteins in these fractions using LCQ MS. In conjunction with size exclusion and chromatofocusing, these methods provide us with a powerful approach to isolate many of the novel proteins along with any known proteins associated with them, from both the environmental microbial community and from biofilm isolates cultured in lab. Multiple fractions from these columns will be initially assayed for several key biochemical activities such as hydrolases, oxidoreductases and polysaccharide synthases that are suspected to play important roles in biofilm formation and function. The development of this type
of systematic approach for the validation and description of hypothetical proteins, used within a
defined biological system, is a test case for the full utility of proteogenomic information.

This Genomics: Genomes to Life project centered at UC Berkeley is performed under the auspices of the U.S. Depart-
ment of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

References

54

**Proteomics Measurements of a Natural Microbial Community Reveal Information about Community Structure and Metabolic Potential**

Nathan C. VerBerkmoes* (verberkmoes@ornl.gov), Rachna J. Ram\(^2\), Vincent J. Denef\(^2\), Gene W. Tyson\(^2\), Brett J. Baker\(^2\), Manesh Shah\(^1\), Michael P. Thelen\(^3\), Jillian Banfield\(^2\), and Robert L. Hettich\(^1\)

\(^1\)Oak Ridge National Laboratory, Oak Ridge, TN; \(^2\)University of California, Berkeley, CA; and \(^3\)Lawrence Livermore National Laboratory, Livermore, CA

Microbial communities play key roles in the Earth's biogeochemical cycles, but for the most part are
very poorly understood. For example, microorganisms from a number of lineages thrive in acid mine
drainage (AMD), one of the most extreme environments on Earth. Recently, we have completed
a proteogenomic investigation of a natural AMD consortia sample to determine the biochemical
proteins by providing the database necessary for identification of peptides from whole cellular, mem-
brane, and periplasmic/ extracellular fractions of the biofilm samples. The proteomic data provided
the first *in situ* analyses of community structure and metabolic potential. MS-based “shotgun” pro-
etomics measurements with multidimensional liquid chromatography – linear trapping quadrupole
mass spectrometry provided confident measurement of 2036 proteins from the AMD sample, with
identifications corresponding to all five dominant species in the biofilm.

For the complete characterization of the proteome of complex natural microbial communities, we
have developed a systematic experimental plan to push the capabilities of current mass spectrometry
techniques as well as integrate the next generation of technology. This will involve evaluation and
implementation of a higher performance MS technology employing a state-of-the-art LTQ-FT-
Orbitrap (Thermo Electron) instrument for deeper and more comprehensive characterization and
quantification of the proteins and protein complexes important for community structure. The ulti-
mate goal will be to define and demonstrate experimental protocols to begin to unravel the molecular
details of natural microbial communities. Specific tasks are directed at examination of spatially and
temporally distinct AMD biofilms, establishment of advanced mass spectrometry methodology for
more comprehensive proteomic information, elucidation of the details of microbial strain variant
diversity with high mass accuracy and MS\(^3\) experiments, and quantification of protein abundances to
understand the major investments of cellular resources.

Recent work has been focused on the proteome characterization of new AMD microbial commu-
nities distinct but related to the original AMD sample. These include the AB front, C pink, and UBA
Ultraback locations of the mine. These proteomes are currently being analyzed by similar methods

* Presenting author
used in the previous study and over 1,500 proteins have been identified from each biofilm. We are currently comparing these datasets to the original proteome dataset to determine conserved protein expression between all samples and to determine those proteins expressed only in given biofilms.

We are in the process of transitioning from current MS technology to more advanced techniques to provide better depth of coverage, smaller sample size requirements, and to unravel the details of microbial strain variation in AMD communities. The proteomic analyses of previous and current samples generated many mass spectra that could not be assigned to any protein. There are several reasons for this, two of which relate to problems with the genomic database. Peptides from proteins that differ significantly from those encoded in the community genomic dataset are unlikely to be matched. If there are substitutions throughout the protein, matching of all peptides (or all but one peptide) may be precluded. In this case, the protein is unidentifiable. Alternatively, a subset of peptides may be found, resulting in (i) assignment of the peptides to a protein that is not identical to the dominant form in the community and (ii) incorrect evaluation of the relative abundance of that protein. For these reasons, proteomic analyses must move away from composite genome sequence-based analyses to utilize the full database of gene types detected in each community. A second reason for unassigned spectra is the presence of organisms in the community that were not sampled in the genomic dataset (due to different organism membership or incomplete genome coverage). It is important to expand the analytical capabilities of the MS-based proteomics methodology for identifying and obtaining detailed biological information for every protein in an AMD sample. This requires identification of a very large number of proteins and their post translations modifications, truncation products, and associated strain variability. The greatest challenges for analysis of a microbial community are the large number and relatively low concentrations of many proteins within the system. For all initial studies, protein separation were conducted on an integrated 2-dimensional nano columns equipped with nanospray MS. For the next phase of this project, a new hybrid LTQ-FT-Orbitrap will be integrated into the pipeline and be used for all sample characterization. High mass accuracy, high sensitivity and high dynamic range capabilities are features of this new instrument, which can be coupled with a linear ion trap mass spectrometer capable of rapid data-dependent MS/MS and MS$^3$. This instrument can easily be coupled with multidimensional chromatography. The high mass accuracy will allow for much more confident identifications of MS/MS spectrum due to the accuracy at which the parent mass of the peptide can be determined (<3 ppm). Furthermore, if desired, MS/MS spectra can also be analyzed in the Orbitrap allowing for <3 ppm mass accuracies on fragment ions in MS/MS or MS$^3$ experiments. This is critical in community samples where databases are larger and sample variability will be much greater. More comprehensive proteomic analysis will be achieved due to the high dynamic range of the instrument, because data-dependent MS/MS events will enable measurement of peptides that cannot be detected in full scan mode on the linear ion trap. The linear ion trap is capable of MS$^3$ experiments, which, along with high mass accuracy of the intact peptide and MS/MS spectra, will greatly increase the accuracy of de novo sequencing techniques outlined below.

To interpret the LC-MS/MS data from the samples, proteome bioinformatics, in particular, Sequest and DBDigger search engines, will be employed and web-based data repositories will be created. One major goal of this subproject is to develop new informatic tools for data analyses, data mining, and data display critical for the challenges associated with these studies. New computational techniques will immediately be implemented into the proteome informatics pipeline. We are currently evaluating a number de novo sequencing algorithms for identification of unmatched peptides. The use of high mass accuracy MS/MS spectra will greatly enhance the accuracy of these algorithms for correctly determining peptide sequences directly from MS/MS spectra.
The final task of this work involves absolute and relative quantification of proteins from AMD biofilms. While exact quantification of proteins from simple isolate microbial proteomes is challenging, the quantitation of proteins directly from environmental samples provides a much greater diversity of challenges. We are currently investigating absolute quantification of proteins in the AMD community samples with isotopically labeled synthetic peptides from well characterized peptides from the community. The use of isotopic encoded affinity tags (ICAT) and $^{18}$O Water labeling will be investigated for relative quantification of proteins between multiple AMD biofilms.

Research sponsored by the Genomics:GTL program, Office of Biological and Environmental Research, U.S. Department of Energy, under contract No. DE-AC05-00OR22725 with Oak Ridge National Laboratory, managed and operated by UT-Battelle, LLC.

Computational Algorithms and Software Tools for Quantitative Shotgun Proteomics

C. Pan, B. Zhang, G. Kora, N.C. VerBerkmoes, D. Tabb, W.H. McDonald, D. Pelletier, E. Uberbacher, G. Hurst, R. Hettich, and N.F. Samatova* (samatovan@ornl.gov)

Oak Ridge National Laboratory, Oak Ridge, TN

Biological organisms respond to many environmental or physiological stimuli by adjusting the expression levels of proteins. The quantification of protein abundance and detection of differential protein expression under various experimental conditions are fundamental and challenging problems in proteomics. Addressing these problems with classical two-dimensional gel electrophoresis (2DE) has some obvious disadvantages due to its low detection sensitivity and linearity, poor solubility of membrane proteins, limited loading capacity of gradient pH strips, low reproducibility of gels, limited throughput, and small linear range of visualization procedures. Recently, alternative approaches based on either stable isotope labeling or label-free mass spectrometry (MS) shotgun proteomics have emerged as a high throughput technique for measuring the relative abundance of thousands of proteins from different cell cultures.

Because of highly complex sample handling in proteome measurements (as described in other abstracts from the Center for Molecular and Cellular Systems, CMCS), proteome quantification requires rigorous statistical approaches. To provide robust quantification of relative protein abundance and sensitive detection of biologically significant differential and correlated protein expression, we are developing advanced statistical methods coupled with suitable software tools that are made available to users as open source (please, pick up the distribution CD with the software or email a request).

ProRata for relative quantification of mixed stable-isotope-labeled proteomes

To extract relative peptide and protein amounts from mass spectrometric measurement of mixed stable-isotope-labeled proteomes, we developed a computer program, called ProRata. To improve both quantification accuracy and quantification confidence, we systematically optimized the core analysis steps for robust quantification: chromatographic peaks detection, peptide relative abundance estimation, and protein relative abundance estimation. Our novel parallel paired covariance algorithm has largely enhanced the signal-to-noise ratio of the two isotopologues’ chromatograms and, as a result, has enabled much more accurate peak detection. Principal component analysis (PCA) was employed to estimate peptide abundance ratios and demonstrated superior estimation accuracy than the tradi-
tional methods based on peak height and peak area. It was observed that the relative quantification of the standard proteome mixtures is of highly variable accuracy for peptides and consequently for proteins. To estimate quantification error, we proposed a novel signal-to-noise measure derived from principal component analysis and showed a linear correlation of this measure with the peptide ratio estimation in the standard mixtures. Finally, maximum likelihood estimation (MLE) was used for protein relative quantification from the PCA-estimated abundance ratios of its proteolytic peptides. MLE not only showed more accurate protein quantification and better coverage than the widely-used RelEx program but also a more robust estimate of a confidence interval for each differential protein expression ratio. For an automated data processing and streamlined data visualization, these algorithms were integrated into a ProRata computer program (see Figure for a sample display).

Detection of differential and correlated protein expression in label-free shotgun proteomics

We performed a systematic analysis of various approaches to quantifying differential protein expression in the eukaryotic *Saccharomyces cerevisiae* and prokaryotic *Rhodopseudomonas palustris* LC-MS/MS label-free shotgun proteomic data. First, we showed that, among three sampling statistics, the spectral count has the highest technical reproducibility followed by the less-reproducible peptide count and relatively non-reproducible sequence coverage. Second, we used spectral count statistics to measure differential protein expression using four statistical tests: Fisher’s exact test, G test, AC test, and t-test. For the yeast data set with spike proteins, the first three tests performed similarly on a pair-wise comparison of multiple experiments. Their False Discovery Rate (FDR) was less than 0.4% for a 10-fold change and less than 0.7% for a 5-fold change, even with a single replicate. For a 2-fold change,

Figure. The display of ProRata analysis results for $^{14}$N-$^{15}$N labeled *R. palustris* proteome. The ProRata graphical user interface consists of four parts: protein table (top left), peptide table (bottom left), graph pane (top right), and text pane (bottom right). The graph pane contains seven graphs: Likelihood Plot, Sequence Coverage Plot, Ion Chromatogram, PPC Chromatogram, PCA Plot, MS1 Scan, and MS2 Scan. All graphs can be detached from the graph pane.
FDR could exceed 10% with one replicate, but was less than 5% or 3% with two or three replicates, respectively. The t-test performed the best with three replicates. Third, we generalized the G test to increase the sensitivity of detecting differential protein expression under multiple experimental conditions. Out of 1,664 detected *R. palustris* proteins in the LC-MS/MS experiment, the generalized G test identified 1,119 differentially expressed proteins under six growth conditions, including photoheterotrophic, chemoheterotrophic, nitrogen fixation, photoautotrophic, stationary phase, as well as benzoate as an alternate carbon source. Unlike 2-fold change under two conditions, the generalized G test differentiated 300 more proteins under six conditions. Furthermore, among the 625,521 protein pairs between these 1,119 differentially expressed proteins, operon pairs were much stronger co-expressed than the non-operon ones. Finally, we identified six protein clusters with known biological significance by combining cluster analysis with functional annotation of these differentially expressed proteins. In summary, the proposed generalized G test using spectral count sampling statistics is a viable methodology for robust quantification of relative protein abundance and for sensitive detection of biologically significant differential and correlated protein expression under multiple experimental conditions in label-free shotgun proteomics.

Research sponsored by the DOE Genomics:GTL Program.

56

High Throughput Comprehensive and Quantitative Microbial Proteomics: Production in Practice


Pacific Northwest National Laboratory, Richland, WA

Significance: Capabilities for quantitative proteomics measurements of steadily increasing throughput and quality have been implemented and are being applied to studies with a range of microbial systems.

With recent advances in whole genome sequencing for a growing number of organisms, biological research is increasingly incorporating higher-level “systems” perspectives and approaches. Key to supporting systems-level advances in microbial and other biological research at the heart of the DOE Genome GTL program is the ability to quantitatively measure the array of proteins (i.e., the proteome) for various organisms under many different conditions.

Among the challenges associated with making useful comprehensive proteomic measurements are identifying and quantifying large sets of proteins whose relative abundances span many orders of magnitude. Additionally, these proteins may vary broadly in chemical and physical properties, have transient and low levels of modifications, and be subject to endogenous proteolytic processing. Ultimately, such measurements and the resulting insight into biochemical processes are expected to enable development of predictive computational models that could profoundly affect environmental clean-up, understandings related to climate, and energy production by e.g., providing a more solid...
basis for mitigating the impacts of energy production-related activities on the environment and human health.

A “prototype high throughput production” lab established in FY 2002 was an early step towards implementing higher throughput proteomics measurements. Operations within this lab remain distinct from technology development efforts, both in laboratory space and staffing. This step was instituted in recognition of the different staff “mind sets” required for success in these different areas, as well as to allow “periodic upgrades” of the technology platform in a manner that does not significantly impact its production operation. The result has been faster implementation of technology advances and more robust automation of technologies that improve overall effectiveness.

The biological applications of the technology and associated activities are the subject of a separate, but interrelated project (J. K. Fredrickson, PI), involving studies of a number of microbial systems (e.g., Shewanella oneidensis, Geobacter sulfurreducens, Rhodobacter sphaeroides) in collaboration with leading experts on each organism. These studies have demonstrated the capability for automated high-confidence protein identifications, broad proteome coverage, and for exploiting both stable isotope labeling and label-free methods to obtain high precision in protein abundance measurements.

With a paradigm established for high throughput proteomic measurements, our primary goal now is to significantly increase data quality, as well as throughput. A significant challenge is how to maximize the information content derived from large and complex data sets such that the researcher can gain novel biological insights. Thus, a key component of our program involves developing the informatics tools needed to quantify and define the quality of data, as well as the tools to make the results broadly available and understandable to the researchers. Efforts currently in progress aim to:

- Significantly increase the overall data production by more than an order of magnitude in conjunction with increased data quality, providing data that are quantitative and have statistically-based measures of quality.
- Extend the application to an increasing number of different kinds of post-translation modifications.
- Provide the infrastructure and informatics tools required to efficiently manage, use, and disseminate large quantities of data generated by GTL “users.”

This presentation will highlight the advances in providing high quality data with statistically-founded measures of quality, while providing increased measurement throughput. The advances will be illustrated in the context of applications to microbes of interest to the GTL program.

Acknowledgements: This research is supported by the Office of Biological and Environmental Research of the U.S. Department of Energy. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute through Contract No. DEAC05-76RLO 1830.

* Presenting author
H$_2$O$_2$-Induced Stress Responses of *Shewanella oneidensis* MR-1

T. Li$^{1,2}$, J. Guo$^2$, D. Klingeman$^1$, L. Wu$^2$, X. Liu$^3$, T. Yan$^2$, Y. Xu$^2$, A. Beliaev$^4$, Z. He$^{5,7}$, T.C. Hazen$^{6,7}$, A. P. Arkin$^{6,7}$, J. Zhou$^{1,5,7}$ (jzhou@ou.edu)

$^1$Oak Ridge National Laboratory, Oak Ridge, TN; $^2$University of Georgia, Athens, GA; $^3$Central South University, Changsha, China; $^4$Pacific Northwest National Laboratory, Richland, WA; $^5$University of Oklahoma, Norman, OK; $^6$Lawrence Berkeley National Laboratory, Berkeley, CA; and $^7$Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

*Shewanella oneidensis* is a facultative anaerobe that possesses a complex electron-transport system, which allows the coupling of metal reduction to bacterial energy generation. The ability of *S. oneidensis* to reduce toxic heavy metals makes it an ideal candidate for bioremediation of contaminated sites. However, many questions remain about how the organism responds to and functions in various environmental conditions. To better understand the molecular biology of this metal-reducing bacterium, we have employed whole-genome microarrays and bioinformatics to investigate the global gene expression profiles of wild type and oxyR mutant strains in response to H$_2$O$_2$ induced oxidative stress.

In wild type *Shewanella* cells, a total of 1,092 genes showed significant changes in the expression level under at least one condition tested. Many genes showed dose-dependent expression pattern and were differentially regulated at different time points. Comparison of the gene expression kinetics suggests that *S. oneidensis* possesses complex regulatory systems to protect the cells from oxidative stress. Among the genes that are immediately up regulated under stress conditions are the genes with established functions in oxidative stress, such as the alkyl hydroperoxide reductase (Ahp) gene, the catalase (Kat) genes, and the stress response DNA-binding protein (dps) gene, as well as various genes that have not been previously described to be involved in the oxidative stress responses of other bacterial species, including some iron- and sulfur- responsive genes. In addition, an oxyR homologue is identified and characterized. Phenotypic and enzymatic studies along with microarray analysis indicate that *S. oneidensis* oxyR serves as a dual function transcription regulator, which activates the expression of many oxidative stress genes in response to H$_2$O$_2$ stress while repressing the expression of a catalase gene (KatB) and the nonspecific DNA binding protein gene (dps) under normal growth condition.

