Mass Spectrometric Analysis of Protein Complexes from *Rhodopseudomonas palustris*

Gregory B. Hurst[1], Nathan C. VerBerkmoes[3], Trish K. Lankford[2], Dale A. Pelletier[2], Frank W. Larimer[2], Robert L. Hettich[1], Michelle V. Buchanan[1], Michael B. Strader[1], Yisong Wang[2], Linda J. Foote[2], Stephen J. Kennel[2]

[1]Chemical Sciences Division and [2]Life Sciences Division
Oak Ridge National Laboratory
[3]Genome Science and Technology Graduate School
University of Tennessee

Overview

Initial target proteins in *Rhodopseudomonas palustris* have been expressed as fusions with affinity labels to enable isolation of protein complexes. A GroELS complex has been isolated and characterized by MS.

Introduction

- The bacterial species *Rhodopseudomonas palustris*
  - occurs widely in the environment
  - survives in a variety of conditions
    - light / dark
    - aerobic / anaerobic
- This species probably has the potential to express different complements of proteins and protein complexes under different growth conditions.
- As part of a center funded by the U.S. Department of Energy Genomes To Life program [1], we are analyzing protein complexes from *R. palustris* by expressing target proteins as fusions with affinity tags to allow subsequent isolation of other proteins associated with the target [2], followed by both "top-down" and "bottom-up" mass spectrometry analysis.
- For *R. palustris* proteomics measurements, see poster ThPV5 #441(Verberkmoes *et al*.)

Methods

- Selected *R. palustris* genes were cloned with affinity tags and expressed in both *E. coli* and *R. palustris* using modified pDEST vectors (Invitrogen).
  - His6 and GST affinity tags
  - N- and C-terminal positions
- Isolation of fusion proteins
  - affinity purification with Ni-NTA or glutathione-bearing agarose beads
  - Expression confirmed using 1-D PAGE and western blots.
- MS analysis of in-gel digested proteins [2]
  - MALDI-TOF (PerSeptive Biosystems Voyager Elite-DE)
  - LC-MS-MS using a quadrupole ion trap mass spectrometer (LC Packings Ultimate/Famos; Finnigan LCQ)
- "Shotgun" analysis: the entire affinity-isolated sample was analyzed by mass spectrometry [3]
  - "top-down" (with FTICR MS)
  - "bottom-up"
    - trypsin digestion
    - Reverse-phase HPLC separation online with electrospray/quadrupole ion trap MS-MS
    - Protein ID's by comparing tandem mass spectral data with sequence database using Sequest or in-house tools.

