Application of high resolution proteomics to characterize microbial systems for metal reduction and photosynthesis

Samuel O. Purvine, Mary S. Lipton, Margie F. Romine, Stephen J. Callister, Kim K. Hixson, Angela D. Norbeck, Matthew E. Monroe, Carrie D. Nicora, LeeAnn McCue, Joshua N. Adkins, Richard D. Smith, and Jim K. Fredrickson
Pacific Northwest National Laboratory, Richland, WA 99352

Overview

❖ The AMT tag based proteomics approach has enabled the rapid characterization of large numbers of proteins directly from cell lysates. The basis of this technology is the detection of distinctive peptides for each expressed protein which are then used to determine the relative abundances of these proteins in cellular cultures.

❖ We have applied this technology to the characterization of three microbial systems *Shewanella oneidensis*, *Geobacter sulfurreducens* and *Rhodobacter sphaeroides*.

❖ Exploiting microbial function for purposes of bioremediation, energy production, and carbon sequestration 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 correctly annotate a genome, determine protein localization and quantitatively measure changes in protein abundance between culture conditions.

Introduction

Microbial systems dominate every facet of life and outnumber eukaryotic species by orders of magnitude. These diverse organisms are responsible for most of the biology on the earth and understanding that biology allows the manipulation of microbial processes for uses in bioremediation, carbon sequestration and energy production. Proteins play a central role in the functionality of these organisms and understanding protein complements and expression patterns will be critical to this undertaking. Utilizing the AMT tag approach, three microbes important to environmental remediation and energy production have been studied.

Shewanella oneidensis

- Respiratory generalist and facultative anaerobe
- Potential applications in the bioremediation of metals
- Can function to enzymatically reduce and precipitate a diverse range of heavy metals and radionuclides
- A thorough understanding of how MR-1 responds to changes in electron acceptor type and concentration and the enzymatic pathways involved in these reactions is critical for effectively using metal-reducing bacteria for remediation

Rhodobacter sphaeroides

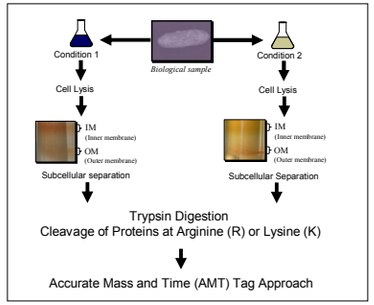
- Gram negative purple nonsulfur eubacterium
- Member of the a-3 subdivision of the *Proteobacteria*
- Within this subgroup, *Rhodobacter sphaeroides* is a well-studied and free-living microbe that is capable of growth under a variety of conditions, e.g.,autotrophic, phototrophic, and anaerobic heterotrophic, utilizing a variety of electron acceptors such as dimethly sulfoxide (DMSO), trimethyl-amine-N-oxide (TMA) and nitrate, among others
- *R. sphaeroides* is known for its diverse abilities, e.g., metal reduction, nitrogen fixation, and the sequestration of carbon dioxide
- Capable of producing copious amounts of hydrogen, potentially providing a source of renewable bioproduced energy
- Many of these abilities are linked to *R. sphaeroides'* photosynthetic apparatus which has served as a model for anoxygenic photosynthesis

Geobacter sulfurreducens

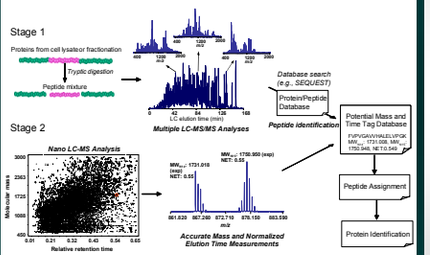
- Dissimilatory metal reducing bacteria that has the potential to reduce U(VI) to U(IV) in ground water
- Effective at reducing a wide range of soluble and solid phase metals during anaerobic respiration
- *Geobacter* appears to be common among contaminated ground water environments and is thus relevant to *in situ* bioremediation in such areas

Methods

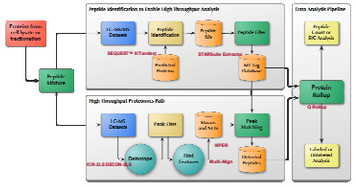
Experimental Design



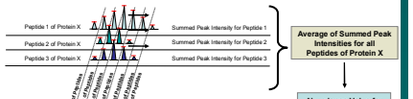
Accurate Mass and Time (AMT) Tag Approach



Data Processing and Management system (PRISM)

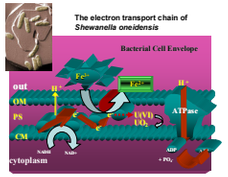


Approach to Calculation of Protein Abundance Values



Results

Validation of Coding Sequence (CDS) Predictions and Mature Protein N-termini in *Shewanella oneidensis* MR-1