Iron is an essential nutrient with limited bioavailability. Due to its ability to react with and catalyze the generation of toxic radicals, iron when overloaded may pose a big threat to living cells and tissues; and so the acquisition of iron is usually tightly regulated in biological systems. In contrast to *E. coli*, the wild type *Shewanella oneidensis* fur gene is not significantly affected by H$_2$O$_2$, while many iron inducible genes showed drastic up-regulation under H$_2$O$_2$ stress conditions. These genes include the TonB1 iron transport system genes, a bacterioferritin gene bfr, the ferrous iron transport genes, and some DNA repair and metabolism genes. To investigate the H$_2$O$_2$-induced regulation of the iron responsive genes in *S. oneidensis*, we also included a fur deletion mutant in our microarray studies. Not surprisingly the fur deletion mutant showed hypersensitivity to H$_2$O$_2$ stress. Further analysis by microarray study of the fur mutant under oxidative stress condition reveals interesting regulation pattern. In fur mutant the iron responsive genes are highly expressed when compared with the wild type strain, which is in agreement with the repressor role of fur. However in contrast to wild type strain fur mutant showed little or diminished induction of the iron inducible genes in response to H$_2$O$_2$, indicating that fur directly or indirectly mediates the upregulation of these genes by H$_2$O$_2$ treatment, and the regulation is apparently independent of the transcription of the fur gene since the expression level of fur does not show significant change under the stress condition. More interestingly, many of the fur regulated iron responsive genes also exhibit unusually high expression in oxyR deletion.
mutant, and when $H_2O_2$ is applied to the oxyR mutant the transcription of the iron genes further increases, which coincides well with the regulation pattern of these genes in wild type strain; these data suggest that oxyR mutation may play a role in modulating fur activity. At the time being we are applying bioinformatics.

58

Application of High Resolution Proteomics to Characterize Microbial Systems for Metal Reduction and Photosynthesis

Mary S. Lipton*, (mary.lipton@pnl.gov), Margie F. Romine1, Stephen J. Callister1, Kim K. Hixson1, Samuel O. Purvine1, Angela D. Norbeck1, Joshua N. Adkins1, Matthew E. Monroe1, Yuri A. Gorby1, Carrie D. Goddard1, Richard D. Smith1, Jim K. Fredrickson1, Derek Lovley2, Richard Ding2, Timothy Donohue1, Miguel Dominguez2, Christine Tavano2, Samuel Kaplan4, Xiaihua Zeng4, Jung Hyeob Roh4, and Frank Larimer5

1Pacific Northwest National Laboratory, Richland, WA; 2University of Massachusetts, Amherst, MA; 3University of Wisconsin, Madison, WI; 4University of Texas Medical School, Houston, TX; and 5Oak Ridge National Laboratory, Oak Ridge, TN

Exploiting microbial function for purposes of bioremediation, energy production, carbon sequestration and other missions important to the DOE requires an in-depth and systems level understanding of the molecular components of the cell that confer its function. Inherent to developing this improved understanding is the ability to rapidly acquire global quantitative measurements of the proteome (i.e. the proteins expressed in the cell). We have applied our state-of-art proteomics technologies based on high-resolution separations combined with Fourier transform ion cyclotron resonance mass spectrometry to obtain quantitative and high throughput global proteomic measurements of Shewanella oneidensis, Geobacter sulfurreducens and Rhodobacter sphaeroides.

Accurate ORF identification and functional annotation are important for post-genome analysis of any organism. The Shewanella Federation has organized an effort to update and refine the predictions of RNA and protein-encoding genes (CDS) for Shewanella oneidensis MR-1. The proteome data has proven to be an exceptional resource for validating hypothetical protein predictions, as well as for positioning start codons and signal peptidase cleavage sites. 1197 hypothetical proteins were manually re-evaluated to access the robustness of predictions. As a result of this evaluation, 525 CDS predictions were dropped and 225 were validated. We have also initiated mining of the Shewanella proteome data to validate start codons. Initial analyses in which we used the original predicted start codons validated 801 N-termini. Our results suggest 306 of these termini are a consequence of cleavage of the N-terminal amino acid by methionylaminopeptidase and that the predominant penultimate amino acids found corresponded well with those found in E. coli. Analyses are currently underway to mine proteome data for evidence of 1) pseudogene expression, 2) additional start codon validation based on new start predictions, 3) leader peptidase cleavage sites, and 4) identification of new CDS.

Geobacter sulfurreducens is a representative of an important genus of metal-reducing bacteria that predominate in a variety of subsurface environments in which Fe(III) oxide reduction is important. Changes in protein expression levels were investigated in Geobacter cultured with different terminal electron acceptors. The abundance of proteins in various subcellular fractions of Geobacter sulfurreducens grown on fumarate or Fe(III) citrate were determined and the results compared to identify proteins associated with these distinct modes of anaerobic respiration. Among the proteins that

* Presenting author
changed, 91 c-type cytochromes were identified. Relative abundance of some c-type cytochromes varied markedly with different growth conditions. Higher abundance of cytochromes during growth on Fe(III) may be indicative of cytochromes that play an essential role in Fe(III) reduction. To better understand the physiology of Geobacter species during growth on Fe(III) oxide, the proteome of G. sulfurreducens grown on Fe(III) oxide was compared with the proteome of cells grown with soluble Fe(III) citrate. Analysis using the accurate mass and time tag (AMT) approach revealed many c-type cytochromes that were significantly more abundant in cells grown with insoluble Fe(III) oxide when compared with cells grown on soluble Fe(III) citrate as the electron acceptor. These cytochromes included the outer-membrane c-type cytochrome, OmcS and OmcG, all of which genetic studies have suggested are required for Fe(III) oxide reduction. Furthermore, several other uncharacterized cytochromes were determined to be significantly up regulated during growth on Fe(III) oxide. A number of other proteins of unknown function were also more abundant during growth on Fe(III) oxide than on soluble Fe(III).

Rhodobacter sphaeroides 2.4.1 is capable of growth under a variety of conditions and has been studied in relation to its ability to reduce metals, fix nitrogen, sequester carbon dioxide, and produce energy through photosynthesis. The complex sensory and regulatory network responsible for the transition of R. sphaeroides to a different metabolic steady state suggests the presence of proteins directly and indirectly involved in the photosynthetic lifestyle; beyond the structural and regulatory proteins transcribed and translated from the well-characterized photosynthetic gene cluster (PGC). We present results that characterize the proteome of aerobic and photosynthetic cell cultures by utilizing: 1) proteins extracted from whole cell lysate, soluble, insoluble, and global fractions, and 2) proteins extracted from sub-cellular fractions that include cytoplasm, cytoplasmic membrane, periplasm, outer membrane, and chromatophore. Both analyses utilized the AMT approach. The first analysis emphasized the role of observed proteins in the photosynthetic lifestyle of R. sphaeroides, such as those involved in electron transport, and compared results with available transcriptome data. The second analysis emphasized the localization of proteins within the cellular matrix. Localization for many of the 4269 proteins predicted from R. sphaeroides’ sequenced genome has not been characterized beyond that implied by prediction algorithms and functional annotation. Therefore, it is important to determine the localization of the proteins within the organism to achieve a clear view of the physiology.

Acknowledgements: This research is supported by the Office of Biological and Environmental Research of the U.S. Department of Energy. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute through Contract No. DEAC05-76RLO 1830.
VIMSS Applied Environmental Microbiology Core Research on Stress Response Pathways in Metal-Reducers


1Lawrence Berkeley National Laboratory, Berkeley, CA; 2Oak Ridge National Laboratory, Oak Ridge, TN; 3Diversa, Inc., San Diego, CA; 4University of Missouri, Columbia, MO; 5University of Washington, Seattle, WA; 6Miami University, Oxford, OH; 7Genomatica, Inc., San Diego, CA; and 8Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

Field Studies

Environmental Characterizations. Nitrate and heavy metals (including uranium and technetium) are a major groundwater contaminant at the U.S. Department of Energy (DOE) NABIR-Field Research Center in Oak Ridge, TN. The sites are marked by acidic conditions, high concentrations of nitrate, chlorinated solvents, and heavy metals. Groundwater bacterial communities were monitored in several wells along a transect that were stimulated via the addition of a potential electron donor (i.e., ethanol). Electron donor was added intermittently over 650 days. By day 535, the nitrate levels in the groundwater had decreased from 10 mM to 0.5 mM, and groundwater uranium levels had declined from approximately 2 mg/l to 0.5 mg/l. Bacterial community composition and structure were characterized via clonal libraries of the SSU rRNA gene sequences. The up-stream and injection well had similar diversity indices, whereas the treatment zone and immediately down-stream well both had increased diversity. When the entire sequence libraries were compared via LIBSHUFF analysis. The results indicated that the bacterial community composition and structure changed upon bio-stimulation for metal-reducing conditions, and that sequences indicative of Anaeromyxobacter and Desulfovibrio were detected in wells that displayed a decline in both nitrate and uranium upon bio-stimulation. The results also suggested that, in addition to the presence of desired populations, an increase in diversity may be important for optimal functionality.

13C-labelled lactate was injected in August 2004 at the Hanford 100H site to biostimulate chromium reduction. After more than 1 year chromium was still at non-detect in the stimulated wells. 16s phylochop analyses showed a dramatic increase in diversity at the stimulated wells, including iron reducers (Geobacter) and sulfate reducers (Desulfovibrio). Sequentially competing terminal electron acceptors were depleted: oxygen, nitrate, iron(III), and sulfate. Methane however was never detected, though 13C was detected in the dissolved inorganic carbon and in the signature lipids (PLFA) of iron reducers and sulfate reducers. Sulfate reduction was still active after more then a year in the deepest parts of the aquifer, and iron(II) still dominated suggesting an active Cr(IV) reducing environment. Desulfovibrio strains have been isolated and are currently being sequenced. Stress responses in these strains will be compared to the pipeline studies on DvH already completed.

Biopanning/Clone libraries. Low biomass samples from nitrate and heavy metal contaminated soils yield DNA amounts which have limited use for direct, native analysis and screening. Multiple displacement amplification (MDA) using φ29 DNA polymerase was used to amplify whole genomes from environmental, contaminated, subsurface sediments. By first amplifying the gDNA, biodiversity
analysis and genomic DNA library construction of microbes found in contaminated soils were made possible. Whole genome amplification of metagenomic DNA from very minute microbial sources enables access to genomic information that was not previously accessible. DNA was extracted from nine subsurface soil samples from five different areas within the DOE NABIR FRC site. The samples represent different geographical areas containing various levels of contaminants and varying subsurface depths. Multiple displacement amplification was used to amplify whole genomes from the extracted DNA. By first amplifying the gDNA, biodiversity analysis and genomic DNA library construction of microbes found in these contaminated soils were made possible. After amplification, SSU rRNA analysis revealed relatively even distribution of species across several major phyla. Clone libraries were constructed from the amplified gDNA and a small subset of clones was used for shotgun sequencing. BLAST analysis of the library clone sequences, and COG analysis, showed that the libraries were diverse and the majority of sequences had sequence identity to known proteins. The libraries were screened by DNA hybridization and sequence analysis for native histidine kinase genes. 37 clones were discovered that contained partial histidine kinase genes, and also partial, associated response regulators and flanking genes.

**Enrichments.** This year we also isolated and characterized a new strain of *D. vulgaris* that should be valuable as a genetic tool and to investigate the role of bacteriophage in microbial stress responses. *D. vulgaris* DePue was isolated from sediments of the metal (Zn) contaminated Lake DePue. Our analysis revealed that this strain did not possess phage genes that were found in the *D. vulgaris* Hildenborough genome and was susceptible to two phages from the former strain. Currently, genome sequencing of strain DePue is underway at the JGI.

**Dual culture systems.** Although sulfate-reducing bacteria (SRB) characteristically respire sulfate, their distribution does not appear to be restricted by sulfate availability. In the absence of sulfate, some SRBs can grow by cooperating syntrophically with hydrogenotrophic methanogens. We established and characterized a syntrophic coculture between the model SRP *Desulfovibrio vulgaris* Hildenborough and *Methanococcus maripaludis* in order to study the physiology and influence of stressors on the growth of *D. vulgaris* in the absence of sulfate. In this interaction, the species must cooperate to transform lactate and carbon dioxide into acetate and methane by transferring reducing equivalents. Interspecies hydrogen transfer is known as a driving process for such interactions, but transfer of formate or other compounds might also occur. Using publicly available genomic and physiological data, we developed a stoichiometric metabolic model that predicts that both hydrogen and formate could be used as electron carriers. This model was able to predict metabolite accumulation of *D. vulgaris* in monoculture and in the syntrophic coculture at multiple stages in their dynamic growth cycle, but there were discrepancies between the experimental data and model predictions at individual time points. These discrepancies may be explained by dynamic oscillations of growth of each species in coculture, a phenomenon which is supported by a dramatically changing ratio of 16S rRNA production of each species over time. We have also compared gene expression of each species under coculture and monoculture conditions. Although experiments with mutants show that it is possible for coculture growth to occur when the capacity to transfer electrons via formate has been deleted, initial gene expression data and the metabolic model both suggest that formate may be used as an electron carrier under syntrophic conditions. Further comparisons of mutant cocultures will be performed to address this issue. Gene expression analyses of co-cultures has revealed several interesting gene expression changes in *D. vulgaris*. Even though there was no sulfate available, *D. vulgaris* genes for the sulfate reduction pathway were expressed. Genes for conversion of lactate to acetate and acetate excretion were downregulated during syntrophic growth compared to mono-culture growth in the presence of sulfate even though lactate was the sole carbon source in each condition. Finally, differential expression of several hydrogenases thought to be important for *D. vulgaris* growth under the different conditions were detected.
Stress Experiments

High Throughput Biomass Production. Producing large quantities of high quality and defensibly reproducible cells that have been exposed to specific environmental stressors is critical to high throughput and concomitant analyses using transcriptomics, proteomics, metabolomics, and lipidomics. Culture of *D. vulgaris* is made even more difficult because it is an obligate anaerobe and sulfate reducer. For the past three years, our Genomics:GTL VIMSS project has developed defined media, stock culture handling, scale-up protocols, bioreactors, and cell harvesting protocols to maximize throughput for simultaneous sampling for lipidomics, transcriptomics, proteomics, and metabolomics. All cells for every experiment, for every analysis are within two subcultures of the original ATCC culture of *D. vulgaris*. In the past three years we have produced biomass for 80 (40 in the last year) integrated experiments (oxygen, NaCl, NO₃, NO₂, heat shock, cold shock, pH, Cr, and mutants Fur and Per) each with as much as 30 liters of mid-log phase cells (3 x 10⁸ cells/ml). In addition, more than 60 adhoc experiments for supportive studies have been done each with 1-6 liters of culture. All cultures, all media components, all protocols, all analyses, all instruments, and all shipping records are completely documented using QA/QC level 1 for every experiment and made available to all investigators on the VIMSS Biofiles database (http://vimss.lbl.gov/perl/biofiles). To determine the optimal growth conditions and determine the minimum inhibitory concentration (MIC) of different stressors we adapted plate reader technology using Biolog and Omnilog readers using anaerobic bags and sealed plates. Since each well of the 96-well plate produces an automated growth curve, over more than 200 h, this has enabled us to do more than 6,000 growth curves over the last two years. Since the Omnilog can monitor 50 plates at a time, this allows us to do more than 5,000 growth curves in a year.

Phenotypic Responses. Phenotypic Microarray™ analysis is a recently developed analytical tool to determine the phenotype of an organism. In the last year we have further refined our phenotyping of DvH to minimize the number of plates necessary. We have also screened 15 knockout mutants of DvH and 10 knockout mutants of *Shewanella* MR1. See (https://vimss.lbl.gov/~jsjacobsen/cgi-bin/Test/HazenLab/Omnilog/home.cgi) for sample data sets and analyses.

Synchrotron FTIR Spectromicroscopy for Real-Time Stress Analysis. This year the stress responses in *Desulfovibrio vulgaris* triggered by oxygen (O₂), nitrate (NO₃), Cr, and sodium chloride (NaCl) were studied using FTIR. The advantage of the FTIR spectroscopy approach is that it allows us to immediately detect *in situ* intracellular molecules or molecular structures, to nondestructively monitor and quantify metabolites produced in response to different stresses, and to rapidly characterize growth-dependence phenomena and stress-response mechanisms. Because the chemical and structural information of molecules associated with cellular processes inside microbes are contained in each infrared spectrum, one can extract chemical and structural information from each spectrum regarding the physiological conditions of a cell or a group of cells. By comparing measurements, we were able to identify tight temporal changes in chemical bonds, functional groups, and chemical substructures in lipids, DNA, proteins, and polyglucose in *D. vulgaris*. For example, when exposed to moderate concentrations of O₂ or NO₃, *D. vulgaris* increases the production of exopolysaccharides but with little change in protein structures. However, when exposed to moderate concentration of NaCl, *D. vulgaris* again increases the production of exopolysaccharides while exhibiting a significant change in protein structures. These results, together with microscopy images, confirmed the importance of exopolysaccharide production in enhancing the stress resistance and survival of *D. vulgaris*. These studies also enabled focusing of VIMSS transcriptomic, proteomic, and metabolomic studies on the best time points to rapidly resolve stress response pathways.
The Virtual Institute of Microbial Stress and Survival (VIMSS): Deduction of Stress Response Pathways in Metal/Radionuclide Reducing Microbes

Carl Abulencia¹,⁸, Eric Alm¹,⁸, Gary Andersen¹, Adam P. Arkin¹,⁸* (APArkin@lbl.gov), Kelly Bender⁵,⁸, Sharon Borglin¹,⁸, Eoin Brodie¹, Romy Chakraborty¹,⁸, Swapnil Chhabra³,⁸, Steve van Dien⁶, Inna Dubchak¹,³, Matthew Fields⁷,⁸, Sara Gaucher¹,⁸, Jil Geller¹,⁸, Masood Hadi¹,⁸, Terry W. Hazen¹,⁸, Qiang He², Zhili He²,⁸, Hoi-Ying Holman¹,⁸, Katherine Huang¹,⁸, Rick Huang²,⁸, Janet Jacobsen¹,⁸, Dominique Joyner¹,³, Jay Keasling¹,³, Keith Keller¹,³, Martin Keller¹,³, Aindrila Mukhopadhyay¹,³, Richard Phan¹,³, Morgan Price¹,³, Joseph A. Ringbauer Jr.⁵,⁸, Anup Singh³,⁸, David Stahl⁶,⁸, Sergey Stolyar³,⁸, Jun Sun⁴, Dorothea Thompson²,⁸, Christopher Walker⁶,⁸, Judy Wall⁵,⁸, Jing Wei⁴, Denise Wolf⁶,⁸, Denise Wyborski³,⁸, Huei-che Yen³,⁸, Grant Zane⁵,⁸, Jizhong Zhou²,⁸, and Beto Zuniga⁶
¹Lawrence Berkeley National Laboratory, Berkeley, CA; ²Oak Ridge National Laboratory, Oak Ridge, TN; ³Sandia National Laboratories, Livermore, CA; ⁴Diversa, Inc., San Diego, CA; ⁵University of Missouri, Columbia, MO; ⁶University of Washington, Seattle, WA; ⁷Miami University, Oxford, OH; and ⁸Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

Introduction

The mission of the Virtual Institute of Microbial Stress and Survival is to understand the molecular basis for the survival and growth of microbes in the environment. Towards this end VIMSS has designed a series of key protocols, experimental pipelines and computational analyses to support and coordinate research in this area. Our flagship project aims to elucidate the pathways and community interactions which underlie the ability of Desulfovibrio vulgaris Hildenborough (DvH) to survive in diverse, possibly contaminated environments and reduce metals. Their ability to reduce toxic Uranium and Chromium, major contaminants of industrial and DOE waste sites, to a less soluble form has made them attractive from the perspective of bioremediation.