Strategy for Identification and Characterization of Protein Complexes

PCR amplification of target gene and cloning into modified expression vector

Transfer expression plasmid into *R. palustris*

Expression of affinity tagged proteins under various physiological conditions

capture protein complexes using affinity tag

Characterize protein complex

mass spectrometry

ID proteins

Stoichiometries

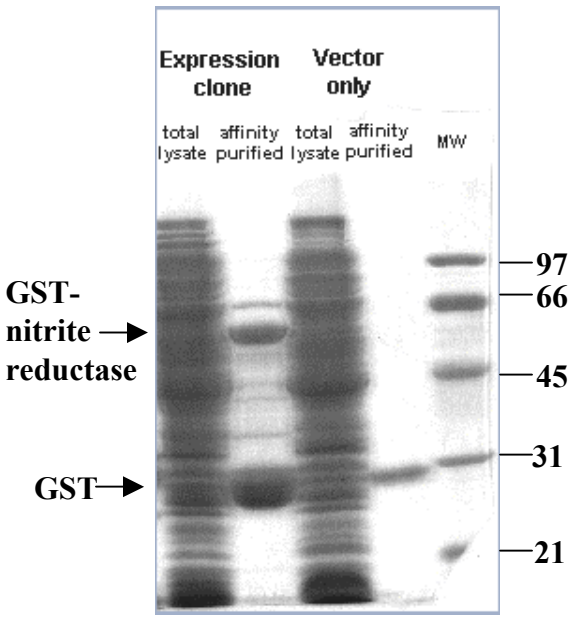
Functional assays

Results: *R. palustris* fusion proteins expressed in *E. coli*

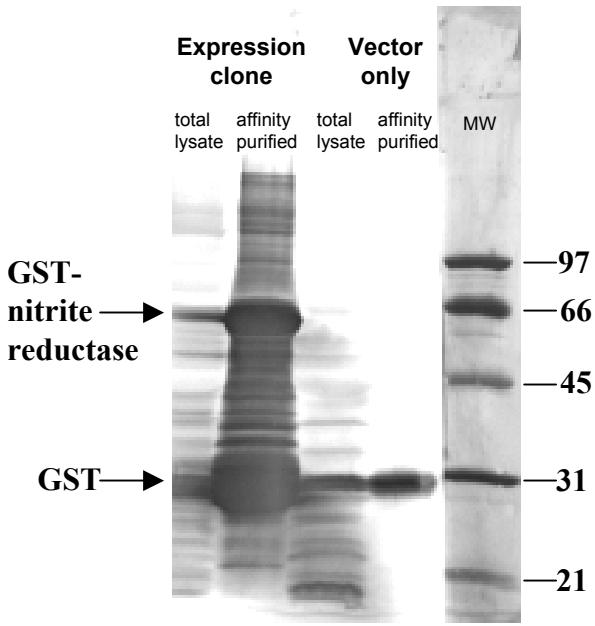
- Expression of the NirK Cu-containing nitrite reductase with an N-terminal GST tag was confirmed (**Figure 1**)
  - western blot to GST
  - MS analysis of in-gel digests of selected bands. Sequence coverage for both the GST and gene product portions of the fusion protein was obtained.
- Three other fusion proteins were verified by digesting the unseparated eluate from the affinity isolation, and analysis by nano-LC-MS-MS. (**Table 1**).
  - Several *E. coli* proteins were also identified.
    - Some contained short runs of histidine residues that could impart affinity toward the Ni-NTA affinity beads.
    - Other *E. coli* proteins presumably isolated due to non-specific interactions with either the affinity medium or the fusion protein.

Figure 1. Verification of affinity tagged *R. palustris* proteins expressed in *E. coli*

Protein stain



Western blot



Mass Spectrometry

GST-nitrite reductase band

	# of peptides identified	
	MALDI	nano LC MS/MS
GST peptides	5	13
GST/linker peptides	2	1
nitrite reductase peptides	9	11
linker/nit. red. peptides	0	1

GST band

GST peptides	10	22
GST/linker peptides	1	1
nitrite reductase peptides	0	0
linker/nit. red. peptides	0	0

Table 1. *R. palustris* fusions in *E. coli*

	No. Of Peptides Identified From:		Other Proteins Identified
Fusion Protein	Target Protein	Affinity Tag	
GroEL2 + N-terminal GST	45	8	0
GroEL2 + C-terminal 6-His, V5 epitope	31	3	19
GroEL1 + C-terminal 6-His, V5 epitope	35	3	8

Results: Native expression of fusion proteins in *R. palustris*

- Plasmids encoding thirteen affinity-tagged fusion proteins were inserted into *R. palustris* (**Table 2**).
- To date, five of these isolated proteins have been analyzed by mass spectrometry and anti-HIS western blotting
- All five proteins showed distinct bands in the western blot
  - **Figure 2**
  - **Figure 3**
- Four of these proteins were successfully identified using LC-MS-MS data, searched against the *R. palustris* database (**Table 3**)
  - Additional proteins were detected, some of which were also observed in a strain not expressing any fusion protein.
  - Use of a second affinity purification step, alternate tags, and optimized conditions will help to eliminate these background proteins.[4]

