



Enhanced Protein Extraction for Microbial (Meta)Proteomics of Defined Laboratory and Environmental Samples

Xiaoxin Liu¹, Karuna Chourey², Silke Nissen³, Abigail Green⁴, Shulei Sun⁵, Stephanie Connon⁴, Victoria Orphan⁴, Frank Löffler³, Robert Hettich^{1,2}
¹UT – ORNL Genome Science & Technology, Knoxville, TN; ²Oak Ridge Natl. Lab, Oak Ridge, TN; ³University of Tennessee, Knoxville, TN; ⁴Caltech, Pasadena, CA; ⁵UCSD, San Diego, CA.



ABSTRACT

We evaluated the efficacy of a detergent - based protein extraction procedure to characterize microbial proteomes from bacterial cultures and environmental samples. This detergent-based protocol was used to measure the proteome profile of *Anaeromyxobacter dehalogenans* strain 2CP-C, with emphasis on a comprehensive coverage of c-type cytochromes expressed under different growth conditions. This work aims at identifying redox-active proteins expressed when strain 2CP-C is grown different electron acceptors. This lysis method was further extended for the *in-situ* protein extraction from methane seep sediments with the goal to identify microbial biomarkers for methane turnover.

EXPERIMENTAL APPROACH

A. dehalogenans was grown with different electron acceptors (Figure 4) and cells were harvested for proteomic analysis. For microbial carbon cycling studies, sediment cores were collected from active methane seeps of Eel River basin, Santa Monica basin and Hydrate Ridge, CA and 10 g of sediment was used for *in-situ* proteome extraction. The samples (cell pellet or sediment) were subjected to detergent-based cell lysis at elevated temperatures followed by acid precipitation of proteins. Following trypsin proteolysis, solid-phase extraction and solvent exchange, the peptides were analyzed with via 2-D-LC-MS/MS on a linear ion trap mass spectrometer (LTQ XL) or a dual pressure Linear Ion trap (LTQ Velos). All MS/MS spectra were searched against genome/metagenome databases.

In Situ Soil Protein Extraction Methodology (SDS-TCA)

Soil dispersed in a detergent based lysis buffer and boiled in water bath for 20 min

↓ brief low speed centrifugation, discard soil, retain supernatant

Supernatant + chilled 100% TCA (to final 25% concentration)

↓ Overnight incubation at 4°C followed by 20 min. centrifugation at top speed to pellet protein, discard supernatant

Wash protein pellet with chilled acetone (twice), dry pellet and solubilized in Guanidine-DTT solution.

↓ Incubate at 60°C for an hour, dilute six fold with Tris-CaCl₂ buffer

Overnight proteolytic digestion (Trypsin) at 37°C followed by denaturation (DTT), desalting and solvent exchange

↓ Peptides interrogated via 24h, 12-step, Liquid Chromatography coupled MS/MS

↓ Datasets analyzed using SEQUEST

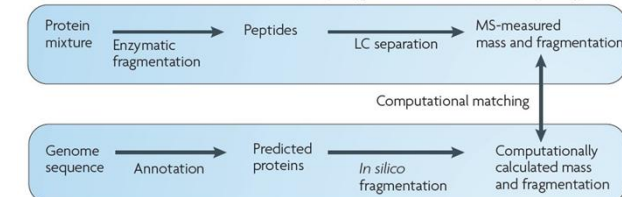
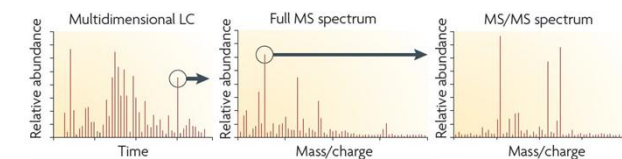
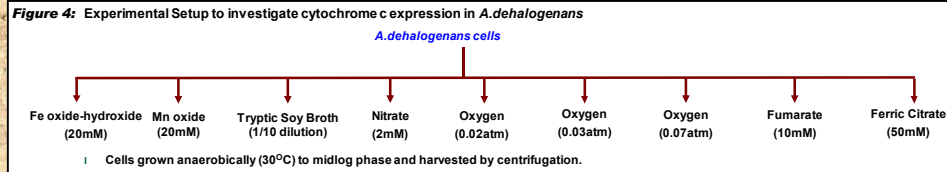


Figure 1. Liquid chromatography-mass spectrometry-based proteomics. Protein mixtures are enzymatically digested into peptides. Then the generated small peptides are sent to liquid chromatography(LC) for separation. Peptides are separated by hydrophobicity and electric charge. Separated peptide are measured by tandem mass spectrometry(MS/MS) generating spectrum providing information of mass to charge ratio and abundance about the peptides. The experimental results are compared with spectra generated computationally from genome sequence of the sample.