We are discovering the molecular basis for the physiology of these organisms first through characterization of the biogeochemical environment in which these microbes live and how different features of these environments affect their growth and reductive potential. We have created an integrated program through the creation of an experimental pipeline for the physiological and functional genomic characterization of microbes under diverse perturbations. This pipeline produced controlled biomass for a plethora of analyses as described below and is managed through workflow tools and a data management and analysis system. The effort is broken into three interacting core activities: The Applied Environmental Microbiology Core; the Functional Genomics Core; and the Computational Core. While the individual accomplishments of these cores may be found in more detail, we summarize the highlights here.

Accomplishments of the Applied Environmental Microbiology Core (AEMC)

Characterization of the Environment. The AEMC has been monitoring natural and stimulated groundwater and soil communities in DOE NABIR Field Research sites. These sites are contaminated with different combinations of heavy metals, low pH, chlorinated solvents, nitrates and other cellular stressors. Strong correlations between the population growth of certain strains of bacteria including DvH like strains and metal reduction dynamics were observed. D. vulgaris strains from multiple sites have been enriched and isolated and some are undergoing sequencing now. A novel cloning method was used to amplify DNA and whole genomes from five separate contaminated sites
to compare and contrast functional diversity among these sites and relate this diversity to the stability and metal-reduction efficacy of the communities at these sites (including *D. vulgaris*). An initial survey of the signal transduction genes is underway.

**Biomass production and Characterization:** In the core pipeline experiments each microbe is first characterized physiologically using Omnilog phenotypic microarrays. A stressor condition is then applied to a large set of batch cultures and samples are collected periodically to obtain a time-series of cellular response. Each time-point is split so that the cells can be imaged, analyzed through synchrotron IR microscopy to measure the bulk physiological changes of the cells during their response, and determine the optimal time points to send to the functional genomics core (FGC) for transcript, protein and metabolite analysis. Twenty-three conditions in DvH and *Shewanella oneidensis* (So) have been produced including mutant studies and co-culture of DvH syntrophically coupled to a hydrogenotrophic methanogen (*Methanococcus maripaludis*). Using publicly available genomic and physiological data, we developed a stoichiometric metabolic model that predicts that both hydrogen and formate could be used as electron carriers. This model was able to predict metabolite accumulation of *D. vulgaris* in monoculture and in the syntrophic coculture at multiple stages in their dynamic growth cycle, but there were discrepancies between the experimental data and model predictions at individual time points which can be explained by interesting observed growth dynamics.

**Accomplishments of the Functional Genomics Core**

**Genetics:** Our bar-coded deletion project proceeds apace and a number of deletions have been phenotyped by omnilog array and through the VIMSS physiological pipeline. New affinity tags have been developed for using in pull-down and molecular complex studies both for this and a collaborating project. We have also expanded our transposon mutagenesis library and have begun to array them for sequencing and phenotyping. Ultimately the most interesting of these will be submitted to pipeline studies.

**Transcriptomics:** We have, to date, characterized seventeen stresses, growth phase conditions, or mutant responses in DvH and six in *S. oneidensis* and results are integrated with the VIMSS MicrobesOnline Database. New regulons and their cis-regulatory sequences have been discovered along with new hypotheses of the pathways by which both organisms respond to these different stressors. A number of papers are in press, submitted or are in preparation around this topic. Compendium analysis for DvH is underway. *Geobacter* stressors are coming online and will be used for the three organism cross comparison.

**Proteomics.** As part of the proteomics mission we have developed and compared and contrasted MS/MS, ICAT MS, ITRAQ MS, and DIGE (Differential In-gel Electrophoresis) MALDI to profile the protein abundance and protein abundance changes in response to stressors. The year has been spent in quantifying reproducibility and accuracy of these results and comparing their predictions to that of microarray. There were significant differences among the methods and with the microarray data, however, these highlight the different sensitivities of the instruments and the long cell-cycle times of the microorganism. Recent, longer term experiments show better agreement between the microarray and proteomic data. We are now targeting our proteomics efforts to track the changes in the key sulfate reducer signature genes (see below) and their linked stress response pathway regulators to generate a more limited but precise measure of both abundance and redox changes.

**Metabolomics:** We have set up and optimized both Capillary electrophoresis (CE) and Liquid chromatography (LC) coupled with Mass spectrometry (MS) methods for characterization of metabolites. Metabolite extraction protocols have been developed for DvH. A new Fourier transform ion cyclotron resonance mass spectrometer has recently come online which will allow a much wider survey of
metabolites without the need for external standards. However, we are still in testing phase of this technology. We have identified the key pathways and metabolites on which we will focus to understand the sulfate-reduction and stress metabolism of DvH in contrast with S. oneidensis and Geobacter.

Accomplishments of the Computational Core

The computational core has continued its core development of MicrobesOnline (web site: http://MicrobesOnline.org) as the core framework for comparative functional microbial genomics and for the dissemination and visualization of VIMSS data. The microarray database is now fully integrated with the website and proteomic, metabolomic and molecular interaction data will shortly follow. A highly curated database of cis-regulatory motifs and their regulators in a variety of organisms from Mikhail Gelfand is undergoing integration with MicrobesOnline and a suite of cis-regulatory prediction tools are being added to the informatics tools available on the site. The team has used the site to provide nearly automated data analysis for the microarray experiments above and has been used to understand the key biological mechanisms in a number of responses including cold and heat shock, low and high pH, low and high oxygen, salt and osmotic shock and adaptation, nitrate and nitrite stress, chromium reducing conditions, growth in co-culture with a methanogen, iron limitation and mutation in the fur gene. A new method called OpWise has aided in this analysis and uses prior computational core work in accurate operon prediction to exploit genomic architecture to get better estimates of gene expression changes and the systematic and non-systematic errors underlying the measurements. We are now in the process of integrating the new proteomic and metabolomic data into the framework. We already have a new interface for analysis and visualization for the phenotype microarray data which we are now using for quality control and analysis of mutant data.

Along with the core platform development and data analysis and interpretation, the computational core has used MicrobesOnline as a platform for more general discovery and annotation. MicrobesOnline is being used as an annotation tool for a number of new genomes and environmental sequences. For example, in collaboration with another project we have identified and annotated a novel sulfate reducing organisms isolated from a deep South African Goldmine. In addition, a set of signature genes that define sulfate reducing bacteria and Archea were inferred and validated against the gene expression database. A set of DvH centered comparative pathway analyses for the Metabolic, global regulatory, dissimilatory nitrogen oxide pathways were accomplished with Mikhail Gelfand’s group. New theories for the formation, maintenance, tuning, and death of operon structure were put forward and a comprehensive study on the evolution of histidine kinases was completed. This latter study also discovered that certain organisms are more likely to generate new kinases by lineage specific expansion and other to acquire them through horizontal gene transfer.

Future Work

In this next year we are focusing on integrating our stress condition measurements and models to the ability of the organism to reduce metals. We will relate the resultant model to the ability of the organism to survive and reduce metals in the various environments that the AEMC has been monitoring. Indeed, we will use water collected from some of these sites as stressors in the lab to see if we can relate our prior results to this more “natural” perturbation. We are focusing efforts on distinguishing between competing theories of energy generation in this organism and on tracking electron flow in some of the inferred pathways. We will also extend our studies on the interaction of DvH with its syntrophic methanogens and derive a better view with how it operates within its community. Based on these studies will attempt to direct field studies to track the role of Dv-like species in affecting biogeochemical changes in the environment. We will also study further the evolution of regulation of key pathways in DvH by comparing and contrasting genomic data across the bacterial kingdom and functional genomic data from Shewanella, Geobacter and a number of other microbes. Ultimately, we
will combine the discovered regulatory pathways and metabolic models of DvH into an integrated model of DvH physiology and begin to understand the evolutionary origins of this network by comparison to other environmental microbes.

### 61

**Nitrate Stress Response in *Desulfovibrio vulgaris* Hildenborough: Whole-Genome Transcriptomics and Proteomics Analyses**

Qiang He¹, Zhili He¹,²,⁶, Wenqiong Chen¹, Zamin Yang¹, Eric J. Alm⁴,⁶, Katherine H. Huang⁴,⁶, Huei-Che Yen⁵,⁶, Dominique C. Joyner⁴,⁶, Martin Keller²,⁵, Jay Keasling⁴,⁶, **Adam P. Arkin**⁴,⁶, Terry C. Hazen⁴,⁶, Judy D. Wall⁵,⁶, and Jizhong Zhou¹,²,⁶* (jzhou@ou.edu)

¹Oak Ridge National Laboratory, Oak Ridge, TN; ²University of Oklahoma, Norman, OK; ³Diversa Corporation, San Diego, CA; ⁴Lawrence Berkeley National Laboratory, Berkeley, CA; ⁵University of Missouri, Columbia, MO; and ⁶Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

Sulfate reducing bacteria (SRB) are of interest for bioremediation with their ability to reduce and immobilize heavy metals. Nitrate, a common co-contaminant in DOE sites, is suggested to inhibit SRB via nitrite. Previous results indicate that nitrite is indeed inhibitory to the growth of *Desulfo-vibrio vulgaris*. However, growth inhibition by nitrate alone was also observed. In this study, growth and expression responses to various concentrations of nitrate were investigated using the Omnilog phenotype arrays and whole-genome DNA microarrays. Changes in the proteome were examined with 3D-LC followed by MS-MS analysis. Microarray analysis found 5, 50, 115, and 149 genes significantly up-regulated and 36, 113, 205, and 149 down-regulated at 30, 60, 120, and 240 min, respectively. Many of these genes (~50% at certain time points) were of unknown functions. By comparison to NaCl stress, transcriptional analysis identified changes specific to NaNO₃ stress. The hybrid cluster protein was among the highly up-regulated genes, suggesting its role in nitrate stress resistance with its proposed function in nitrogen metabolism. The up-regulation of phage shock protein genes (*pspA* and *pspC*) might indicate a reduced proton motive force and the repression of multiple ribosomal protein genes could further explain the growth cessation resulting from nitrate stress. A glycine/betaine transporter gene was also up-regulated, suggesting that NaNO₃ also constituted osmotic stress. Osmoprotectant accumulation as the major resistance mechanism was validated by the partial relief of growth inhibition by glycine betaine. Proteomics analyses further confirmed the altered expression of these genes, and in addition, detected increased levels of several enzymes (Sat, DvsB, and AprB) in the sulfate reduction pathway, indicative of the increased energy production during nitrate stress. In conclusion, excess NaNO₃ resulted in both osmotic stress and nitrate stress. *D. vulgaris* shifted nitrogen metabolism and energy production in response to nitrate stress. Resistance to osmotic stress was achieved primarily by the transport of osmoprotectant.

Many of the proteins that are candidates for bioenergetic pathways involved with sulfate respiration in *Desulfovibrio* spp. have been studied, but complete pathways and overall cell physiology remain to be resolved for many environmentally relevant conditions. In order to understand the metabolism of these microorganisms under adverse environmental conditions for improved bioremediation efforts, *Desulfo-vibrio vulgaris* Hildenborough was also used as a model organism to study stress response to nitrite, an important intermediate in the nitrogen cycle. Previous physiological studies demonstrated that growth was inhibited by nitrite and that nitrite reduction was observed to be the primary mechanism of detoxification. Global transcriptional profiling with whole-genome microarrays revealed a coordinated cascade of responses to nitrite in pathways of energy metabolism, nitrogen metabolism, oxidative stress...
response, and iron homeostasis. In agreement with previous observations, nitrite stressed cells showed a decrease in expression of genes encoding sulfate reduction functions in addition to respiratory oxidative phosphorylation and ATPase activity. Consequently, the stressed cells had decreased expression of ATP-dependent amino acid transporters and proteins involved in translation. Nitrite detoxification also appeared to shift the flow of reducing equivalents from oxidative phosphorylation to nitrite reduction. Increased demand for iron, resulting from these regulatory events and the chemical oxidation of available Fe$^{2+}$, likely contributed to iron depletion and the derepression of the Fur regulon.

62

**Bacterial Nanowires: Novel Electron Transport Machines that Facilitate Extracellular Electron Transfer**

Yuri A. Gorby\(^*\) (yuri.gorby@pnl.gov), Svetlana Yanina\(^1\), Dianne Moyles\(^2\), Matthew J. Marshall\(^1\), Jeffrey S. McLean\(^1\), Alice Dohnalkova\(^1\), Kevin M. Rosso\(^1\), Anton Korenevski\(^1\), Terry J. Beveridge\(^1\), Alex S. Beliaev\(^1\), In Seop Chang\(^4\), Byung Hong Kim\(^4\), Kyung Shik Kim\(^4\), David E. Culley\(^1\), Samantha B. Reed\(^1\), Margaret F. Romine\(^1\), Daad A. Saffarini\(^3\), Liang Shi\(^1\), Dwayne A. Elias\(^1\), David W. Kennedy\(^1\), Grigoriy E. Pinchuk\(^1\), Eric A. Hill\(^1\), John M. Zachara\(^1\), Kenneth H. Nealson\(^3\), and Jim K. Fredrickson\(^1\)

\(^1\)Pacific Northwest National Laboratory, Richland, WA; \(^2\)University of Guelph, Guelph, Ontario; \(^3\)University of Wisconsin, Milwaukee, WI; \(^4\)Korea Institute of Science and Technology, Seoul, Korea; and \(^5\)University of Southern California, Los Angeles, CA

Coordinated transfer of electrons from electron donors to electron acceptors provides means for harvesting and transforming electrochemical potential energy into forms that sustain life. Electron transport components are truly 'global regulators' of cellular processes in the sense that energy flow influences all aspects of activity, response, and regulation and, as such, warrant scientific attention. Identifying electron transport components, characterizing the mechanisms of electron transduction, and relating electron transfer to bioenergetics are necessary to advance a systems level understanding of microorganisms. Dissimilatory metal reducing bacteria, such as *Shewanella oneidensis* strain MR-1 and *Geobacter sulfurreducens*, have enjoyed such scientific scrutiny since their discovery in the mid 1980's. Beyond obvious potential for these organisms to catalyze biogeochemical processes in natural environments, much of the research was driven toward developing a fundamental understanding of the processes that facilitate and limit electron transfer from bacteria to solid phases.

Although the controlling mechanisms of electron transfer remain poorly understood, electron acceptor availability under metal reducing conditions is typically a growth-limiting condition. Dissimilatory metal reducing bacteria produce electrically conductive appendages, which we call bacterial nanowires, in direct response to electron acceptor limitation. Nanowires produced by *S. oneidensis* strain MR-1, which served as our primary model organism, are functionalized by decaheme cytochromes MtrC and OmcA that are distributed along the length of the nanowires. Mutants deficient in MtrC and OmcA produce nanowires that were poorly conductive as determined by Scanning Tunneling Microscopy (STM). These mutants also differed from the wild type in their inability to reduce solid phase iron oxides, poor power production in a mediator-less microbial fuel cell, and failure to form complex biofilms at air-liquid interfaces.

Preliminary observations suggest that nanowires are also produced by other bacteria, including the oxygenic, phototrophic cyanobacterium *Synechocystis* PCC6803. This organism can produce highly conductive nanowires in response to excess light, which serves as the energy source to split water into
protons and electrons, and limited CO₂, which otherwise serves as an electron sink during biomass production. Although additional work is needed to characterize the components of nanowires in other organisms, these results demonstrate that electrically conductive nanowires are not restricted to any single genus or even to particular metabolic guilds, such as dissimilatory metal reducing bacteria. Indeed, we hypothesize that nanowires are distributed throughout the bacterial world where they and serve as structures for efficient electron transfer and energy distribution. The electron carriers associated with nanowires in different metabolic guilds is likely to vary significantly due to differences in the redox couples utilized by the various organisms. High throughput methods such as those being proposed for new DOE biology user facilities would greatly facilitate the identification and analysis of these extracellular molecular machines. Further collaborative investigation into the complete composition of nanowires, mechanisms of electron flow through the wires, and interaction of nanowires between and among organisms in natural microbial communities is warranted in order to completely realize the implications of these structures in areas of alternative energy, carbon sequestration, bioremediation, and possibly pathogenicity and human health.

VIMSS Functional Genomics Core Research on Stress Response Pathways in Metal-Reducers


In collaboration with the Applied Environmental and Microbiology Core (AEMC), the Functional Genomics Core (FGC) has now fully optimized the generation of biomass for parallel experiments using the various genomics techniques. Outlined below are the most significant accomplishments from the different units of the FGC.

A. Microarray Analysis

* Desulfovibrio vulgaris* Hildenborough was used as a model organism to study a wide variety of stresses that are environmentally important. They include hyper-salinity, extreme pH conditions, Chromate, Nitrite, Nitrate and cold shock. Additionally, stationary phase physiology was also examined. Microarray analysis has now also been initiated for characterization of *D. vulgaris* mutants and the ∆fur mutant has been compared with wild type. Each transcriptome analysis has led to a vast repository of information regarding each stress. Several comparative stress response analyses are being conducted collaboratively between the FGC and the Computational Core (CC). Additionally, transcriptomics profiling has also been completed for several stresses in *S. oneidensis*. Initial studies...
have also been started for *G. metallireducens*. This sets the stage for important comparative studies that compare similar stress responses across different organisms.

**B. Proteomics**

*a. Peptide tagging Quantitative proteomics using ICAT and ITRAQ strategies*

Methods were optimized for extensive separation of peptide pools so as to enable detection of a large number of peptides and corresponding proteins. Using a strong cation exchange coupled with reverse phase and tandem high resolution mass spectroscopy greater than 800 proteins were identified and their relative amounts quantified. The highly reproducible internal replicates allow stringent statistical analysis to be conducted on these data sets. This method had been applied to both oxygen stress and nitrate stress in *D. vulgaris*.

*b. Comprehensive proteomics using 3D-LC/MS/MS*

Using the powerful 3D separation technique and the high throughput mass spectroscopy at Diversa, Corp., a large number of proteins (40-60% of the proteome) were identified for several stress response comparisons. In this method spectral counting was used to estimate relative quantities of proteins. With proteomics data from several stress responses it now becomes possible to conduct a comparative proteomics analysis.