Figure 2: Anti-His Western Blot of *R. palustris* clones

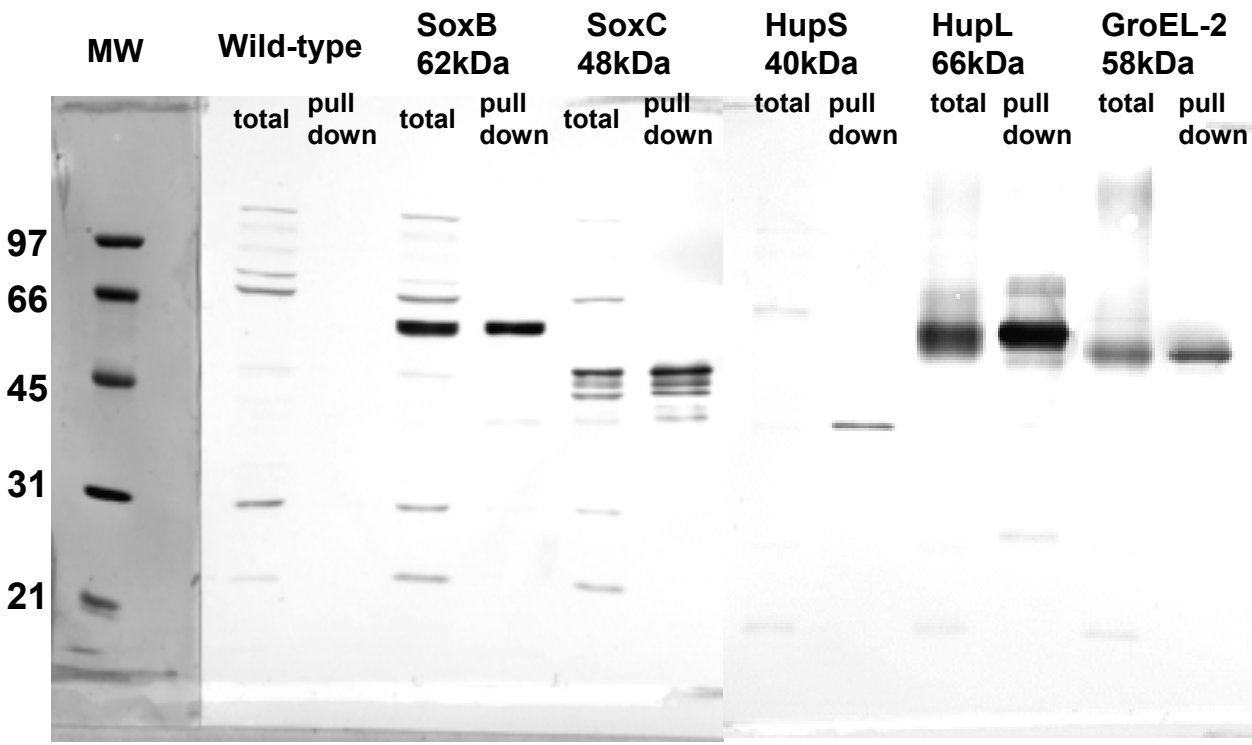