RESULTS FOR Anaeromyxobacter dehalogenans STRAIN 2CP-C PROTEOMICS



Figure 2: *Anaeromyxobacter dehalogenans* is a Gram-negative rod-shaped, motile, spore-forming bacteria found in the soil. (Image source: JGI)



- Cells grown anaerobically (30°C) to midlog phase and harvested by centrifugation.
- Cell pellet subjected to heat assisted detergent based lysis to extract whole cell proteome.
- Protein proteolysis was achieved using Trypsin, followed by peptide clean up and solvent exchange.
- Peptides were analyzed via 12-step LC-MS/MS using LTQXL mass spectrometer.
- Database searches were done using SEQUEST against the isolate genome database.

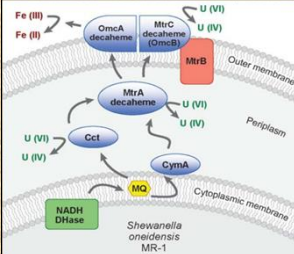
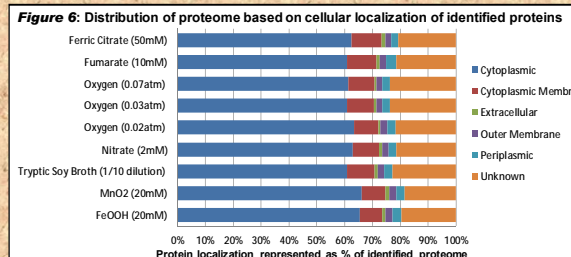


Figure 3: A model for possible electron transport pathways for U(VI) reduction in *Shewanella oneidensis* MR-1. (Wall and Krumholz 2006)

Table 1: Summary table of proteome profile of *A. dehalogenans* grown with different electron acceptors

Electron Acceptors	Proteins IDs	Total Spectral Count	Number of Cytochrome C
FeOOH (20mM)	889	14136	10
MnO ₂ (20mM)	1044	27412	16
Tryptic Soy Broth (1/10 dilution)	1691	72230	23
Nitrate (2mM)	1544	62327	24
Oxygen (0.02atm)	1647	73141	21
Oxygen (0.03atm)	1687	81899	23
Oxygen (0.07atm)	1617	76783	20
Fumarate (10mM)	1543	62516	29
Ferric Citrate (50mM)	1282	61219	20



- Under each electron acceptor amended growth conditions, the number of identified proteins ranges from 889 to 1687 (Table 1).
- Out of the 69 predicted cytochrome c proteins from the genome analysis, we detected expression of 48 cytochromes among all growth conditions.
- 10 to 29 cytochrome c expressions were detected for each growth conditions (Table 1).
- Identified proteins represent different cellular locations within cell indicating efficient cellular lysis, protein extraction and good representation of all sites in a bacterial cell (Figure 6).
- The most abundant proteins are from 'energy production and conversion' and 'amino acid transport and metabolism' (Figure 7). There is a large number of hypothetical proteins that have no known function at present and belong to 'none' category.

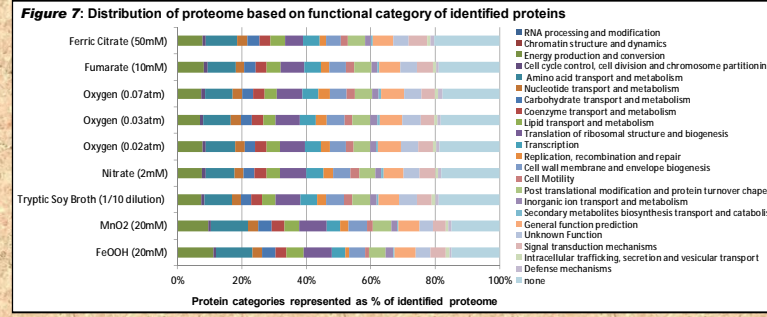


Table 2: Summary table of common (red) and unique (green) cytochromes detected in *A. dehalogenans* grown with different electron acceptors. Numbers in cells represent spectral counts for different c-type cytochromes.