*c. Study of Protein Complexes and Protein-Protein Interaction*

Since proteins generally function within the cell through strong and weak interactions with other protein partners, it was considered necessary to characterize these complexes to increase our understanding of the functional proteomics picture. Two contrasting approaches for the isolation of protein complexes from *D. vulgaris* have been implemented. The endogenous approach involves *D. vulgaris* mutants containing an engineered tag that can be captured by affinity chromatography. Using lysates from these cells the tagged protein, in complex with its associated proteins, are captured and selectively eluted. In the exogenous approach, heterologously expressed His-tagged bait proteins from *D. vulgaris* are purified and coupled to affinity beads which are then incubated with *D. vulgaris* lysate to capture interacting proteins. Components of well characterized homologous protein complexes in *E. coli*, e.g. rpoB and rpoC have been used to validate methods. Several *D. vulgaris* proteins which have already been used in these strategies include Dnak, ClpX and CooX. Several ORFsgenes in Sulfate reducing bacteria are now being developed for similar studies.

**C. Metabolomics**

Fully developed methods have now been optimized by the FGC to study a vast majority of commonly encountered metabolites, which for methodological purposes are broadly broken down into molecules that are ideally resolved in positive ion mode or in negative ion mode. Given the scarcity of established methods in this area, our optimized strategies will be published as significant advancement in the field of metabolomics. A survey of metabolic extracts from *D. vulgaris* is being conducted as proof of concept. Noteworthy advance has also been made in the use of the Fourier Transform ion cyclotron resonance (FTICR) to detect and characterize metabolites based almost entirely on their exact mass. FTICR coupled Mass spectroscopy will enable the identification of hundreds of metabolites. In addition, fragmentation information can be used to characterize unknown peaks.

**D. Integrating Genomics Data**

Work from two separate stress response studies; namely Salt stress and Heat shock, were used in authoring integrated genomics manuscripts. The salt stress study was further complemented by phospholipid fatty acid analysis and several osmoprotection assays that led to a comprehensive model

* Presenting author
for salt stress in *D. vulgaris*. Following the model of rigorous checks for consistency and accuracy for all microarray data, an important collaborative project has been started between the FGC and CC to create a similar set of computational tools to analyze the Proteomics data. Data from the Oxygen stress and Nitrate stress are being used for this.

**64**

**Comparative Analysis of Bacterial Gene Expression in Response to Environmental Stress**

Eric J. Alm1,2*, Edward E. Baidoo1,2, Peter I. Benke1,2, Sharon C. Borglin1,2, Wenqiong Chen1,3, Swapnil Chhabra1,4, Matthew W. Fields1,9, Sara P. Gaucher1,4, Alex Gilman1,2, Masood Hadi1,4, Terry C. Hazen1,2, Qiang He1,5, Hoi-Ying Holman1,2, Katherine Huang1,2, Rick Huang1,2, Zhili He1,5, Dominique C. Joyner1,2, Jay D. Keasling1,7, Martin Keller1,3, Keith Keller1,2, Aindrila Mukhopadhyay1,2, Paul Oeller1,2, Francesco Pingitore1,2, Alyssa Redding1,7, Anup Singh1,4, David Stahl1,8, Sergey Stolyar1,8, Jun Sun1,2, Zamin Yang1,5, Judy Wall1,6, Grant Zane1,6, Jizhong Zhou1,5, and Adam P. Arkin1,2,7

1Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov; 2Lawrence Berkeley National Laboratory, Berkeley, CA; 3Diversa, Inc., San Diego, CA; 4Sandia National Laboratories, Livermore, CA; 5Oak Ridge National Laboratory, Oak Ridge, TN; 6University of Missouri, Columbia, MO; 7University of California, Berkeley, CA; 8University of Washington, Seattle, WA; and 9Miami University, Oxford, OH

The transcriptional response of bacterial species to environmental stress has been the subject of considerable research, fueled in part by the widespread availability of gene expression microarray technology. Previous studies have established the similarity of gene expression networks across a wide range of organisms, yet in these studies different experiments were performed on different species preventing a direct comparison. We have compiled a core set of ‘standard’ stressors including salt, pH, temperature, and oxygen and nitrite/nitrate levels and applied these stressors systematically to a phylogenetically diverse group of metal-reducing bacteria. We compare the expression patterns of orthologous genes and regulons in *Desulfovibrio vulgaris*, *Geobacter metallireducens*, and *Shewanella oneidensis* after exposure to these stressors. We observe that while the overall network may be conserved (genes in the same pathways have high correlations over all conditions), the response of the network to the same perturbations can be very different in different species (pathways may respond to the same stressor in different ways). Differences between species can arise from differential behavior of the same regulons and because ‘orthologous’ regulons may comprise different sets of (non-orthologous) genes, both of which may lead to insights in the ecological factors that shape gene expression.
Evaluation of Stress Responses in Sulfate-Reducing Bacteria Through Genome Analysis: Identification of Universal Responses

J.D. Wall1,7 (wallj@missouri.edu), H.-C. Yen1, E.C. Drury1,7, A. Mukhopadhyay2,7, S. Chhabra3,7, Q. He4, M.W. Fields5,7, A. Singh1, J. Zhou4,7, T.C. Hazen2,7, and A.P. Arkin2,6,7

1University of Missouri, Columbia, MO; 2Lawrence Berkeley National Laboratory, Berkeley, CA; 3Sandia National Laboratories, Livermore, CA; 4University of Oklahoma, Norman, OK; 5Miami University, Oxford, OH; 6Howard Hughes Medical Institute, Berkeley, CA; and 7Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

The application of the toxic metal metabolism of the anaerobic sulfate-reducing bacteria to bioremediation of contaminated environments requires a broad understanding of the effects of environmental stresses on the organism. The model bacterium, Desulfovibrio vulgaris Hildenborough, for which the genome sequence has been fully determined, is being examined for its responses to a variety of stresses that may be expected to be encountered in natural/contaminated settings. We have examined the preliminary transcriptional data from ten treatments to learn whether there are general responses or common themes for responses to stresses by D. vulgaris. This anaerobe apparently does not have an ortholog encoding RpoS implicated in the universal stress response in γ-Proteobacteria. Interestingly genes predicted to be controlled by the global regulator Fur appear to be among the most frequently responsive in the genome. The transcriptional responses to increased concentrations of sodium and potassium overlapped strongly, as would be predicted. Curiously, it was not predicted that these salt responses would be shared by the response to reduced temperature. Also counter to our prediction, the response to nitrate was not a simple sum of the responses to sodium and nitrite. Further insights into general patterns of transcription during stresses will be discussed.

Cellular Responses to Changing Conditions in Desulfovibrio vulgaris and Shewanella oneidensis

M.E. Clark1, Q. He2, Z. He2,7, K.H. Huang3,7, E.J. Alm3,7, X. Wan1, T.C. Hazen3,7, A.P. Arkin3,4,5,7, J.D. Wall6,7, J. Zhou7, J. Kurowski1, A. Sundararajan1, A. Klonowska1, D. Klingeman2, T. Yan2, M. Duley1, and M. W. Fields1,7 (fieldsmw@muohio.edu)

1Miami University, Oxford, OH; 2Oak Ridge National Laboratory, Oak Ridge, TN; 3Lawrence Berkeley National Laboratory, Berkeley, CA; 4University of California, Berkeley, CA; 5Howard Hughes Medical Institute, Berkeley, CA; 6University of Missouri, Columbia, MO; and 7Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

Temporal Transcriptomic Analysis of Desulfovibrio vulgaris Hildenborough Transition into Stationary-Phase Growth during Electron Donor Depletion

Desulfovibrio vulgaris was cultivated in a defined medium and biomass was sampled over time for approximately 70 h to characterize the shifts in gene expression as cells transitioned from exponential to stationary phase growth during electron donor depletion. In the context of in situ bioremediation, nutrient scarcity and/or depletion may be a common obstacle encountered by
microorganisms due to the oligotrophic nature of most groundwater and sediment environments. In addition to temporal transcriptomics; protein, carbohydrate, lactate, acetate, and sulfate levels were measured. The microarray data was used for statistical expression analyses, hierarchical cluster analysis, and promoter element prediction. As the cells transitioned from exponential to stationary-phase growth a majority of the down-expressed genes were involved in translation and transcription, and this trend continued in the remaining time points. Intracellular trafficking and secretion, ion transport, and coenzyme metabolism showed more up-expression compared to down-expression as the cells entered stationary-phase. Interestingly, most phage-related genes were up-expressed at the onset of stationary-phase. This result suggested that nutrient depletion may signal lysogenic phage to become lytic, and may impact community dynamics and DNA transfer mechanisms of sulfate-reducing bacteria. The putative foAAB system (in addition to other putative iron-related genes) was significantly up-expressed, and suggested the possible importance of Fe\(^{2+}\) acquisition under reducing growth conditions for sulfate-reducing bacteria. A large subset of carbohydrate-related genes had altered gene expression, and the total carbohydrate levels declined during the growth phase transition. Interestingly, the *D. vulgaris* genome does not contain a putative rpoS gene, a common attribute of the δ-Proteobacteria genomes sequenced to date, and other putative rpo factors did not have significantly altered expression profiles. The elucidation of growth-phase dependent gene expression is essential for a general understanding of growth physiology that is also crucial for data interpretation of stress-responsive genes. In addition, to effectively immobilize heavy metals and radionuclides via sulfate-reduction, it is important to understand the cellular responses to adverse factors observed at contaminated subsurface environments, such as the changing ratios of electron donors and acceptors. Our results indicated that genes related to phage, internal carbon flow, outer envelop, and iron homeostasis played important roles as the cells experienced electron donor depletion.

**Deletion of a Multi-Domain PAS Protein Causes Pleiotropic Effects in *Shewanella oneidensis* MR-1**

*Shewanella oneidensis* MR-1, a Gram-negative facultative anaerobe, can utilize a wide array of alternative electron acceptors during anaerobic respiration, and the ability to reduce soluble forms of heavy metals to insoluble forms makes it a potential candidate for bioremediation studies. Understanding the physiological responses of *S. oneidensis* to environmental stresses (e.g., nutrients, oxygen) is important for the assessment of potential impacts on metal-reducing activity. Here we describe the physiological role of a presumptive signal transduction protein in *Shewanella oneidensis* MR-1. The predicted ORF (SO3389) encoded a GGDEF, EAL, and two PAS domains. The deduced amino acid sequence was not closely related to previously described proteins, but presumptive proteins with similar domain architectures were observed in metabolically diverse microorganisms. An in-frame, deletion mutant was constructed (ΔSO3389), and the mutant displayed an extended lag period (30 h) when transferred from aerobic to anaerobic medium. The mutant was also defective in motility, cytochrome content, and was drastically defective in biofilm formation. These pleiotropic phenotypes were observed with multiple growth substrates. During the transition from aerobic to anaerobic conditions, the mutant was deficient in three c-type cytochromes (57, 33, and 20 kDa). In addition, mutant biofilms produced less carbohydrate compared to wild-type cells. Bacterial motility was affected only in aerobic conditions, and this result suggested that SO3389 was involved in O$_2$ responses and was important for anoxic and biofilm growth.
Probing Gene Expression in Single *Shewanella oneidensis* MR-1 Cells

**X. Sunney Xie*** (xie@chemistry.harvard.edu), Paul Choi, Jie Xiao, Ji Yu, Long Cai, Nir Friedman, Jeremy Hearn, Kaiqin Lao, Luying Xun, Joe Zhou, Margaret Romine, and Jim Fredrickson

1Harvard University, Cambridge, MA; 2Applied BioSystems, Foster City, CA; 3Washington State University, Pullman, WA; 4Oak Ridge National Laboratory, Oak Ridge, TN; and 5Pacific Northwest National Laboratory, Richland, WA

Our objective is to make real-time observations of gene expression in single live *Shewanella oneidensis* MR-1 cells with high sensitivity and high throughput. Available technology is only sufficient for the detection of gene expression at high expression levels. However, many important genes are expressed at low levels. New techniques are needed to probe gene expression that produces only a few protein molecules in a single cell, and to follow the expression in real time. Our efforts are summarized as follows:

**Real time imaging of the production of single YFP molecules in a live cell**

We have developed a reporter system for observing real-time production of single protein molecules in individual bacterial cells. A membrane-targeting sequence (tsr) was fused to the gene of fast maturing yellow fluorescent protein (YFP) under the control of a promoter on the chromosome DNA (Figure 1). Gene expression under a repressed condition generates membrane-localized YFP molecules that can be detected one at a time (Figure 2). We found that the protein molecules are produced in bursts and each burst originates from a stochastically transcribed single mRNA molecule. Protein copy numbers in the bursts follow an exponential distribution.

**Real time monitoring of protein production in a live cell using β-galactosidase**

We demonstrate another technique that allows measurements of low level protein expression in individual cells with single molecule sensitivity by taking advantage of the enzymatic properties of β-galactosidase (β-gal), the time-honored reporter for gene expression. β-gal can hydrolyze a wide range of synthetic substrates in addition to lactose, its native substrate. By hydrolyzing fluorogenic substrate a single enzyme molecule can produce a large number of fluorescent product molecules, as was first demonstrated by Rotman in 1961. However, live cell measurements have not been possible because efflux pumps on the cell membrane actively expel foreign organic molecules from the cytoplasm, inhibiting retention of fluorescent products in the cell.

To circumvent the efflux problem, we trap cells in closed microfluidic chambers, such that the fluorescent product expelled from a single cell can accumulate in the small volume of a chamber, recovering the fluorescence signal due to enzymatic amplification.

Figure 1. Detection of single membrane-immobilized YFP molecules that are generated under a repressed condition.
Figure 2. Sequence of fluorescent images of single molecules of fast maturing and membrane-immobilized YFP (yellow) overlaid with bacterial images.

We observed discrete changes of the slopes of fluorescence signal (Figure 3A) indicating that protein production occurs in bursts (Figure 3B), again with the number of molecules per mRNA following an exponential distribution (Figure 3C). We show that the two key parameters of protein expression, the burst size and frequency, can be either determined directly from real time monitoring of protein production or extracted from a measurement of the steady-state copy number distribution in a population of cells with a simple theoretical model.

These studies not only provide new methods for quantification of low-level gene expression, but also yields quantitative understanding of the working of transcription and translation in live cells.

**Real time RT-PCR for sensitive mRNA quantification**

In collaboration with ABI, we use real time RT-PCR to quantitatively measure low mRNA copy numbers inside single bacterial cells. The expression of the reporter gene at the mRNA level can be correlated with that at the protein level.

Compared to the existing methods for characterization of gene expression, such as DNA microarrays and mass spectrometry, the above three methods allow high sensitivity and measurements of gene expression profiling in single live cells. It enables studies of gene expression at the uncharted low expression levels, providing a complete picture of global gene expression.

**Library construction of Shewanella oneidensis MR-1**

We have developed an efficient system for targeted insertions of YFP-gene translational fusions into the *Shewanella oneidensis* MR-1 chromosome using the Gateway technology. Our initial library of over a hundred sequence-confirmed strains includes cytochrome and biofilm-related genes. We are applying the above new techniques to study gene expression of *Shewanella oneidensis*.
Phosphoproteome of *Shewanella oneidensis* MR1

Gyorgy Babnigg* (gbabnig@anl.gov), Cynthia Sanville-Millard¹, Carl Lindberg¹, Sandra L. Tollaksen¹, Wenhong Zhu², John R. Yates III², Jim K. Fredrickson³, and Carol S. Giometti¹

¹Argonne National Laboratory, Argonne, IL; ²Scripps Research Institute, La Jolla, CA; and ³Pacific Northwest National Laboratory, Richland, WA

Protein phosphorylation plays an important role in the regulation of cell physiology in both prokaryotes and eukaryotes. Although bacterial protein phosphorylation on histidine residues is well known to be involved in signal transduction, phosphorylation of serine, threonine or tyrosine residues has only recently been described. We are using a suite of proteomics tools, including affinity chromatography, two-dimensional gel electrophoresis (2DE), Western blotting, tryptic peptide mass analysis, and phosphopeptide characterization to identify phosphoproteins expressed by *Shewanella oneidensis* MR-1 cells grown under different conditions and to determine whether or not serine, threonine, or tyrosine phosphorylation events are involved in the regulation of *S. oneidensis* MR-1 metabolism. By probing 2DE patterns of whole cell lysate patterns with antibodies directed against specific phosphorylated-amino acids, proteins containing phosphoserine, phosphotyrosine, and phosphothreonine have been detected. In a comparison of cells grown aerobically and with oxygen limitation (suboxic), differential expression of phosphoproteins has been observed (Figure 1).

Sixteen phosphoproteins have been identified by tryptic peptide mass analysis using LC-MS/MS; identified proteins include GGDEF domain protein, translation elongation factor Tu, and formate acetyltransferase. Phosphorylated proteins have also been enriched from cell lysates using immobilized metal affinity chromatography prior to analysis by 2DE or LC-MS/MS. The isolation of phosphoproteins by affinity chromatography is also enabling characterization by Fourier transform MS/MS to identify the actual sites of phosphorylation and to determine whether or not there is site-specific phosphorylation in response to different growth conditions. Such characterization is of particular interest in the case of formate acetyltransferase (Figure 2), since this protein has more phosphorylated forms under suboxic than under aerobic growth.

---

**Figure 1.** 2DE patterns of MR-1 phosphoproteins isolated by affinity chromatography and detected by silver stain (A,C) or by reaction with anti-phosphoserine antibody (B,D). Cells were grown aerobically (A,B) or with limited oxygen (C,D).
conditions, suggesting a regulatory function (Figure 2). This work will be extended to the study of cells harvested from the *Shewanella* Federation chemostat experiments in order to follow the phosphorylation events through transition between different growth conditions.

This work is supported by the U.S. Department of Energy, Office of Biological and Environmental Research, Genomics:GTL Program, under contract number W-31-109-ENG-38.

**Figure 2.** 2DE pattern of formate acetyltransferase from MR-1 cells grown aerobically (A) or with limited oxygen (B); detected using anti-phosphoserine.