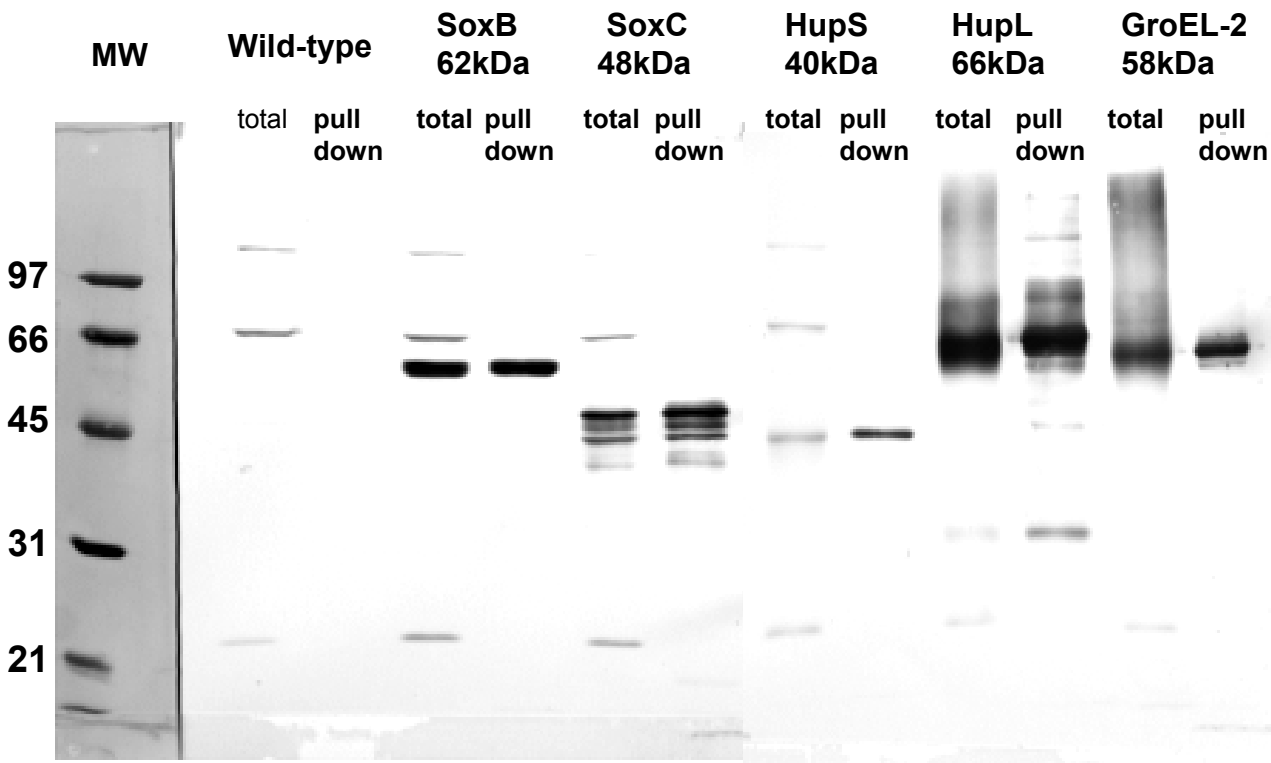
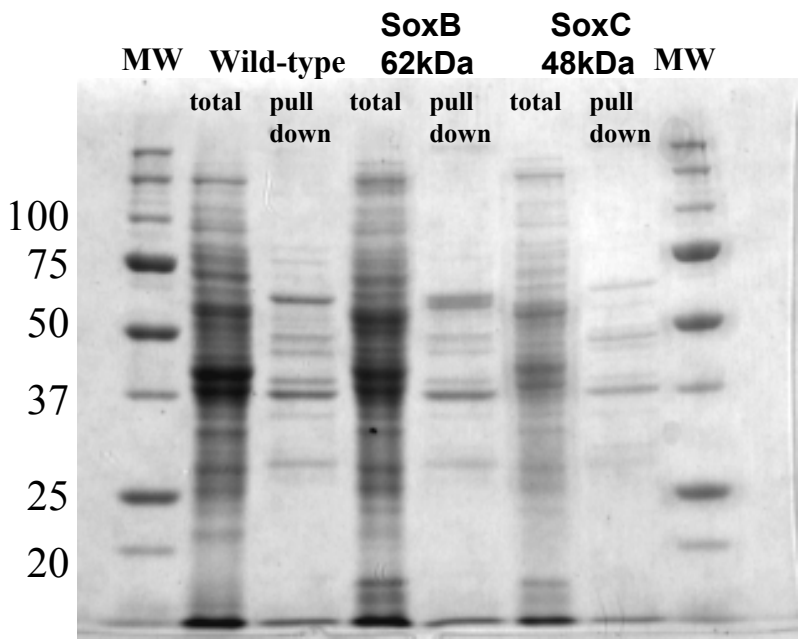