Chromosome locus ID	GenBank Annotation	Predicted C-type cytochrome motifs	Cellular Localization	FeOOH (20mM)	MnO ₂ (20mM)	Tryptic Soy Broth (1/10 dilution)	Nitrate (2mM)	Oxygen (0.02atm)	Oxygen (0.03atm)	Oxygen (0.07atm)	Fumarate (10mM)	Ferric Citrate (50mM)
Adch_0795	quinol:cytochrome c oxidoreductase pentaheme cytochrome subunit	5 motifs	Unknown									
Adch_0799	quinol:cytochrome c oxidoreductase monoheme COG2010_CocA	1 motif	Periplasmic									
Adch_0909	hypothetical protein	4 motifs	Periplasmic									
Adch_0910	NrFA Nitrite reductase (cytochrome: ammonia-forming)	5 motifs	Periplasmic									
Adch_0918	hypothetical protein	1 motif	Periplasmic									
Adch_1278	cytochrome c oxidase, mono-heme subunit/FixO	2 motifs	Periplasmic	10	19	44	53	20	16	5	44	48
Adch_1891	hypothetical protein	3 motifs	Unknown									
Adch_2004	hypothetical protein	3 motifs	Periplasmic									
Adch_2169	hypothetical protein preceds Nitrate reductase alpha subunit	1 motif	Unknown									
Adch_2272	hypothetical protein	10 motifs	Extracellular	2	24	80	58	42	16	7	20	91
Adch_2621	cytochrome c class I	1 motif	Periplasmic									
Adch_2621	hypothetical protein	2 motifs	Unknown									
Adch_2682	hypothetical protein cytochrome c-554 homolog	4 motifs	Unknown									
Adch_2918	NrFA Nitrite reductase (cytochrome: ammonia-forming) EC:1.7.2.2	5 motifs	Periplasmic	51	25	57	51	37	50	45	124	59
Adch_2960	hypothetical protein	2 motifs	Unknown									
Adch_2960	putative lipoprotein	1 motif	Periplasmic									
Adch_3065	cytochrome c family protein	6 motifs	Unknown									
Adch_3090	hypothetical protein	8 motifs	Periplasmic									

SUMMARY

- A total of 2521 proteins (out of the annotated 4485 ORFs) were identified across all growth conditions, leading to 56.2% coverage of *A. dehalogenans* genome.
- There were 512 proteins expressed under all treatments, and likely constitute the core proteome of *A. dehalogenans*.
- Cytochrome c expression was highest in fumarate grown cells and lowest in Fe oxide-hydroxide and Mn oxide amended cultures
- Three cytochromes (*Adch 1172*, *Adch 2216* and *Adch 2902* (NrFA)) were common to all growth conditions and are likely to be constitutively expressed by the microbe irrespective of the type of electron acceptor in the media.

RESULTS FOR DEEP SEA METHANE SEEP SEDIMENT MICROBIAL PROTEOMICS



Figure 8: Image of deep sea methane seep.

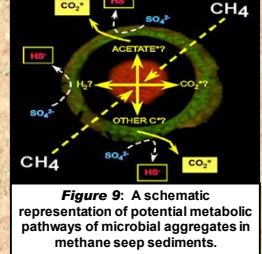


Figure 9: A schematic representation of potential metabolic pathways of microbial aggregates in methane seep sediments.

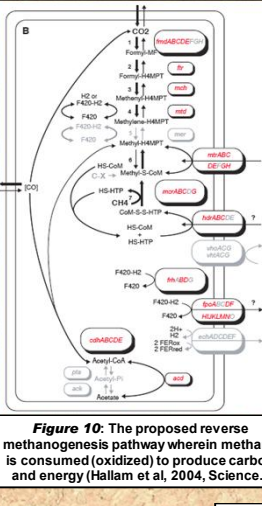


Figure 10: The proposed reverse methanogenesis pathway wherein methane is consumed (oxidized) to produce carbon and energy (Hallam et al., 2004, Science.)

Table 3: Representative proteins shared between all methane seep samples analyzed and microbes with closest gene similarities indicating core metabolic pathways and potential species of methane seep environments.