---

### The *Shewanella* Federation: Functional Genomic Investigations of Dissimilatory Metal-Reducing *Shewanella*

**James K. Fredrickson**\(^*\) (Jim.Fredrickson@pnl.gov), Margaret F. Romine\(^1\), Carol S. Giometti\(^2\), Eugene Kolker\(^3\), Kenneth H. Nealson\(^4\), James M. Tiedje\(^5\), and Jizhong Zhou\(^6,11\), Monica Riley\(^7\), Shimon Weiss\(^8\), Christophe Schilling\(^9\), and Timothy S. Gardner\(^10\)

\(^1\)Pacific Northwest National Laboratory, Richland, WA; \(^2\)Argonne National Laboratory, Argonne, IL; \(^3\)BIATECH, Bothell, WA; \(^4\)University of Southern California, Los Angeles, CA; \(^5\)Michigan State University, East Lansing, MI; \(^6\)Oak Ridge National Laboratory, Oak Ridge, TN; \(^7\)Marine Biological Laboratory, Woods Hole, MA; \(^8\)University of California, Los Angeles, CA; \(^9\)Genomatica, Inc., San Diego, CA; \(^10\)Boston University, Boston, MA; and \(^11\)University of Oklahoma, Norman, OK

*Shewanella oneidensis* MR-1 is a motile, facultative γ-Proteobacterium with extensive metabolic versatility with regards to electron acceptor utilization; it can utilize O\(_2\), nitrate, fumarate, TMAO, DMSO, Mn, Fe, and S\(^0\) as terminal electron acceptors during anaerobic respiration. The ability to effectively reduce polyvalent metals including solid phase Fe and Mn oxides and radionuclides such as uranium and technetium as well as to function as a catalyst in mediator-less microbial fuel cells has made *Shewanella* an excellent model organism for understanding biogeochemical cycling of metals and anaerobic electron transport pathways. Complete sequencing of the MR-1 genome has enabled the application of high throughput functional genomics methods for measuring gene and protein expression. The *Shewanella* Federation (SF), a collaborative scientific team, is applying these approaches to achieve a system-level understanding of how MR-1 regulates energy and material flow and to utilize its versatile electron transport system to reduce metals and transfer electrons to electrode surfaces. The SF has developed an integrated approach to *Shewanella* functional genomics that capitalizes on the relative strengths, capabilities, and expertise of the various members. A major emphasis has been placed on integrating prediction and experiment in an iterative fashion to unravel the network structure that controls the flow of energy and materials through cells. SF members share information, resources and collaborate on projects that range from a few investigators focused on a defined topic to more complex “Federation-level” experiments that utilize combined SF capabilities to address more global scientific questions. The SF is organized into integrated working groups focused on: (1) the functional genomics of energy flow and electron transport regulation; (2) the characterization and modeling of metabolic and regulatory networks; (3) the comparative genomic and physiologic analyses of multiple *Shewanella* species to develop an evolutionary model for
Milestone 2

*Presenting author*
The SF holds biannual meetings that are rotated among the various partners. These meetings include PIs, collaborators, staff, and students and are used to review and plan collaborative research and to discuss a variety of technical issues ranging from data sharing mechanisms to growth and mutagenesis protocols. These meetings are open and often include non-DOE funded researchers who have an interest in the biology of *Shewanella*. SF members share experimental data, protocols and materials (mutants, reporter gene constructs, clones etc.) and maintain a collective annotation of the MR-1 genome that is based on input and data from multiple sources. Future directions of the SF include investigations into the population and community biology of *Shewanella* to begin to link genomics to populations, communities and evolution.

### 70

**An Integrated Knowledge Resource for the *Shewanella* Federation**

Nagiza F. Samatova\(^1\)* (samatovan@ornl.gov), Nicole Baldwin\(^1\), Bing Zhang\(^1\), Buyng-Hoon Park\(^1\), Denise Schmoyer\(^1\), Yuri Gorby\(^1\), Grigoriy E. Pinchuk\(^2\), Timothy S. Gardner\(^3\), Mary S. Lipton\(^2\), Samuel Purvine\(^2\), Angela D. Norbeck\(^2\), Margaret F. Romine\(^2\), Gyorgy Babnigg\(^4\), Carol S. Giometti\(^4\), James K. Fredrickson\(^4\), and Ed Uberbacher\(^1\)

\(^1\)Oak Ridge National Laboratory, Oak Ridge, TN; \(^2\)Pacific Northwest National Laboratory, Richland, WA; \(^3\)Boston University, Boston, MA; and \(^4\)Argonne National Laboratory, Argonne, IL

The *Shewanella* Federation (SF) represents a distributed group of investigators formed to understand the metabolic potential of *Shewanella* species, particularly with relevance to metal reduction and bioremediation applications. The Federation is generating large volumes of data of many different types, however, current data management strategies primarily rely on localized solutions and ad hoc data exchange procedures. The lack of integrative bioinformatics solutions is an impediment to bringing this major DOE investment to its full potential. Based on highest priority needs presented by many individual SF researchers, this project aims to construct a data and knowledge integration environment that will allow investigators to query across the individual research domains, link to analysis applications, visualize data in a cell systems context, and produce new knowledge, while minimizing the effort, time and complexity to participating laboratories. Specifically, major goals are: (1) to develop strategies for capturing and integrating diverse data types into common data models that support systems biology investigation, (2) to develop tools and processes to catalog and retrieve high-throughput data from warehoused and non-local data storage, (3) to construct a data and knowledge base that integrates gene, protein, expression and pathway-level knowledge, and (4) to incorporate interfaces for navigation and visualization of the multi-dimensional data produced.

During the initial three months of the project the following progress has been made. We have configured and built a server hardware and software infrastructure to provide efficient and reliable (24/7; 100% reliability; zero data loss) future access to the data generated by SF. We collected sample data files and associated meta-data including those for experimental protocols, raw experimental data, pre-processed and computationally analyzed data from the majority of SF sites to help with system design. Based in part on this input, we designed a comprehensive relational schema, compliant with community accepted data standards, that currently is capable of capturing fermentation, microarray, proteome, and interactome data. This high-level schema design captures information about Projects, Cell Culture, Experiments (fermentation, microarray, proteomics, MS pull-downs, etc.), Computational Analyses, and User(s) (figure, next page). For some sites, where local relational schemas are available, the proper schema mapping strategies have been developed. In
addition, we designed a schema for publicly available data sources and integrated this schema with the one for SF experimental data and meta-data. We wrote various tools for data ingest from these public databases, data parsing, and upload to the *Shewanella* knowledgebase including sequence databases (GenBank, TIGR, RefSeq, UniProt, InterPro, Pfam, ProDom, SMART), structure databases (PDB, COILS, SOSUI, PROSPECT), pathway databases (KEGG), and protein interaction databases (STRING, 3DID, DIP). Each of these data sources can be queried through a common interface that supports both simple and advanced search capabilities suitable for both browsing and for complex queries. To accept data and meta-data from various researchers in SF, we developed a web-based system for users to upload the information relevant to their experiments, and automatically update the database, thus making this information searchable across all the previously entered information. The system currently supports fermentation experiments and is being extended to other protocols (see figure, above). The system provides the capability for SF sites to edit and annotate the information, to save experimental protocol description in Excel format, to restrict the choices using controlled vocabularies, and to validate entered fields. The system is based on XML, XSD, and JSP/Bean technologies. On the analysis side, we developed tools for quantification of data from either stable isotope labeling or label-free LC-MS/MS shotgun proteomics. The tools will be freely available from http://MSProRata.org. In addition, we applied our protein-protein interaction prediction tools to *Shewanella* and are making the results of these predictions available through the *Shewanella* knowledgebase http://modpod.csm.ornl.gov/shew.

Research sponsored by the DOE Genomics:GTL Program.
Physiological Characterization of Genome-Sequenced *Shewanella* Species

Radu Popa, Rodica Popa, Anna Obraztsova, A. Hsieh, and Kenneth H. Nealson* (knealson@usc.edu)

University of Southern California, Los Angeles, CA

The genus *Shewanella* consists of a widely distributed group of facultatively anaerobic bacteria that are renowned for their ability to reduce many different electron acceptors and in particular for their ability to reduce solid phase iron and manganese oxides. *S. oneidensis* MR-1 was the first of this genus to be sequenced and annotated, followed by a number of other strains through DOE’s Joint Genome Institute in cooperation with the *Shewanella* Federation (SF), providing a great opportunity for combined comparative physiology and genomic analyses. That is, to ask, among other questions, the extent to which the physiology of a given bacterium can be predicted by analysis of its genome, and if not, what must we learn to be able to accurately do so. A necessary step in this co-analysis is the physiological analysis of the sequenced strains. This report includes a summary of physiological characterization of several sequenced species and strains of *Shewanella*. A companion abstract (J. Tiedje et al.) describes the initial results of genomic and proteomic analyses of a subset of the sequenced strains.

The ability to respire different electron donors was tested via the Biolog™ phenotype array system. This system utilizes 96-well plates, in which each well contains a different energy source. If a substrate is respired, a dye (formazan) is reduced, producing a blue color that is scored using a plate reader. Thus, the system scores respiratory ability, but not necessarily growth. Two different Biolog™ plates were utilized, providing a total of approximately 190 different substrates. Of these, approximately 10 were respired by all 7 strains, and another 15 were respired by 6 of the eight strains tested. Some strains, such as MR-1 utilized only 22 different substrates, while others had a much wider substrate range, up to 41 different compounds, including many hexoses and complex molecules.

The ability to grow on substrates that can be respired was tested on all substrates that were positive for respiration. This work is in progress, but relates to the issue of whether a given substrate might be susceptible to oxidation, yet not utilized as a substrate for growth. For strain MR-1, a surprisingly large number of substrates that were respired were not capable of supporting growth in the standard minimal medium.

The ability to grow anaerobically utilizing different electron acceptors was tested on a variety of different electron acceptors. This work is also in progress, and involves the screening of each strain for the ability of several substrates that were known to support growth aerobically to grow anaerobically on various electron acceptors. While the physiological mechanisms are not yet elucidated, it is clear that many strains utilize some electron donors quite well for some electron acceptors, but not for others. These results suggest that the internal electron transport webs of MR-1, and other shewanellae, are more complex than has been appreciated. A current objective of the *Shewanella* Federation is to elucidate the role of various c-type cytochromes in anaerobic respiration, in *S. oneidensis* MR-1, via a combination of bioinformatics and experimental analyses. This will greatly facilitate cross-genome comparisons and testing of hypotheses regarding the role to specific cytochrome homologs across the different strains.

All in all, the strains so far analyzed exhibit a wide range of physiological diversity, consistent with observations of genetic diversity at the genome scale, thus providing ample fodder for additional detailed genomic/physiology analyses.

* Presenting author
The Use of Microbial Fuel Cells (MFCs) for the Study of Electron Flow to Solid Surfaces: Characterization of Current Producing Abilities of Mutants of *Shewanella oneidensis* MR-1, of Other Strains and Species of *Shewanella*

Orianna Bretschger*, Byung Hong Kim, In Seop Chang, Margaret F. Romine, Jim K. Fredrickson, Yuri A. Gorby, Samantha B. Reed, David E. Culley, Soumitra Barua, and Kenneth H. Nealson

1 University of Southern California, Los Angeles, CA; 2 Korea Institute of Science and Technology, Seoul, Korea; 3 Pacific Northwest National Laboratory, Richland, WA; 4 Oak Ridge National Laboratory, Oak Ridge, TN

*Presenting author*  

*Shewanella oneidensis* MR-1 is a Gram negative facultative anaerobe capable of utilizing a broad range of electron acceptors, including several solid substrates. In addition, it has been known for many years that MR-1 can catalyze current production in microbial fuel cells (MFCs) without the addition of electron shuttles. We have examined a number of mutants of strain MR-1 for their ability to produce current, in an effort to determine whether the pathway of electron flow to solid metal oxides might be the same, or similar to, that used for current production. If so, it may be possible to use the MFC as a quantitative measure of electron flux: essentially as a proxy for the ability to reduce solid substrates. We summarize herein the results of MFC screening of mutants altered in both structural and regulatory genes.

**Our microbial fuel cell is a two chambered cell** separated by a proton-permeable membrane. On the anode side, *S. oneidensis* MR-1 is contained under anaerobic conditions, where it catalyzes the oxidation of substrate, producing electrons that flow through the anode electrode and across an external circuit. The protons produced by MR-1 diffuse through the membrane to the cathode side. A platinum catalyst is used at the cathode to convert the electrons, protons and molecular oxygen to water.

**Mutants defective in various cytochromes, as well as various regulatory elements has been made by members of the Shewanella Federation (SF).** These mutants are targeted deletions constructed by either homologous cross-over using host-encoded recombinases (PNNL group), or by introduced phage cre-lox recombinases (ORNL group). These mutants display a wide range of properties, but the over-riding feature seen is that mutants that inhibit the ability of MR-1 to reduce solid phase iron also inhibit the ability to produce current in the MFC, and to a similar degree, suggesting that the mechanisms involved in both processes require similar regulatory and structural components. In particular, we note that mutants in the *mtrA*, *mtrB*, *mtrC* or *omcA* structural genes are always deficient in both iron reduction and current production.

**In addition, a number of Shewanella species and strains were screened for their ability to produce current.** These included, in addition to *S. oneidensis* MR-1, 7 strains that were recently sequenced by the JGI. Of interest in this study was the strain OS217, *S. denitrificans*, which is lacking the cassette of genes (*mtrA*, *B*, and *C*, and *omcA*) that are believed responsible for current production and solid iron reduction in the other *Shewanella* strains. All of the strains, including OS217 produced comparable current, suggesting that a different mechanism must be utilized for both metal reduction and current production by this organism.
In addition to serving as a platform for electron flux studies, the MFC was also used for the analysis of nutrients during current production. Monitoring of reactants and products during current production allowed real time assessment of metabolic flow as related to current production.

73

Characterization of c-Type Cytochromes in Shewanella oneidensis MR-1

Margaret F. Romine1* (margie.romine@pnl.gov), Samantha B. Reed1, Soumitra Barua2,5, Samuel O. Purvine1, Alex S. Beliaev1, Haichun Gao2–3, Zamin Yang2–5, Yunfeng Yang2–5, David E. Culley1, David W. Kennedy1, Yuri Londer3, Tripti Khare3, Sandra Tollaksen3, Claribel Cruz-Garza3, Jun Li6, Mandy Ward6, Jizhong Zhou2,5, Frank Collart4, Mary S. Lipton1, James M. Tiedje4, Carol S. Giometti3, and Jim K. Fredrickson1

1Pacific Northwest National Laboratory, Richland, WA; 2Oak Ridge National Laboratory, Oak Ridge, TN; 3Argonne National Laboratory, Argonne, IL; 4Michigan State University, East Lansing, MI; 5University of Oklahoma, Norman, OK; and 6Johns Hopkins University, Baltimore, MD

Shewanella oneidensis MR-1 is a facultative aerobe that is capable of using fumarate, nitrate, nitrite, dimethylsulfoxide (DMSO), thiosulfate, trimethylamine oxide (TMAO), S°, and a variety of solid-phase and complexed metals including Fe(III) oxides, Mn(IV) oxides, Tc(VII), U(VI), and V(V) to drive respiration. Such a unique respiratory versatility of dissimilatory metal reducing bacteria like MR-1 is largely due to the abundance of c-type cytochromes that are revealed upon sequencing of their genomes. The predicted localization of a subset of these cytochromes to the outer membrane is thought to enable these organisms to utilize insoluble electron acceptors such as metal oxides and S°. Thus far four c-type cytochromes, OmcA, MtrC, MtrA, and CctA have been linked to respiration of Fe(III) and Mn(IV) by Shewanella sp. and an additional 10 are predicted to be necessary for respiration of oxygen, fumarate, nitrate, DMSO, TMAO, or S° by MR-1. An additional cytochrome c, CymA, is required for respiration of Fe, fumarate, DMSO, nitrite, and nitrate, but not TMAO. The roles of the remaining predicted cytochromes (nearly 75% of the total number) are have not yet been elucidated. This observation confirms that the mechanisms of electron transport in Shewanella remain poorly understood. Consequently, the Shewanella federation has developed a strategy to explore this metabolism further using an integrated approach that involves targeted gene knock-outs as well as comparative genomics, physiology, and gene/protein expression patterns. Summarized herein is our progress in determining the roles of S. oneidensis MR-1 c-type cytochromes in cellular respiration.

Prediction of Genes Encoding c-Type Cytochromes. Computational analysis of the MR-1 genome sequence revealed that the deduced amino acid sequences from 71 genes contain the signature CXXCH heme c binding motif. Functions of orthologs in other bacteria, the occurrence of conserved domains, and the presence of an expected sec leader peptide were used to narrow the number of predicted cytochromes to 44. Comparison to proteins deduced from genome sequences of 10 other Shewanella sp. enabled us to determine that SO3141 is degenerate, requiring 4 frameshifts to reconstruct the proper reading frames to produce the expected intact outer membrane decaheme cytochrome c. Furthermore, it was revealed that the 3’ and S’ end of genes SO4570 and SO4569, encoding a cytochrome and NfrC-like FeS protein, respectively, are replaced by 6 repeats of CAATGCTGA. A third gene, SO3623 encodes a split tetraheme flavocytochrome c and is intact, but the upstream gene encoding the flavin subunit is interrupted by an IS element. Consequently, it is unlikely that these proteins participate in respiratory processes in MR-1 without genomic rearrangement. In summary we have identified 41 genes that are predicted to encode c-type cytochromes.
Validation of $c$-Type Cytochrome Predictions. The occurrence of a CXXCH motif and a signal peptide is not sufficient to confidently determine that a protein is a $c$-type cytochrome since a large number of functionally unrelated proteins also possess a CXXCH motif. The covalently attached heme characteristic of $c$-type cytochromes can easily be detected by staining proteins separated by SDS-PAGE with a chemiluminescent ELISA substrate and therefore used as a method to validate our predictions and to map their mobility in acrylamide gels to facilitate their identification in ongoing 2D PAGE-based differential protein expression studies. We have successfully overexpressed, in Escherichia coli, 15 of the proteins predicted to bind 4 or fewer hemes and demonstrated that they bind heme in 1D acrylamide gels. An additional 3 proteins were expressed but not soluble. In addition to location mapping in 2D gel separation systems, these protein preparations will also be characterized by AMT mass tag technology to identify signature peptides that uniquely identify these proteins in cellular protein extracts.

Conditions that Promote Expression of Predicted $c$-Type Cytochromes. The identification of conditions that uniquely enable expression of RNA or protein from these genes provide useful clues of their function and determine the conditions necessary for detecting aberrant phenotypes in mutants. Microarray analyses of MR-1 cells grown with fumarate were evaluated for changes in gene expression after a shift to 10 alternate electron acceptors (Beliaev, et al. 2005). Results revealed a surprisingly widespread induction of cytochrome $c$ genes with thiosulfate and conversely the paucity of genes induced by DMSO. Many of the genes induced include those which are functionally uncharacterized suggesting that growth or reduction of thiosulfate and related sulfur containing compounds should be further investigated. AMT tag proteome analysis of MR-1 extracts collected from aerobic, suboxic, and anaerobic cultures using fumarate as the electron acceptor have confirmed expression of 22 of these genes, including several encoded by genes whose expression did not change significantly in microarray analyses. The combined results from these expression experiments allow us to distinguish the cytochromes that are required for multiple respiratory metabolisms from those that are unique to one or a few types.