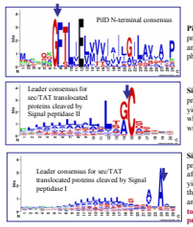
The genome sequence of *S. oneidensis* MR-1 was first released in 2002 and predicted to encode 4,931 protein-encoding open reading frames (CDS) by the Institute for Genome Research (TIGR). Of this total, 2032 were annotated as hypothetical proteins, 1161 of which were unique as well as 871 that were also similar to predicted proteins in other bacteria. In 2003, Daraselia, et. al. used an alternative ORF calling strategy that resulted in removal of 415 of the original hypothetical predictions and addition of 34 new ones. These combined CDS predictions have been manually reevaluated using the following information:

- Overlap with adjacent genes, promoters, enhancer elements, attenuators, and RNAs
- Overlap with IS elements and other repeats
- Extent of sequence similarity to CDS in other genomes as well as proteins deduced from genome sequences of 10 additional *Shewanella* species
- Occurrence of an upstream Shine-Dalgarno site
- Identification of peptides in tryptic digests of cellular lysates
- Identification of peptides that match predicted N-terminal peptides

The proteome data has proved useful for both validation of hypothetical CDS predictions and determination of the N-termini of nearly one quarter of the revised list of CDS encoded by this genome. **A summary of these analyses are presented here.**

Current Status of CDS predictions in the *S. oneidensis* MR-1 Genome

- 537 hypothetical CDS predictions deleted
- 211 other CDS predictions removed
- 35 new CDS predictions added
- 4183 total intact CDS
- 210 total pseudogenes
- 461 CDS start sites changed
- 200 out of 659 hypothetical CDS validated with proteome data
- 384 out of 797 conserved hypothetical CDS validated with proteome data



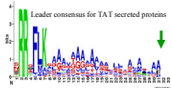
PilD is predicted to cleave 21 proteins after glycine leaving an N-terminal methylated phenylalanine.

Signal peptidase II (LspA) is predicted to cleave 148 proteins yielding an N-terminal cysteine which is then further modified with a lipoyl group.

Signal peptidase I (LepB) is predicted to cleave 680 proteins after a conserved ala-X-ala motif yielding an unmodified N-terminus that can be detected by proteome analysis. **Peptides corresponding to the N-termini of 130 of these proteins have been identified.**

Map is predicted to remove the N-terminal methionine from 1369 proteins with a G, A, S, T, or C at the N+2 position. **N-terminal peptides were detected for 258 of these proteins.** N-terminal peptides for an additional 5 proteins with an asparagine at N+2 were also detected. No N-terminal peptides were detected supporting cleavage at valines found at position+2 and consequently none of these proteins are included in this count.

AmpP is predicted to remove the N-terminal methionine from 170 proteins with a proline at position N+2. **N-terminal peptides were detected for 22 of these proteins.**

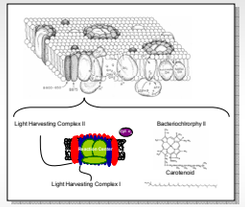


Unlike the Sec secretion pathway, the **twin arginine system (TAT)** secretes folded proteins across the inner membrane. 30 proteins with the characteristic TAT leader were identified in MR-1. **N-terminal peptides were identified for 6 of these proteins validating PetA ubiquinol-cytochrome c reductase iron-sulfur subunit (SO0608) cleavage by Map rather than LepB, as has been seen in other bacteria, and 4 of the remaining 5 proteins have leaders longer than is typical of Sec leaders (53-68 aa).**

A total of 721 protein N-termini were validated by proteome analysis! Future studies aim to identify N-terminal peptides associated with N-terminally modified lipoproteins and PilD substrates identified here.

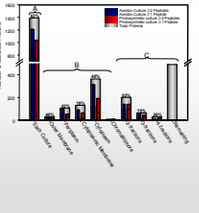
Elias, D. A., et al. Global detection and characterization of hypothetical proteins in *Shewanella oneidensis* MR-1 using LC-MS based proteomics. Proteomics. **2005**. 5:3120-3130.
Kolker, E., et. al. Global Profiling of *Shewanella oneidensis* MR-1: Expression of hypothetical genes and improved functional annotations. Proc. Natl. Acad. Sci. U. S. A. **2005**. 102(6):2099-2104.
Romine, M. F., et. al. Validation of *Shewanella oneidensis* MR-1 small proteins by AMT tag-based proteome analysis. OMICS. **2004**. 8(3):239-254.