Microbes common to all Methane Seep sediments*	Proteins identified in all sediment samples*
<i>Thiobacillus denitrificans</i> ATCC 25259	Dissimilatory adenylylsulfate reductase subunit alpha
<i>Thiomicrospira crunogena</i> XCL-2	Adenylylsulfate reductase
<i>Syntrophobacter fumaroxidans</i> MPOB	Glyceraldehyde-3-phosphate dehydrogenase
<i>Sphingomonas wittichii</i> RW1	Chaperonin GroEL (Hsp60, Cpn10)
<i>Pseudoaeromonas atlantica</i> T6c	Membrane protein involved in aromatic hydrocarbon degradation
<i>Planctomycetes limophilus</i> DSM 3776	Glutamine synthetase Type I
<i>Pirellula staleyii</i> DSM 6068	FOF1 ATP synthase subunit beta
<i>Methylophilus trichosporium</i> OJB3b	Carbon-monoxide dehydrogenase catalytic subunit
<i>Methylthiobacterium</i> DSM 11000	4Fe-4S Ferredoxin, iron - sulphur binding protein
<i>Magneto-coccus</i> sp. MC-1	Translation elongation factor Tu
<i>Acidobacterium</i> sp. MP5ACTX9	DNA-directed RNA polymerase Beta subunit
<i>Desulfobulbus propionicus</i> DSM 2032	Maltose-binding protein (MalE)
<i>Syntrophobacter fumaroxidans</i> MPOB	Alkaline Phosphatase (AprA)
<i>Desulfotalea psychrophila</i> L5v54	Chaperone Protein DnaK
<i>Desulfococcus oleovorans</i> Hxd3	5,10-methylene tetrahydropteroyltransferase
<i>Desulfobacterium autotrophicum</i> HRM2, DSM 3382	Methylenetetrahydrodihydropteroyl transferase
<i>Desulfatibaculum alkenivorans</i> AK-01	Methyl-coenzyme M reductase, gamma subunit
<i>Desulfotomaculum acetoxidans</i> DSM 4771	Methyl-viologen-reducing hydrogenase delta subunit
<i>Dehalococcus</i> sp. SPH-1	
<i>Chloroflexus aggregans</i> DSM 9485	

Table 4: Total number of proteins identified in each sample and how they correlate with the proteome profile of other samples. Size of proteins identified ranged from 411 kDa to 8 kDa.

Sample	Protein IDs	Santa Monica basin 0-4 cm	Eel River basin 0-10 cm	Hydrate Ridge 0-6 cm	Santa Monica basin 8-12 cm	Eel River basin 10-20 cm
Santa Monica basin 0-4cm	465	100.0%	24.7%	57.4%	51.8%	2.6%
Eel River basin 0-10cm	160	71.9%	100.0%	61.9%	56.9%	6.2%
Hydrate Ridge 0-6 cm	479	55.7%	20.7%	100.0%	47.8%	2.5%
Santa Monica basin 8-12 cm	387	62.3%	23.5%	59.2%	100.0%	4.1%
Eel River basin 10-20 cm	27	44.4%	37.0%	44.4%	59.3%	100.0%

- Sediments sampled from the three methane seep sites show many common proteins (Table 3 left), some of which are well known sulfate reducers and methane oxidizers.
- Table 3 (right) also lists the proteins identified in all samples, which lend support to the proposed reverse methanogenesis pathway (Figure 10).
- Data obtained from bacterial 16S rRNA analysis (Figure 11) matches well with metaproteomics data (Table 3 left).
- Proteomes from shallow subsurface sediments show higher biomass and more proteins than sediments from deeper zone.
- Proteins involved in methane oxidation, sulfate reduction, and CO₂ uptake were detected in all sampled sites, although each sediment had its own unique proteome (Table 4).

SUMMARY

- A total of 1014 non-redundant microbial proteins were identified in the sediments of three distinct methane seeps.
- The size of identified proteins ranged from 8 - 411 kDa, showing efficient extraction of a wide range of proteins from these sediments (which include humics and other interfering substances).
- The proteomic profiles of methane seep microbes suggest an anaerobic lifestyle utilizing CO₂ (RuBisCo), methane (Methyl-coenzyme M reductase), and sulfide (coupled with sulfate reduction by sulfate reducing bacteria (Adenylylsulfate reductase)). All these proteins match up to the predicted metabolism for chemoheterotrophic bacteria in cold seeps.
- These results help elucidate pathways of carbon cycling in deep sea methane seep and identify the microbial groups behind these efforts.
- The 16S gene data suggest that methane seep environments are composed of a wide diversity of bacteria that span a large number of different phyla but are dominated by species from within the Proteobacterial phylum. These 16S gene profiles also correlate strongly with the species indicated by the proteomics data.

ACKNOWLEDGMENTS

- Research on carbon cycling in Deep sea Methane Seep is supported by Genomic Science program under U.S. Department of Energy's Office of Biological and Environmental Research.
- Research on proteomics of *A. dehalogenans* is supported by Subsurface Biogeochemical Research (SBR) under U.S. Department of Energy's Office of Biological and Environmental Research.
- Oak Ridge National Laboratory is managed and operated by the University of Tennessee-Battelle, LLC., for U.S. Department of Energy.