Targeted Deletion of $c$-Type Cytochromes. Targeted deletion of all but 5 of the predicted intact cytochrome $c$ encoding genes have been successfully constructed by either homologous cross-over with host-encoded recombinases (PNNL) or with introduced phage cre-loxP recombinases (ORNL). Each mutant was tagged with a unique bar code to facilitate tracking individual strains in planned competitive growth studies. Preliminary analyses on the ability of these mutants to grow in microtiter plates with different electron acceptors revealed 5 mutants with growth defects with nitrate. Surprisingly the $\Delta$napB mutant grew better in LB/nitrate medium than wild type cells, while a $\Delta$napA mutant could not grow at all. Nitrite, which is toxic, accumulated after 12 hours of growth by the WT strain only. We hypothesize that in the absence of napB, an alternative cytochrome $c$ supplies electrons to napA for subsequent reduction of nitrate. Removal of nitrite by this alternative pathway is more efficient thereby enabling cells to attain a higher biomass yield. Mutants in the high-affinity $cbb_3$ cytochrome oxidase components exhibit a defect in both O$_2$ and TMAO electron acceptors suggesting a role for this complex in both suboxic and anaerobic respiratory processes. Defects in the reduction of Mn(IV) relative to WT MR-1 was evident in 10 different mutants suggesting a complex network of electron transfer reactions. Mutants were also evaluated anaerobically for energy taxis to Fe(III)-citrate, nitrate, nitrite, TMAO, DMSO, and fumarate using swarm plate assays. Initial test results suggest that 2 mutants were defective in all 6 assays, while others showed defects for only selected substrates providing new clues of function for several uncharacterized cytochromes.

Reference

Genome-Wide Transcriptional Responses to Metal Stresses in *Caulobacter crescentus*

Gary L. Andersen* (GLAndersen@lbl.gov), Ping Hu†, Eoin L. Brodie‡, Yohey Suzuki§, and Harley H. McAdams

1Lawrence Berkeley National Laboratory, Berkeley, CA; 2Stanford University School of Medicine, Stanford, CA; and 3Research Center for Deep Geological Environments, Ibaraki, Japan

Effective bioremediation of metal contaminated sites requires knowledge of genetic pathways for resistance and biotransformation by component organisms within a microbial community. The bacterium *Caulobacter crescentus* and related stalk bacterial species are known for their distinctive ability to live in low nutrient environments, a characteristic of most heavy metal contaminated sites. *Caulobacter crescentus* is also a model organism for studying cell cycle regulation with well developed genetics. We have identified the pathways responding to heavy metal toxicity in *C. crescentus* to provide insights for possible application of *Caulobacter* to environmental restoration. Using a custom Affymetrix GeneChip array designed by the McAdams laboratory at Stanford University, analyses of genome wide transcriptional activities of *C. crescentus* cells post exposure to four heavy metals (chromium, cadmium, selenium and uranium) presented significant knowledge how *Caulobacter crescentus* activates different mechanisms in response to various metal stresses. Surprisingly, at the uranium concentration close to the highest observed at the NABIR Field Research Center (200 µM), *Caulobacter* growth rate was not significantly affected and it was not until a concentration of 1 mM uranium that *Caulobacter* growth slowed. Under the same conditions, growth of *E. coli* K-12 was completely stopped and the growth of *Pseudomonas putida* KT2440 was drastically reduced. To investigate the possible uranium resistance mechanism utilized by *Caulobacter crescentus* we performed transmission electron microscope (TEM) with energy-dispersive x-ray spectroscopy (EDX) analysis and demonstrated that *C. crescentus* did not form any uranium-containing phosphate granules intracellularly. However, TEM images of whole cells of *C. crescentus* revealed extracellular precipitates associated with the cells. EDX spectra from cells and extracellular precipitates showed that while uranium is almost absent within cells, extracellular precipitates contain high concentrations of uranium, phosphorus and calcium, suggesting that the extracellular precipitates are composed mainly of these elements. Based on the chemical composition, the precipitates are thought to be the uranyl phosphate mineral autunite, a major source of naturally occurring secondary uranium ore and is known to persist under oxidizing conditions on a geological time-scale. Transcriptional analysis did show a protein candidate, which may involved in the uranium precipitation process. Two two-component systems were identified to be specifically up-regulated in response to uranium stress. One pair of knockout mutants was studied to identify their possible targets. We also identified differentially expressed transcripts from antisense strand of a predicted gene responding specifically to metal stresses. Further studies may elucidate functions of these transcripts. The combination of whole genome transcriptional analysis, phenotypic and genetic studies and advanced imaging provided powerful insights into mechanisms of uranium resistance mechanisms by *Caulobacter crescentus*. 

* Presenting author
Host Gene Expression Responses: Unique Identifiers of Exposure to Biothreat Agents

Rasha Hammamieh* (rasha.hammamieh@na.amedd.army.mil), Steven Eker\(^2\), Patrick Lincoln\(^2\), and Marti Jett\(^1\)

\(^1\)Walter Reed Army Institute of Research, Silver Spring, MD and \(^2\)SRI International, Inc., Menlo Park, CA

We have developed bioinformatic tools to facilitate mining of high throughput genomic and proteomic data. Today’s fast-growing sphere of bioinformatics, where retrieving precise information on massive datasets can reveal in-depth understanding of systems biology. Using our program GeneCite, scientists can interconnect two input files via any of the three available Boolean operators at NCBI web domain. After completion of a given search, GeneCite provides a summary of result briefing total number of hits etc., and two output files. First file provides literature citation counts for each given search key, while the other file offers hyperlinks for each query connecting the appropriate result page of the data source. The other tool, PathwayScreen, takes a list of Gene ID numbers and outputs a file listing the pathways that those genes are in and a link to any appropriate resources, namely BioCarta.com. The SRI team has developed approaches for model building to identify unique gene patterns that can distinguish among biothreat pathogenic agents.

High Throughput Fermentation and Cell Culture Device

David Klein and Stephen Boyer* (sboyer@gener8.net)

Gener8 Inc., Mountain View, CA

The focus of our Phase II SBIR project is the creation of a high-throughput screening platform capable of delivering the essential controls of a stirred-vessel bioreactor (pH, dissolved oxygen, and temperature) in a small-scale, inexpensive, robust, easy-to-use, disposable format. The result of our work to date is an array of 24 x 10ml reactors in an SBS plate format. The capabilities of current MicroReactor systems will be reviewed and industrial as well as GTL research applications will be discussed. Industrial applications include clone and media screening as well as simple factorial design-of-experiments studies. GTL research applications include parametric studies of microbial physiology, cultivation of poorly characterized organisms; other applications might include controlled and reproducible sample generation for coarse studies of the microbial proteome, transcriptome, physiome, and metabolome.
Genome-Wide Biochemical Characterization of Plant Acyl-CoA Dependent Acyltransferases

Chang-Jun Liu* (cliu@bnl.gov), and Xiao-Hong Yu
Brookhaven National Laboratory, Upton, NY

Acyl-CoA dependent acyltransferases catalyze the transfer of aliphatic and/or aromatic acyl moiety from CoA thioester donor to the nucleophile (OH- or NH-) of acceptor molecules. In plants enzymatic O- or N-acylation reactions are central to both primary and secondary metabolism, and are essential for plant growth and development, and plant environmental interactions. Particularly, acyl-CoA dependent O-acylation participates in plant cell-wall polysaccharide biosynthesis (Teleman et al., 2003), and the formation and deposition of heartwood-forming secondary metabolites in tree species, implicating significant biotechnological applications in genetic manipulation of lignocellulosic properties and carbon sequestration. Consistent with their multifaceted biological roles in plant metabolism, development and disease resistance, acyl-CoA dependent O-acyltransferases comprise a large and highly divergent protein family, known as BAHD superfamily (St-Pierre and De Luca, 2000). This family of enzymes consists of two conserved motifs, HXXXD and DFGWG in their primary sequences. Based on these sequence signatures, we employed tblastn algorithm searching poplar genomics sequences (http://genome.jgi-psf.org) and Tomato (SOL genomics network), Medicago truncatula (http://www.tigr.org/tigr-scripts/tgi/) EST databases and identified approximate 48 BAHD members in poplar genome, 48 Unigene in tomato unigene database, and 38 Tentative Consensus in Medicago truncatula EST database. Functional annotation of these large numbers of putative acyltransferase genes only based upon the sequence similarity with a few function known acyltransferases was difficult and extremely unreliable, due to the function diversity and primary sequence divergence. To unequivocally characterize the biochemical functions of putative acyltransferases in genome-wide, we developed a high throughput biochemical assay procedure that consists of efficient Gateway cloning for expression vector construction, magnetic Ni-particles and microdialysis for rapid protein extraction and purification, 96 well-plate formatted in vitro assay, and single-well, on-line product detection and identification by High Performance Liquid Chromatography-Electron Spray Ionization Mass Spectrometry. Adapting the developed method and combining with structural homology modeling based functional prediction, we have characterized 5 novel acetyl-CoA and malonyl-CoA dependent, (iso)flavonoid and anthocyanin biosynthetic acyltransferases from M. truncatula EST clones. The details of biochemical characterization will be discussed.

* Presenting author
The MAGGIE Project: A Mass-Based Platform for Protein and Metabolite Characterization

Gary Siuzdak1 (siuzdak@scripps.edu), Sunia Trauger1, Francis E. Jenney Jr.2, Steven M. Yannone3, Daojing Wang3, Nitin S. Baliga4, Steven R. Holbrook3, Chris H.Q. Ding3, Angeli Menon2, John A. Tainer1,3, and Michael W.W. Adams2

1Scripps Research Institute, La Jolla, CA; 2University of Georgia, Athens, GA; 3Lawrence Berkeley National Laboratory, Berkeley, CA; and 4Institute for Systems Biology, Seattle, WA

We have developed mass-based approaches for characterizing molecular complexes from Pyrococcus furiosus that include microLC ESI-MS/MS, microLC ESI-TOF, MALDIMS and DIOS-MS as well as a novel nonlinear software platform. These approaches initially examine proteins from the bacteria using multi-dimensional LC ESI-MS, followed by analysis of separated complex fractions via DIOS-MS, MALDI-MS and microLC ESI-MS/MS. Intact proteins of the separated complexes are also examined using MALDIMS. Among the approaches being developed are perfluorinated surfactant-enhanced desorption/ionization on fluorinated silicon to provide greater protein coverage for PTM analysis. In addition, a new affinity approach based on the selective fluorous-fluorous interaction between fluorous tagged analytes and the fluorous silylated porous silicon (pSi) surface. By employing a simple washing procedure, a mixture containing target analytes deposited on the fluorous silylated pSi surface are selectively captured and enriched by affinity purification thereby facilitating its analysis. Metabolite data is simultaneously being generated with these analytical tools to investigate their role in these interactions as well as metabolite characterization through a newly created online database of metabolite information (http://metlin.scripps.edu/). This research is largely synergistic with the efforts of Michael Adams and his ability to trap reversible and dynamic complexes enabling their purification. The platform along with XCMS software developments and initial results will be presented on the protein and metabolite characterization.

Organisation of Heavy Metal Resistance Genes in the Four Replicons of Cupriavidus metallidurans CH34

Sébastien Monchy1, Nicolas Morin1, Mohammed A. Benotmane1, Sébastien van Aelst1,2, Max Mergeay1,2, Tatiana Vallaey1,2, Ruddy Wattiez1,2, Safyih Taghavi2, John Dunn3, and Daniel van der Lelie* (vdlelie@bnl.gov)

1Belgian Center for Nuclear Studies, SCK/CEN, Mol, Belgium; 2Université Libre de Bruxelles, Belgium; 3INRA, Jouyen-Josas, France; 4Université Mons-Hainaut, Belgium; and 5Brookhaven National Laboratory, Upton, NY

Cupriavidus metallidurans (formerly Ralstonia) belongs to the phylum β-Proteobacteria and includes various isolates of soil bacteria adapted to harsh industrial biotopes. The genome of strain CH34 has been sequenced (Joint Genome Institute) and contains four replicons: two large plasmids pMOL28 (171 Kb) and pMOL30 (234 kb), and two megareplicons – a chromosome (3.9 Mb) that is especially rich in biosynthetic genes, and a megaplasmid (2.6 Mb). Analysis of C. metallidurans on the genome level against existing databases and through phylogenetic approaches, the transcriptome level using...
quantitative PCR and microarrays analysis, and the proteome level (using 2-D gel electrophoresis and mass spectrometry) indicate a high number of heavy metal resistance or -detoxification genes in comparison to other sequenced bacteria. These genes seem to be mainly associated with the two plasmids and the smaller megareplicon of *C. metallidurans*.

Before the CH34 genome sequence became available, research focused on the plasmid-encoded *czc* (Cd, Zn, Co), *cnr* (Co, Ni), *cbr* (Cr), *cop* (Cu), *pbr* (Pb) and *mer* (Hg) metal resistance operons, who had been identified via phenotypical analysis of plasmid-cured derivatives and cloning strategies. The whole genome sequence turned out to be a valuable resource for the identification of additional metal resistance systems, such as the *ars* operons located on the chromosome and several additional heavy metal efflux systems and their regulators.

In this communication we concentrate on the metal resistance determinants located on pMOL28 and pMOL30. Microarrays containing all ORFs as identified on the JGI draft sequence were hybridized with Cy3 and Cy5 labeled cDNA obtained from CH34 after induction with several heavy metals (30 min. induction with 0.5mM Cd, 5µM Hg, 0.8mM Zn, 0.4 mM Pb, 0.1mM Cu and 0.6mM Ni in 284 gluconate minimal medium at 30°C). We found 80 ORFs on a total of 161 located on pMOL28 and 134 ORFs on a total of 242 located on pMOL30 that were over-expressed in at least one metal condition.

Plasmid pMOL28 contains three clusters conferring resistance to nickel and cobalt (*cnr*), to chromate (*cbr*) and to mercury (*mer* of Tn4378). These three clusters constitute a 35 kb region which is flanked by IS1071 on the *mer* side and a deleted form of IS1071 on the *cnr* side. As expected, the *cnrYXHCBAT* cluster responded to Ni, but surprisingly also to Cu and Cd.

On plasmid pMOL30, heavy metal resistances are clustered opposite of the replication origin of the plasmid. Among other minor determinants, this region contains the *czc* *NICRADRE* cluster, a mercury transposon (Tn4380), the *pbr* *TRABCD* cluster, the newly identified *silABC* operon that responds to Cu, and a large cluster of 19 genes comprising the Cu-resistance operon (*copVTM-KNSRABCDIJGFLQHE*). Expression analysis revealed that the *czc* cluster was specifically induced by Zn and Cd. For the other metal resistance clusters, cross-responses inductions were observed to a much broader range of metals than expected from phenotypical analysis. Transcription of the *cop* cluster was not only induced by Cu but also by Zn, Cd and Ni, while transcription of the *pbr* cluster was not only induced by Pb but also Zn.

For both plasmids, the *mer* clusters’ transcription responded to Hg but also to Zn, Cd and Pb. In addition, the microarray analysis allowed use to identify hypothetical ORFs (including several potential signal peptides) whose expression was highly induced in the presence of specific heavy metals. We hypothesize that these peptides could be involved in coordinating a general metal response by *C. metallidurans* CH34 against heavy metals.

Sequencing of the *C. metallidurans* genome has been carried out at the Joint Genome Institute (JGI) under the auspices of the U.S. Department of Energy’s Office of Science, Biological and Environmental Research Program and the by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, Lawrence Berkeley National Laboratory under contract No. DE-AC03-76SF00098 and Los Alamos National Laboratory under contract No. W-7405-ENG-36. The draft sequence of the genome is now available at the http://genome.jgi-psf.org/draft_microbes/ralme/ralme.home.html web site.

References


* Presenting author

80

Geobacter Project Subproject II: Expression of Geobacteraceae Genes Under Diverse Environmental Conditions

Dawn Holmes1* (dholmes@microbio.umass.edu), Brad Postier1, Regina O’Neil1, Kelly Nevin1, Barbara Methe2, Jessica Butler1, Shelley Haveman1, and Derek Lovley1

1University of Massachusetts, Amherst, MA and 2The Institute for Genomic Research, Rockville, MD

The purpose of this research is to describe genome-wide patterns of gene expression in Geobacteraceae species exposed to a variety of environmental conditions. These results are important not only for defining the metabolic state of Geobacteraceae in environments of interest but also provide the data needed to further elucidate gene function, determine regulatory networks, and further refine in silico models to predict the metabolism and growth of Geobacteraceae during in situ uranium bioremediation or harvesting electricity from organic wastes. Gene expression patterns from a diversity of pure cultures of Geobacteraceae grown under a wide range of environmentally relevant conditions was evaluated. This information was used to identify key genes which could aid in diagnosing the physiological state and rates of metabolism of Geobacteraceae in subsurface environments by quantifying in situ levels of transcripts for these genes.

Key to the inexpensive analysis of gene expression in multiple Geobacteraceae genomes was the development of methods to take advantage of the Combimatrix® microarray technology. With the Combimatrix® electrochemical synthesis method, arrays can be custom-made for individual experiments at lower cost than for other platforms. This affords the flexibility of readily evaluating gene expression in any pure culture for which a genome is available or custom-designing arrays for analysis of environmental transcripts, without the large investment in microarray synthesis required with other platforms. Gene expression studies conducted with the Combimatrix® arrays gave results comparable to several other platforms we had routinely used in the past with lower cost and higher flexibility.

Numerous microarray studies were conducted on pure cultures of Geobacter sulfurreducens, Geobacter metallireducens and Pelobacter carbinolicus grown under different conditions of electron donor or electron acceptor availability in order to elucidate genes involved in electron transfer to various electron acceptors or in the metabolism of electron donors. For example, with the recent availability of the complete genome sequence of G. metallireducens it was possible to conduct microarray studies under a variety of growth conditions that helped elucidate mechanisms for environmentally significant processes such as dissimilatory nitrate reduction to ammonia and the metabolism of aromatic compounds. Comparison of cultures of P. carbinolicus grown under fermentative versus Fe(III)-reducing conditions are helping to identify components in electron transfer to Fe(III) that are conserved among the Geobacteraceae. Studies in which chemostat cultures of G. sulfurreducens were provided with low levels of oxygen or limited by nutrient availability revealed genes diagnostic of oxidative stress or nutrient limitation. Comparison of electron donor versus electron acceptor limiting conditions have elucidated gene expression patterns diagnostic of Geobacter species becoming limiting for electron acceptor during in situ uranium bioremediation.