Determination of Protein Localization in *Rhodobacter sphaerodes*

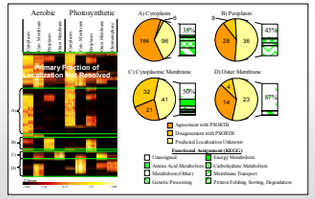


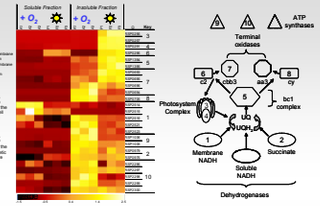
The Intracytoplasmic Cell Membrane (ICM) contains the photosystem complex and additional components necessary to capture electrons, transporting them through the electron transport chain driving the formation of a proton gradient required for ATP synthesis. Electrons are transported through the light harvesting complexes to the reaction center which consists of three subunits

Global Characterization of *Rhodobacter* Proteins. ~8300 high quality peptides were detected from 54 LC-FTICR MS analyses, which corresponded to 1391 proteins or 32% coverage of predicted proteins. 62% of the total proteins identified by 2 or more peptides were observed in both cell cultures (**A**). Only 19% of proteins identified in the cytoplasmic membrane were common to both cell cultures (**B**). This generally correlates to the different proteins utilized for energy gathering and conversion. Proteins detected in 2-, 3-, and 4-fractions (**C**) were identical for both cell cultures. A majority of proteins identified in multiple sub-cellular fractions were not consistent in terms of the fractions identified. Proteins observed in multiple-fractions are possibly 1) those that span, or 2) protrude into adjacent fractions, or 3) were observed in multiple fractions due to purity issues related to each sub-cellular fraction.



Localization of *Rhodobacter* Proteins To determine the most likely fraction of localization for proteins observed in multiple sub-cellular fractions, *Z*-scores were calculated from protein abundances across a given cell culture condition and clustered using hierarchical algorithms. **A)** ***Localization of proteins to the cytoplasm.*** 266 proteins were resolved, with a majority of them showing consistent results for both cell cultures. 164 out of 170 proteins with predicted localization (PSORTB) agreed with observed results. Predicted localization for 96 proteins could not be determined. The largest fraction of these had an unassigned function based on KEGG functional categories (Pie Chart). **B)** ***Localization of proteins to the periplasm.*** 73 proteins were resolved, with 28 of 37 proteins in agreement with predicted localization. Roughly 50% of proteins observed in the periplasm had no predicted localization information (Pie Chart). **C)** ***Localization of proteins to the cytoplasmic membrane.*** 96 proteins were resolved. However, for this sub-cellular fraction disagreement between observed and predicted results was greater than agreement between observed and predicted results (Pie Chart). **D)** ***Localization of proteins to the outer membrane.*** 41 proteins resolved. 23 of these proteins had no predicted localization information. Within this set of proteins almost all had an unassigned function which suggested that the protein of the outer membrane is relatively uncharacterized.

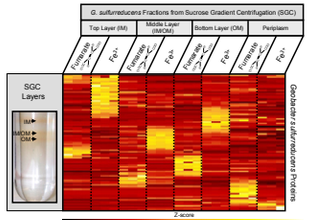




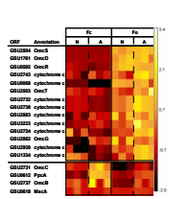
Predicted localization for 437 proteins (196 from proteins observed in multiple sub-cellular fractions + 241 proteins observed in a single sub-cellular fraction) could not be determined by the applied algorithms. Although annotations and functional assignments for these proteins provided some localization evidence, many proteins were annotated as hypothetical/conserved hypothetical, putative, or had an unassigned function. A sample of these proteins are listed along with their observed localization and the relevant culture condition. Each protein listed here was identified by a minimum of 2 peptides. In parentheses are the percents of protein amino acid sequences covered by detected peptides. The sub-cellular fraction assigned to some proteins corresponded with proteins uniquely detected within a given cell culture. However for this study, a higher confidence was given to sub-cellular fraction assignments where observations were consistent for both cell cultures. A next step is to evaluate the relative abundance of these proteins in the context of aerobic and photosynthetic cell states.

Callister, S. J. et al. **2006.** Comparison of aerobic and photosynthetic *Rhodobacter sphaeroides* 2.4.1 proteomes. *Journal of Microbiological Methods* (in press).
Callister, S. J., et al. **2006.** Application of the accurate mass and time tag approach to the proteome analysis of sub-cellular fractions obtained from *Rhodobacter sphaeroides* 2.4.1 aerobic and photosynthetic cell cultures. *Journal of Proteome Research* (in press).
Zeng, X., et al. **2006.** Proteomic and transcriptional characterization of intracytoplasmic membranes purified from *Rhodobacter sphaeroides* 2.4.1. (*In preparation*).
Callister, S. J., et al. **2006.** Proteome comparison of transitional cell states from aerobic and photosynthetic *Rhodobacter sphaeroides* 2.4.1. (*In preparation*).