* Presenting author
Although chemostat cultures are able to provide physiologically consistent cells which are ideal for comparing gene expression under different growth conditions, chemostats are not the most accurate representation of the subsurface sediments in which *Geobacter* species grow during *in situ* uranium bioremediation or on the surface of energy-harvesting electrodes. A major advance has been the development of techniques for extracting mRNA and quantifying gene transcript levels from cultures grown in subsurface sediments or on electrodes. For example, *G. metallireducens* was grown in an artificial sediment of synthetic poorly crystalline Fe(III) oxide or in actual subsurface sediments from the *in situ* uranium bioremediation field site in Rifle, Colorado. In both instances there was significantly higher expression of genes associated with motility and chemotaxis in these cultures than in cultures grown on soluble electron acceptors. These results are significant because they provide insight into the unexpected result in field studies that many of the *Geobacter* involved in *in situ* uranium bioremediation are planktonic, rather than attached to the sediment particles, as was previously considered.

Down-regulated genes during growth in the sediments included several NADH dehydrogenase proteins, a number of ribosomal proteins, and other proteins involved in translation, transcription, and cell division. This reflects the slower growth of *G. metallireducens* under these more environmentally relevant conditions and reveals genes whose levels of expression might be used to estimate rates of growth of *Geobacter* species in the subsurface. During growth in sediments, genes for nitrogen fixation and phosphate uptake were more highly expressed, demonstrating that *in situ* uranium bioremediation might be limited by nutrient availability. This corresponds with measurements of transcript levels during *in situ* uranium bioremediation. Methods for overcoming these limitations can now be studied with genome-wide monitoring of gene expression in the sediment cultivation systems.

Analysis of gene expression during growth on electrodes was expanded from the studies reported last year to include a greater diversity of environmental conditions. As will be detailed in the presentation, the results from these studies have resulted in a model for the electrical contacts between the cell surface and electrodes.

These pure culture studies have provided significant guidance as to how best to monitor the *in situ* metabolic state of *Geobacteraceae* in environments of interest. As noted in our abstract on *Geobacteraceae* genomic sequences in these environments, the heterogeneity in genome sequences across time and environments requires that focus be placed on key, highly conserved genes that are representative of key types of metabolism. Our library of primers that can be used for quantitative RT-PCR analysis to track transcript levels of key genes in subsurface environments continues to build. Examples, of how measuring the transcript levels of genes diagnostic of limitation for nitrogen and phosphorous, oxidative stress, and rates of central metabolism have helped define the physiological state of *Geobacter* species during *in situ* uranium bioremediation at the field study site in Rifle, Colorado will be detailed.
Comprehensive Study of recA Expression in Deinococcus radiodurans with Single Cell and Population Level Analyses

Emily H. Turner* (emilyt@u.washington.edu), Haley Pugsley, and Norman J. Dovichi

University of Washington, Seattle, WA

The goal of this project is the development of instrumentation and methodologies for the analysis of biologically relevant proteins in single bacteria. We have built an instrument that performs capillary electrophoresis with laser-induced fluorescence, capable of detecting fluorescently-tagged proteins at the sub-zeptomole level (1 zeptomole = 600 molecules) in single eukaryotes and prokaryotes. The data obtained at the single cell level is complimented with population level analyses by capillary electrophoresis, flow cytometry and fluorescence microscopy. The combination of single cell and population level data provides a comprehensive analysis of protein expression and distribution across a population.

We are currently studying expression of the RecA protein in Deinococcus radiodurans. D. radiodurans displays an extraordinary capacity to repair DNA damage. Binding of the RecA protein to single-stranded DNA induces the SOS response to DNA damage; RecA has been identified as key to the survival of Deinococcus following high levels of DNA damage. The population-wide expression distribution of RecA, and the relationship of this distribution to DNA damage tolerance, is not well characterized. We have produced a novel recA/eGFP construct in D. radiodurans, strain MaHa01, to analyze recA expression with established and newly developed proteomic technologies.

Capillary electrophoresis with laser-induced fluorescence detection. Capillary electrophoresis is used to detect recA/eGFP expression in single D. radiodurans and across a population. In the single bacterium studies, a cell of interest is identified with fluorescence microscopy and injected into a small bore capillary (10 µm inner diameter), where lysis occurs. Separation of cellular components occurs with the application of high voltage, and results in rapid and ultrasensitive detection of released eGFP. This analysis resolves eGFP fluorescence from native autofluorescent components, and is capable of detecting eGFP expression in a single Deinococcus at the sub-zeptomole level.

Capillary electrophoresis is also used to characterize relative expression distributions of recA across a population. Intact cells are continuously injected into the capillary; as the bacteria exit the capillary and pass through the laser, intracellular eGFP fluorescence is detected. Changes in recA distribution across a population are rapidly determined after inducing DNA damage. Performing this analysis, which is analogous to flow cytometry, produces population-level data from the same instrument that measures our single cell data.

Flow cytometry. To induce DNA damage in MaHa01, cultures are exposed to mitomycin C (MMC). In future studies, UV radiation and bleomycin exposure will be used to induce DNA double-strand breaks. Flow cytometry is used to sort a MMC-exposed MaHa01 population into groups with high and low recA expression. After a 96 hour recovery period, these groups are re-exposed to MMC, and their recA expression is measured. An adaptive response is observed for both high and low recA groups. Preliminary data suggests the presence of two recA expression distributions within a population; further study of these distributions is planned.

Fluorescence microscopy. The growth of D. radiodurans in single, pair and tetrad groupings presents a challenge in quantifying protein expression across a population. Image analysis of fluorescence...
micrographs provides recA expression distributions for single, pair and tetrad MaHa01 populations. Using fluorescence microscopy, we have also observed stochastic recA expression between members of a tetrad.

The use of the recA/eGFP construct with these technologies provides comprehensive determination of the recA expression distribution in D. radiodurans and the effects of exposure to DNA damage on this distribution. We have recently developed a D. radiodurans construct that expresses eGFP under the control of the heat-shock pGro promoter. Similar analyses of pGro expression will be carried out in future work.

82

Directed Evolution of Radioresistance in a Radiosensitive Species

J.R. Battista* (jbattis@lsu.edu) and Michael M. Cox

1Louisiana State University and A & M College, Baton Rouge, LA and 2University of Wisconsin, Madison, WI

Several bacterial species exhibit extraordinary resistance to ionizing radiation, surviving doses of 5,000Gy or higher without loss of viability. Although many hypotheses have been advanced to explain this radioresistance, very little is known about the specific mechanisms involved. To further our efforts to explain ionizing radiation resistance, we have taken the radiosensitive species, Escherichia coli strain MG1655, and subjected it to high dose ionizing radiation with the intent of generating a radioresistant strain that can be more easily studied than naturally radioresistant species. The protocol consisted of a series of selective steps in which successive exponential phase cultures were exposed to increasing doses of gamma radiation. The initial dose applied killed approximately 90% of the culture. The survivors were diluted into fresh growth medium and allowed to propagate. This process of irradiation and outgrowth was repeated for 21 generations. As the culture became more resistant to the effects of ionizing radiation the dose administered was increased. At the end of the study a purified resistant strain was recovered and its capacity to survive ionizing radiation evaluated. The evolved strain, which was designated 21-9 was approximately 500-fold more resistant at 5000Gy than its parent. The strain does not exhibit obvious phenotypic differences from MG1655, growing with normal kinetics in rich media at 37C. Analysis of genome restitution post-irradiation indicates the cells suffer DNA double strand breaks and those breaks are repaired with kinetics similar to those reported for irradiated cultures of D. radiodurans. 21-9 was re-sequenced using comparative genome sequencing, a microarray hybridization-based method developed by NimbleGen Systems Incorporated, to find mutations in the strain's genome. Sixty three differences were found in the genome of 21-9 relative to MG1655; 62 point mutations and one large deletion associated with the excision of the e14 prophage. Current efforts are focused on evaluating the role of these changes in ionizing radiation resistance.
83

Development of a *Deinococcus radiodurans* Homologous Recombination System

Sanjay Vashee* (SVashee@venterinstitute.org), Ray-Yuan Chuang, Christian Barnes, Hamilton O. Smith, and J. Craig Venter

J. Craig Venter Institute, Rockville, MD

A major goal of our Institute is to rationally design synthetic microorganisms that are capable of carrying out any required functions. One component of this effort entails the packaging of the designed pathways into a cohesive genome. Our approach to this problem is to develop an efficient in vitro homologous recombination system based upon *Deinococcus radiodurans* (Dr). This bacterium was selected because it has the remarkable ability to survive 15,000 Gy of ionizing radiation. In contrast, doses below 10 Gy are lethal to almost all other organisms. Although hundreds of double-strand breaks are created, Dr is able to accurately restore its genome without evidence of mutation within a few hours after exposure, suggesting that the bacterium has a very efficient repair mechanism. The major repair pathway is thought to be homologous recombination, mainly because Dr strains containing mutations in *recA*, the bacterial recombinase, are severely sensitive to ionizing radiation.

Since the mechanism of homologous recombination is not yet well understood in Dr, we have undertaken two general approaches to study this phenomenon. First, we are utilizing information from the sequenced genome. For example, homologues of *E. coli* homologous recombination proteins, such as *recD* and *ruvA*, are present in Dr. Thus, one approach is to assemble the homologous recombination activity by purifying and characterizing the analogous recombinant proteins. However, it is probable that not all genes that play a major role in homologous recombination have been identified by annotation. To overcome this potential obstacle, we are also establishing an endogenous extract that contains homologous recombination activity. This extract can then be fractionated to isolate and purify all proteins that perform homologous recombination. Progress made towards our goals will be presented.

84

Studies on the Fe Acquisition Mechanisms in *Nitrosomonas europaea* Derived from the Genome

Xueming Wei¹, Neeraja Vajrala¹, Norman Hommes* (hommesn@onid.orst.edu), Cliff Unkefer², Luis Sayavedra-Soto¹, and Daniel J. Arp¹

¹Oregon State University, Corvallis, OR and ²Los Alamos National Laboratory, Los Alamos, NM

*Nitrosomonas europaea* is a chemolithotrophic bacterium that can grow solely on ammonia (NH₃) and carbon dioxide (CO₂). The genome of *N. europaea* consists of ~2460 protein-encoding genes. We have been utilizing the genome information to guide the studies on its unique Fe requirement and uptake systems. Fe is often a limiting factor for the growth of most bacteria because in aerobic environments, Fe exists predominantly in the insoluble ferric form (solubility in H₂O at pH 7.0 is 10⁻¹⁸ M).

The *N. europaea* genome reveals that up to 4% of the coding genes are dedicated to the transport of Fe, yet it noticeably lacks genes for siderophore biosynthesis. There are 22 sets of genes that are
organized similarly to the \( \text{fecI/fecR/fecA} \) system (genes for \( \sigma \)-factor/anti \( \sigma \)-factor/TonB-dependent Fe-siderophore receptor/transducer), and 20 additional \( \text{fecA} \)-type genes that do not have associated \( \text{fecI/fecR} \) genes. All 22 genes in the first group encode outer membrane (OM) siderophore transducers that have an N-terminal extension, while 18 of the 20 genes of the second group code for OM siderophore receptors (lack of N-extension). These OM siderophore transducers/receptors are biochemically and phylogenetically diverse. 13 of the 42 OM receptor/transducer genes are either truncated or interrupted by IS elements and frame shifts and are likely not functional. Parallel to the large number of Fe acquisition-related genes, over 2% of the genome encodes proteins for heme and cytochrome biosynthesis and proteins with Fe-S centers. This high number of genes for Fe acquisition and Fe-containing proteins is consistent with the life style, especially the energy metabolism, of \( \text{N. europaea} \). Our study has confirmed that \( \text{N. europaea} \) has higher contents of cellular Fe and cytochromes than common species such as \( \text{E. coli} \).

We have determined the Fe requirement for \( \text{N. europaea} \) growth, and the effect of Fe limitation on cell physiology. Reverse transcriptase (RT)-PCR showed that 60% of the functional genes were expressed under either Fe-limited (0.2 \( \mu \)M) or Fe-replete (10 \( \mu \)M) conditions. The mRNA levels of a few genes appeared to be higher in Fe-replete cell than in Fe-limited cell. Four of the genes were expressed at much higher levels under Fe-limited condition than Fe-replete condition, all of which are genes encoding the siderophore receptors highly induced in Fe-limited cell that were identified by MS/MS. The expression of these genes at transcriptional level shows a diverse response to Fe availability. PAGE analysis also showed that several OM proteins were expressed at much higher levels under Fe limitation (0.2 \( \mu \)M Fe) than under Fe-replete (10 \( \mu \)M Fe) conditions. We have determined the identities of the differentially expressed OM proteins by HPLC tandem mass spectrometry (LC/MS/MS) analysis. Majority of these proteins are TonB-dependent receptors for siderophores such as ferrichrome and catechol-type siderophores. Included in these were proteins encoded by the four genes identified by RT-PCR as being highly expressed under Fe-limited condition. Interestingly, all of these OM proteins are true siderophore receptors that lack the N-terminal extension characteristic of the OM siderophore transducer family, and all are encoded by genes that do not have cognate \( \sigma \)-factor/anti \( \sigma \)-factor genes. Three of the six genes encoding these receptors are the only ones of the 29 intact siderophores transducer/receptor genes that are preceded by a putative Fur box, suggesting the possibility of regulation by Fur (ferric uptake regulator). An OM porin OmpC, a multicopper oxidase, and a type II secretion pathway protein were also among the highly expressed proteins in Fe-limited cells. Both OM porin and multicopper oxidase could be involved in Fe uptake. These results provide evidence that under Fe deficient conditions, \( \text{N. europaea} \) up regulates the expression of certain Fe-acquisition-related proteins.

The addition of exogenous siderophores to Fe-limited medium increased \( \text{N. europaea} \) growth (total cell mass), suggesting its capability of using external siderophores for efficient Fe uptake. By LC/MS/MS analysis, we have also identified two OM transducers (encoded by NE1097/1088, \( \text{foxA} \) homologues) specific for the siderophore desferal, and they were expressed only in Fe-limited, desferal-containing medium, confirming that the expression required the induction by desferal. Both single and double mutants with disrupted desferal transducer genes have been created. Characterization of these mutants showed that the double mutants (with both genes inactivated) could not grow in Fe-limited, desferal-containing medium. Interestingly, the mutant with a disrupted gene NE1097 has the same phenotype as the double mutant, but single mutants with a defective gene NE1088 was able to grow in desferal-containing medium only when Fe level was raised to \( 0.1 \mu \)M (5\( \times \)Fe-limited medium). These results suggest that the acquisition of desferal-bound Fe needs functional desferal transducers, providing direct biochemical and genetic evidence for the functionality of the putative siderophore transducer genes in \( \text{N. europaea} \). This result, together with the results from siderophore feeding experiments and elevated production of OM siderophore receptors under Fe...
limitation, re-enforces the notion that *N. europaea* can compete for Fe-loaded siderophores secreted by other microbes in its natural environments.

A putative regulatory mechanism for the desferal uptake system in *N. europaea* is proposed based on the genetic analysis and phenotypical behaviors of the desferal transducer mutants. The genome sequence shows that NE1097 exists with cognate *fecIR*-type genes, while there is no cognate *fecIR*-type gene preceding NE1088. Both OM desferal receptors are TonB-dependent transducers which can interact with anti $\sigma$-factor, which affects sigma factor, for the regulation of the expression of the systems. Based on these results and genetic information, it is likely that binding of Fe-loaded desferal to the transducers triggers the interaction between the transducers and the anti $\sigma$ factor, and only the transducer encoded by NE1097 could interact with its cognate anti $\sigma$ factor to turn on the expression of both genes. It also likely that the sigma factor encoded by the gene cognate to NE1097 could also interact with the promoter of gene NE1088 to turn on its expression.

References


85

Large Scale Genomic Analysis for Understanding Hydrogen Metabolism in *Chlamydomonas reinhardtii*

**Michael Seibert** *(mike_seibert@nrel.gov)*, Florence Mus, Alexandra Dubini, Maria L. Ghirardi, Matthew C. Posewitz, and Arthur R. Grossman

1National Renewable Energy Laboratory, Golden, CO; 2Carnegie Institution of Washington, Stanford, CA; and 3Colorado School of Mines, Golden, CO

While many taxonomically diverse microbes have the ability to produce $H_2$, only certain photosynthetic organisms, including the green alga, *Chlamydomonas reinhardtii*, are able to directly couple water oxidation to the photoproduction of $H_2$. A fundamental understanding of the metabolism in this prototype alga might enable the future development of a sustainable system for biological $H_2$ production. To work toward this understanding, we are exploiting whole genome sequence information and genomic tools that have been newly developed for *C. reinhardtii*. The completion of the *C. reinhardtii* genome sequence as part of the DOE Office of Science’s Genomics:GTL Program facilitated the recent development of a high-density DNA microarray at the Carnegie Institution. This array is based on synthetic ~70 mers that represent approximately 10,000 unique *C. reinhardtii* genes, and is a powerful tool that can be used to thoroughly explore genome-wide changes in cellular transcript levels that accompany the acclimation of *C. reinhardtii* to conditions facilitating $H_2$ production.

Hydrogenase (the enzyme that catalyzes $H_2$ production) activity in *C. reinhardtii* is induced by anaerobiosis, achieved either in the dark by inert gas purging of $O_2$ or in the light by depriving cultures of sulfate. The latter attenuates photosynthetic $O_2$ evolution, which allows *C. reinhardtii* to metabolize any residual $O_2$ in the culture vessel. The resultant anaerobic environment sustains photoproduction of volumetric amounts of $H_2$ for 4 days in batch cultures. Under sulfur-deprived conditions, algal cultures...
exhibit a mixed metabolic state in which anaerobic fermentation, oxygenic photosynthesis and aerobic respiration co-occur. Our goal is to understand the underlying physiological processes that enable *C. reinhardtii* to sustain H$_2$ production, and toward this goal we are examining changes in transcript abundance and the establishment of protein networks that accompany the development of algal H$_2$ production. The use of high density DNA microarrays provides an initial view of the ways in which a cell may modulate its metabolism under different environmental conditions, and a 3,000 element array was recently used to examine sulfur- and phosphorus-deprivation responses in *C. reinhardtii*. The new 10,000 element array represents over half of the *C. reinhardtii* transcriptome and will provide more comprehensive information on ways in which this alga adjusts to conditions that sustain H$_2$ production.

Our initial studies are focused on characterizing differential gene expression in WT cultures that are aerobically grown and then anaerobically acclimated. Experimental protocols using qPCR have been established to rigorously quantify transcript levels for the HydA1 and HydA2 structural [FeFe]-hydrogenase genes and the HydEF and HydG [FeFe]-hydrogenase assembly genes in *C. reinhardtii*. Moreover, we have used qPCR to investigate relative transcript abundance of several genes involved in glycolysis and associated with the acclimation of the cells to sulfur-deprived growth conditions. Using *C. reinhardtii* strain CC425, the qPCR data indicate that there is a dramatic increase in abundance of the transcript encoding pyruvate ferrodoxin oxidoreductase (85-fold under appropriate conditions) during dark anaerobiosis, and that increased levels of transcripts for HydA structural genes (45-fold for HydA1 and 25-fold for HydA2) and hydrogenase assembly genes (90-fold for HydG and 40-fold for HydEF) develop under anoxic conditions. Initial microarray data examining the differential expression of genes following dark-anaerobic induction were obtained. These data demonstrate significant changes in the transcript levels of several genes associated with signal transduction, transcriptional regulation, translational regulation, posttranslational modification, as well as photosynthesis, electron transport, proton transport, fermentation, stress response physiology, and a variety of other metabolic processes. Interestingly, the transcripts for several ribosomal proteins increase when the cells experience anaerobic conditions, indicating the possibility of significant changes in protein synthesis during anoxia. We also observe elevated levels of nitrate reductase transcript, which may reflect the establishment of a competing pathway for electrons away from H$_2$ production. Finally, transcripts from genes associated with oxidative stress also rise during anaerobiosis. These data provide the first insights into the metabolic pathways utilized by *C. reinhardtii* and the genome-wide changes in gene transcription that occur as this alga acclimates to an anoxic environment.

In addition to examining WT cultures, we have isolated several *C. reinhardtii* mutants at NREL (under another DOE Office of Science program) with attenuated H$_2$-photoproduction activity. We will compare gene expression profiles from these mutants with the WT under appropriate conditions. One such mutant lacks a functional HydEF gene, which is required to assemble an active hydrogenase enzyme. This hydEF-1 mutant is the only reported *C. reinhardtii* strain that is unable to produce any H$_2$ at all. Since hydEF-1 is specifically disrupted in its ability to synthesize an active hydrogenase, any gene that is differentially expressed in this mutant should be a consequence of the mutant’s inability to photoproduce H$_2$. Another *C. reinhardtii* mutant, sta7-10, is unable to accumulate intracellular starch. Interestingly, this mutant shows aberrant induction of hydrogenase transcript accumulation and attenuated H$_2$-photoproduction activity during anaerobiosis.

In sum, we have begun to develop a global understanding of factors that promote H$_2$ production during anaerobiosis by analyzing transcript profiles from WT cultures of *C. reinhardtii*. This work is elucidating the biochemical pathways utilized by *C. reinhardtii* during anaerobiosis and will provide insights into how mutants, altered in normal H$_2$ metabolism, acclimate to anaerobiosis. More detailed knowledge of the metabolic and regulatory context that facilitates H$_2$ production will be necessary to understand and ultimately correct current limitations in H$_2$-production yields.
86

Single-Molecule Imaging of Macromolecular Dynamics in a Cell

Jamie H.D. Cate and Haw Yang (hawayang@berkeley.edu)

Lawrence Berkeley National Laboratory, Berkeley, CA and University of California, Berkeley, CA

We are taking several approaches to make the tools that will be needed for single-molecule imaging of macromolecule dynamics in living cells. An apparatus that tracks a single moving nanoparticle in 3D while providing concurrent sequential spectroscopic measurements has been developed. The design is based on confocal microscopy and is the first step towards correlating the reactivity of a single molecule with its spatial location in cells. One critical element in tracking single molecules in cells is the optical probes that contrast the molecule against cellular background. This goal is approached by luminescence engineering of semi-conducting quantum dots (Qdot). We report detailed characterizations of suchengineered Qdots at the single particle level. The experimentally obtained lifetime-intensity correlation maps suggest that Qdot charging states are continuously distributed, and provide the physical foundation for luminescence engineering by synthesis. To place exogenous probes inside a bacterial cell in a controlled way, reliable methods to overcome the membrane barrier have to be developed. We characterize the two most commonly used methods, electroporation and heat shock, at the single-cell level. It was found that probes introduced via electroporation enter a cell primarily through the pole whereas those introduced via heat shock through the newly synthesized membrane. These observations provide important clues for controlled placement of exogenous probes into bacterial cells. Finally, we report the successful construction of a microfluidic mixer that allows studies of biological macromolecules under crowding conditions that are comparable to those inside a cell. This mixer platform is critical in validating any models that may form from our future in cellular studies. With the high-resolution ribosome structure solved, the new tools we are developing will form the basis for in cellular studies of protein synthesis machinery.

87

Probing Single Microbial Proteins and Multi-Protein Complexes with Bioconjugated Quantum Dots

Gang Bao, Grant Jensen, Shuming Nie, and Phil LeDuc

Georgia Institute of Technology, Atlanta, GA; Emory University, Atlanta, GA; California Institute of Technology, Pasadena; and Carnegie Mellon University, Pittsburgh, PA

We have been developing quantum-dot (QD) based strategies for imaging and identification of individual proteins and protein complexes in microbial cells. Currently, there is a lack of novel labeling reagents for visualizing and tracking the assembly and disassembly of multi-protein molecular machines. There is no existing method to study simultaneous co-localization and dynamics of different intra-cellular processes with high spatial resolution. As shown in Figure 1, the multifunctional quantum-dot bioconjugates we develop consisting of a quantum dot of 2-6 nm in size encapsulated in a phospholipid micelle, with delivery peptides and protein targeting ligands (adaptors) conjugated to the surface of the QD through a biocompatible polymer. After internalization into microbial cells, the adaptor molecules on the surface of QD bioconjugates bind to specific target proteins or protein complexes that are genetically tagged. Optical imaging is used to visualize the localization, trafficking
and interaction of the proteins, resulting in a dynamic picture but with a limited spatial resolution (~200 nm). The same cell is imaged by EM to determine their detailed structures and localize the target proteins to ~4 nm resolution. For each protein or protein complex, selected tags are tested to optimize the specificity and signal-to-noise ratios of protein detection and localization. This innovative molecular imaging approach integrates peptide-based cellular delivery, protein targeting/tagging, light microscopy and electron microscopy.

To achieve the goals of this DOE GTL project, we have successfully synthesized core-shell and alloyed CdHgTe quantum dots (QDs) for dual-modality optical and EM imaging. This new class of QDs contains Hg, a heavy element that is often used in x-ray and electron scattering experiments, allowing studies of cellular structures at nanometer resolution. We have also linked QDs to a chelating compound (nickel-nitrilotriacetic acid or Ni-NTA) that quantitatively binds to hexahistidine-tagged biomolecules with controlled molar ratio and molecular orientation.

We have tested a number of methods for delivery of QD probes into living cells, and identified the advantages and limitations of each method. For example, we explored the possibility of delivering QD probes into yeast and E. coli with high efficiency using different methods, including peptide-based delivery, heat shock, and the use of anti-microbial/permeabilizing agents. Specifically, we performed a preliminary study of peptide-based delivery of QD bioconjugates into yeast and E. coli using three different peptides, TAT, polyArg, and a peptide (ArgSerAsnAsnProPheArgAlaArg) that has been used for delivering GFP into yeast S. cerevisiae. We have also tested different tagging strategies including tetracysteine/FIAsH, SNAP tag, Histidine/Ni-NTA and Histidine-peptide.

As part of our effort to develop QD-based technologies to identify and track individual protein complexes in microbial cells, we have performed preliminary optical imaging studies of single QDs delivered into living cells. Using a spinning-disk confocal microscope, we have succeeded in imaging single QD probes delivered into the cytoplasm of living cells. Several lines of evidence support that the QDs in cells are indeed single: (a) these QDs have similar brightness and spot size; (b) the brightness of these QDs is not higher than that of single QDs on a coverslip; and (c) the intracellular QDs show intermittent on/off light emission (called blinking), a characteristic of single dot behavior.

We have also developed computation algorithms for two-color colocalization and correlation tracking of QD probes. As an alternative, we successfully imaged individual 10 nm gold nanoparticles and established the darkfield optical imaging capability for cellular studies.

We are advancing electron tomography as a promising new tool to image protein complexes both in-vitro and in-vivo within small microbial cells. A new helium-cooled, 300kV, FEG, “G2 Polara” FEI

![Figure 1](image_url)

Figure 1. (A) Schematic illustration of a multifunctional quantum dot bioconjugate consisting of encapsulated QD with targeting adaptor and delivery peptide on its surface; (B) correlated optical and EM imaging of the same cell gives both temporal and spatial information on a protein complex; (C) possible conjugation and tagging strategies for optimizing detection specificity and sensitivity. Note that molecules are not drawn to the exact scale.
TEM at Caltech was used to image purified protein complexes, viruses, and whole bacterial cells. We pioneered the use of a new “flip-flop” cryorotation stage that allows dual-axis cryotomography, and developed a simple Perl-based system for distributed computation to handle the massive image processing demands that arise from imaging intact bacteria in 3D. These technological advances have allowed us to visualize directly cytoskeletal elements within small microbial cells and the domain structure of purified multi-enzyme complexes, both are key imaging goals of the genomes to life program. For example, we produced three-dimensional reconstructions of several different types of bacteria, including some of great interest to the DOE (Magnetospirillum magneticum, Mycoplasma pneumonia, and Caulobacter crescentus), with unprecedented resolution and authenticity. We imaged chemoreceptor clusters, flagella, pili, polyribosomes, and other ultrastructural details, and identified five unique patterns of cytoskeletal filaments bundles likely involved in cell shape determination, establishment of polarity, and chromosome segregation in C. crescentus.

As a model system to study protein localization, we have been investigating the migration of Dictyostelium discoideum under defined extracellular stimuli. We have utilized custom-fabricated microfluidic devices to stimulate a cell in local domains both with 2D and 3D control while simultaneously visualizing its response with fluorescent microscopy using quantum dots. We targeted quantum dots to CRAC and G-actin and analyzed them for co-localization of the GFP and quantum dot signals in Dictyostelium. This technique will be further combined with high-resolution electron microscopy imaging to visualize individual proteins and protein complexes.

Acknowledgement: This research is funded by a grant from DOE (DE-FG02-04ER63785).

Hyperspectral Imaging of Photosynthetic Pigment Molecules in Living Cyanobacterial Cells

Jerilyn A. Timlin¹, Michael B. Sinclair¹, Howland D.T. Jones¹, Linda T. Nieman¹, Sawsan Hamad², David M. Haaland* (dmhaala@sandia.gov), and Wim F.J. Vermaas²

¹Sandia National Laboratories, Albuquerque, NM and ²Arizona State University, Tempe, AZ

Hyperspectral confocal imaging has the potential to revolutionize the quality and quantity of energy transfer information derived in situ from living, photosynthetic systems. We have developed this technology at Sandia, and applied it to intact, living cyanobacterial cells to follow energy transfer within the light-harvesting antenna of the photosynthetic apparatus. Light is absorbed by a number of different pigments in the cell, including phycobilins, chlorophyll a, and carotenoids. Phycobilins bound to specific proteins serve as an antenna complex (the phycobilisome) to funnel excitation energy to lower-energy pigments and eventually to a chlorophyll a molecule in a special protein environment, the photosynthetic reaction center. Phycobilin pigments and chlorophylls in excited states will usually transfer their energy to the reaction center chlorophyll, but can also decay by fluorescence emission or internal conversion. The fluorescence is in principle a signature of the pigments in the excited state, but the emission wavelengths of the various pigments are very close together.

Using hyperspectral fluorescence microscopy of intact cyanobacterial cells and multivariate analysis technology, we have identified five different fluorescent spectral signatures with maxima between 640 and 700 nm and mapped their location within the living cell (see Figure 1 for an example using
wild-type cells). This unique imaging system captures a full emission spectrum at each image pixel and uses multivariate curve resolution (MCR) technology\textsuperscript{1,2} to deduce the fluorescence spectrum and location of individual pigment species with single molecule sensitivity in three dimensions with a spatial resolution of 0.24 \textmu m in X and Y directions, and 0.6 \textmu m in the Z direction\textsuperscript{3}. A single cell can be imaged in all three dimensions with 50 ms temporal resolution.

Upon excitation at 488 nm, six fluorescence components are resolved from the hyperspectral images of intact cells: three phycobilins (phycocyanin (PC), allophycocyanin (APC), and allophycocyanin-B (APC-B)), chlorophyll \textit{a}, a weaker long-wavelength chlorophyll \textit{a} component peaking at \textasciitilde 698 nm and an extremely weak short-wavelength component. The latter has resonance-Raman bands detectable with our imaging system that correspond to carotenoids, primarily \textbeta-\textit{carotene}. These six spectral signatures are common to the wild type and mutant strains, but differences in their relative intensities and spatial locations between mutant strains make them useful for developing insight as to molecular level localization within the cell. For example, the 698 nm component is due primarily to photosystem I (PS I)-related chlorophyll (it is greatly decreased in mutants lacking PS I or strains depleted in chlorophyll) and is distributed much more evenly through the cell than the phycobilin components, which are concentrated around the periphery. By comparing spectra and distribution of components from the wild type and strains lacking PS I and/or ChlL, an enzyme needed for light-independent conversion between biosynthetic precursors of chlorophyll, individual fluorescence components can be assigned to specific emitting pigments.

The results obtained indicate a preferential localization of fluorescing phycobilin components around the periphery of the cells, whereas chlorophyll emission is more evenly distributed within the cell. Based on these and other results, our current explanation is that unattached, highly fluorescent phycobilisomes (\textit{i.e.}, phycobilisomes that do not transfer energy to photosynthetic reaction center complexes) are toward the periphery of the concentric stack of thylakoids in the cell, whereas photosynthetic reaction center complexes are more evenly distributed among thylakoids.

We show here that the hyperspectral confocal imaging approach provides highly detailed information regarding sub-cellular localization of pigments in living cells with unparalleled resolution. We anticipate to be able to further develop and refine the

\* Presenting author
technique and deconvolutions to be able to visualize and follow a large number of fluorescing cellular components in the cell.

References

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

The Cyanobacterium *Synechocystis* sp. PCC 6803: Membrane Biogenesis, Structure and Function

Wim Vermaas* (wim@asu.edu), Robert W. Roberson, Sawsan Hamad, Bing Wang, Zhi Cai, and Allison van de Meene

Arizona State University, Tempe, AZ

Cyanobacteria play a key role in global carbon fixation and energy conversion, and selected strains lend themselves very well to metabolic engineering and synthetic biology. One strain that is particularly useful in this respect is the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 that is readily amenable to detailed structural and functional investigations as it has a known genome sequence, is easily transformed, can grow under a wide variety of environmental conditions, and has developed into the premier cyanobacterial model system for photosynthesis and respiration studies. A major advantage of *Synechocystis* is its ready availability of informative mutants with knock-outs of one or more (up to seven) genes and/or with overexpression of introduced genes. As *Synechocystis* was the first cyanobacterium to be sequenced, another major advantage of this organism is the abundance of bioinformatics resources geared toward this cyanobacterium (e.g., CyanoSeed developed with Genomics:GTL funding [http://theseed.uchicago.edu/FIG/organisms.cgi?show=cyano] and CyanoBase, [http://www.kazusa.or.jp/cyanobase/Synechocystis/index.html]). In our GTL-supported studies we focus on cell structure and cell physiology in *Synechocystis*, with particular emphasis on thylakoid membrane formation and on metabolism related to photosynthesis and respiration. New results on (a) thylakoid membrane biogenesis, (b) fluxes through central carbon utilization pathways, and (c) distribution mechanisms between carbon storage compounds will be presented in subsequent paragraphs. Together, these results help pave the way for metabolic engineering efforts resulting in improved bioenergy production and carbon sequestration.

The internal thylakoid membrane system of *Synechocystis* comprises about 80% of the total membrane content of the cell, and contains membrane protein complexes involved in both photosynthesis and respiration. Thylakoid organization is rather sophisticated, with membranes occurring in multiple layers along the periphery of the cell while often connected to a rod-like structure, the thylakoid center. Thylakoid formation seems to be critically correlated with the presence of chlorophyll and not with the presence of photosynthetic complexes, as in a mutant where chlorophyll synthesis is under strict light control thylakoids are essentially absent after prolonged growth in virtual darkness, whereas thylakoid membranes form rapidly upon exposure to light. In contrast, mutants lacking both photosystems retain a significant amount of thylakoids. The molecular mechanism
of thylakoid formation remains largely unknown, but some proteins that may be involved with thylakoid generation have been identified. Upon overexpression of one of these proteins we found more and closer spaced thylakoids protein in *Synechocystis*, along with novel membrane structures that may be instructive in understanding membrane biogenesis. These results indicate that we now are able to generate cyanobacterial strains with increased levels of the photosynthetic apparatus and thylakoids (useful biomass) per cell.

Metabolic engineering is aided by detailed insight into fluxes through main metabolic pathways. Thus far, flux data regarding central carbon utilization pathways are derived largely from a rather indirect approach of isotope labeling and monitoring the isotopic composition of end products such as amino acids. However, cyanobacterial carbon utilization is very complex, with sugar utilization by the pentose phosphate pathway and glycolysis immediately connected with carbon fixation pathways through the Calvin-Benson-Bassham cycle. Most steps in the pathways are reversible, and several steps involve the reorganization of C-C bonds, causing rapid isotope scrambling when providing uniformly labeled $^{13}$C-glucose. With improved LC/MS (liquid chromatography/mass spectrometry) techniques we have now been successful in monitoring the mass redistribution of several sugar phosphates as a function of time after adding labeled glucose. Comparing these results under different growth conditions and in specific mutants lacking particular steps in the pathways yields a detailed insight regarding *in vivo* rates of key metabolic reactions in carbon utilization.

*Synechocystis* sp. PCC 6803 has two main carbon storage compounds: glycogen and polyhydroxybutyrate (PHB). The compound to be accumulated has been found to primarily depend on the environmental conditions. Glycogen is found under many conditions, but under –for example– nitrate limitation PHB can make up 5–10% of the dry weight of the cell. We have explored the metabolic reasons for this apparent dichotomy in the preferred carbon storage compound in *Synechocystis*.

Based on results of comparative PHB accumulation analysis as a function of environmental conditions in different mutant strains, the level of reduction of specific redox carriers in the cell is found to be a key determinant for PHB accumulation. With this knowledge PHB production in *Synechocystis* can now be optimized.

These aspects of the project build on an excellent foundation of genomic and functional data regarding *Synechocystis* sp. PCC 6803, and together provide a solid basis for metabolic engineering of this cyanobacterium to enhance solar-powered carbon sequestration and bioenergy conversion.

**Reference**