Quantitative Characterization of the Proteomic Response of *Geobacter sulfurreducens* during growth on Fe(III) oxides

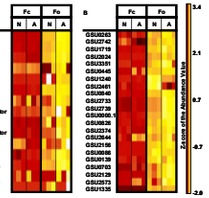


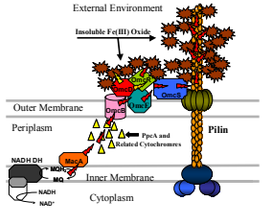
Relative abundance data plot for 91 cytochromes with 1 or more unique peptides identified per O/RF from *G. sulfurreducens*. Cell fraction preparation (inner membrane, hybrids of inner and outer membrane, outer membrane and periplasmic space) from *G. sulfurreducens* grown in either fumarate or soluble Fe(III) reduction in batch condition is also shown side by side. The darkest colored data represent proteins that have a significantly lower abundance than average. The lighter color represents higher relative z-score and thus a higher relative abundance for each observed protein. A z-score of 0 represents a protein abundance value that is equal to the average abundance value of all abundances for that given protein in the eight culture conditions.



c-type cytochromes showing significant changes in the Fe(III) oxide reduction condition compared to soluble Fe(III) reduction conditions. The darkest colored data represent proteins that have a significantly (z-score ≥ 1.96) lower abundance than average. The lighter color data are proteins with a significantly (z-score ≤ 1.96) higher abundance from average. Fc: cells grown in late-log phase during Fe (III) citrate reduction; Fo: cells grown during Fe (III) oxide reduction; A: protein samples alkylated with iodoacetamide; N: protein samples not alkylated.

Cluster Analysis of *Geobacter* grown on Fe(III): Selected signal transduction and regulatory proteins (A) and hypothetical proteins (B) were detected by the AMT tag approach and clustered between the Fe(III) oxide reduction and soluble Fe(III) reduction conditions. The darkest colored data represent proteins that have a significantly (z-score ≥ 1.96) lower abundance than average. The lighter colored data represent proteins with a significantly (z-score ≤ 1.96) higher abundance from average. Fc: cells grown in late-log phase during Fe (III) citrate reduction; Fo: cells grown during Fe (III) oxide reduction; A: protein samples alkylated with iodoacetamide; N: protein samples not alkylated.





A proposed model of electron transfer for insoluble Fe(III) reduction based on proteins with significant expression found specifically with growth on Fe (III) oxides.

Ding Y-HR, et al. **2006.** "The Proteome of Dissimilatory Metal-Reducing Microorganism Geobacter sulfurreducens Under Various Environmental Conditions", submitted to *BBA Proteins and Proteomics*.
Ding Y-HR, et al. **2006.** "Proteomic and functional genomic evaluation of proteins differentially displayed between insoluble and soluble Fe(III) reduction from *G. sulfurreducens*", submitted to *J. of Bacteriology*.

Conclusions

❖ The use of the proteome data to adjust the number and start position of predicted proteins encoded by MR-1 has enabled us to produce a more accurate prediction of operon structure and protein localization. This information in turn will be especially useful for subsequent differential expression studies where these predictions impact interpretation of both microarray and proteome data as well as in prediction of functions of expressed proteins.

❖ Insights into the cellular locations of proteins have been effectively obtained via the AMT tag approach as applied to *Rhodobacter sphaeroides*, leading to a deeper understanding of the dynamic cellular environment and ultimately towards a systems level picture of the living organism.

❖ The high throughput nature of the AMT tag approach has led to a better understanding of the electron transport pathway in *Geobacter sulfurreducens*, an important organism for biogeochemical remediation and environmental improvements.

Acknowledgements

This research was supported by the U. S. Department of Energy (DOE) Office of Biological and Environmental Research Genomics:GTL program. We greatly appreciate the contributions from our collaborators:

Shewanella oneidensis: Frank Larimer, Chris Hemme (Oakridge National Laboratory), Sean Conlan (Wadsworth Institute), and Gretta Serres (Woods Hole Marine Biological Laboratory)

Rhodobacter sphaeroides: Timothy Donohue, Miguel Dominguez, Christine Tavano (University of Wisconsin-Madison), Samuel Kaplan, Xiaihua Zeng, and Jung Hyeob Roh (UT-Houston Medical School)

Geobacter sulfurreducens: Derek Lovley and Richard Ding (University of Massachusetts)

Experimental portions of this research at Pacific Northwest National Laboratory (PNNL) were performed in the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy (DOE) national scientific user facility at PNNL in Richland, Washington. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under Contract No. DE-AC05-76RLO 1830.

Contact Information

Samuel O. Purvine
Instrument Development Laboratory, K8-91
Pacific Northwest National Laboratory
P.O. Box 999, Richland, WA 99352
e-mail: samuel.purvine@pnl.gov
Phone: (509) 376-3013