Figure 3: Anti-V5 Western Blot of *R. palustris* clones



Protein gels of *R. palustris* pulldowns



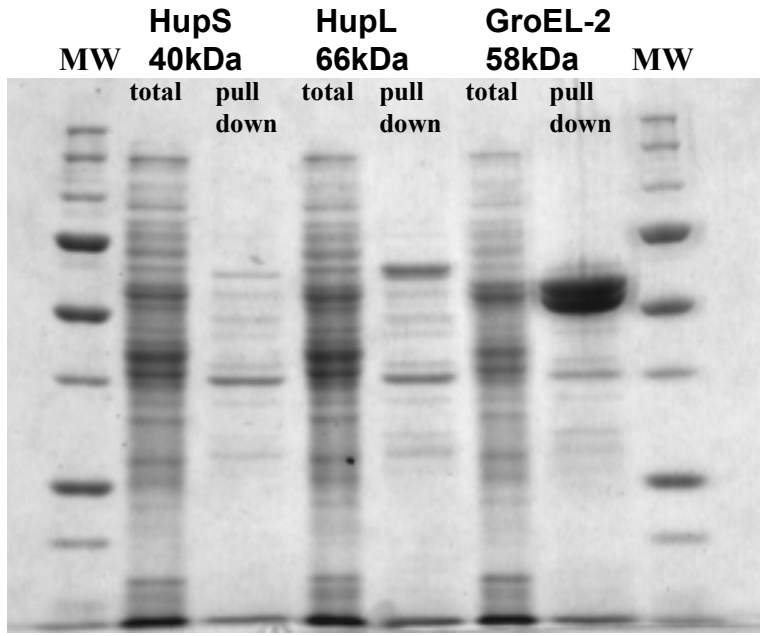


Table 2. *R. palustris* target genes PCR amplified and cloned into destination vectors

gene#:	gene product
or3490	**NirK** Cu-containing nitrite reductase(39.9kDa/368aa)
or4709	**GroEL-2** (57.8kDa/547aa/Pi=5.38) Has ½ of CIRCE element
or4708	**GroES-2** (11.2kDa/104aa)
or5426	**GroEL-1** (57.6kDa/547aa/Pi=5.46) Has CIRCE element
or5425	**GroES-1** (10.6kDa/98aa)
or6580	**SoxB** (61.5kDa/565aa) sulfite dehydrogenase/Mn-dep. hydrolase
or6581	**SoxC** (47.8kDa/431aa) molybdopterin subunit sulfite oxidase
or1645	**HupS** (40.3kDa/370aa) uptake hydrogenase small subunit
or1646	**HupL** (65.7kDa/597aa) uptake hydrogenase large subunit
or3960	**ClpA** (87.5kDa/794aa) endopeptidase
or6448	**NifH** (32kDa/298aa) nitrogenase iron-protein
or6429	**CbbM** (50.5kDa/461aa) RuBisCo FormII
or6072	**FtsY** (54.7kDa/516aa) signal recognition particle

Table 3: Proteins identified in Affinity Isolations (LC-MS-MS)

		Tagged Gene											
Gene#	Identification	HupS		HupL		GroEL-2		SoxB		SoxC		no tag	
		%	UP	%	UP	%	UP	%	UP	%	UP	%	UP
1646	hydrogenase large chain			49	21								
4709	Chaperonin GroEL 2 (HSP60 family)	33	10	28	7	72	34	31	9	31	9	5	1
6580	sulfur/thiosulfate oxidation protein SoxB							42	14				
6581	sulfite dehydrogenase									36	13		
1581	Putative GTPases (G3E family) / (CobW protein)	24	6	24	5	17	3	30	7	15	4	20	6
1754	possible GTP cyclohydrolase II, riboflavin	28	8	40	11	23	5	43	12	26	6	38	11
3623	Dihydroxyacid dehydratase/phosphogluconate dehydratase	44	14	33	13	22	9	31	12	29	12	47	20
4473	Ribonucleases G and E	12	9	17	12	4	3	21	14	12	10	13	9
5913	Fe2+/Zn2+ uptake regulation proteins	17	2	34	4	34	4	22	2	22	2	22	2
6133	F0F1-type ATP synthase beta subunit	24	7	12	4	11	3	11	3	22	7	6	2
6428	fructosebisphosphate aldolase	72	14	62	12	46	9	39	9	32	7	72	13
7323	UDP-N-acetylmuramate-alanine ligase	21	6	17	5	27	7	28	9	27	8	37	11
1809	hypothetical protein			45	3	35	2	45	3	32	2	45	3
5056	hypothetical protein			39	4	24	3	39	4	24	3	66	5
6814	Uncharacterized BCR / conserved hypothetical protein	27	2			27	2	41	3	27	2	54	3
6924	Formyltetrahydrofolate hydrolase /deformylase	40	6	13	2			13	2	20	3	19	3
4140	Ribosomal protein S2	20	3	48	9	33	4			10	2	15	2
5938	Lactate dehydrogenase and related dehydrogenases	18	3	29	5	19	3			13	2	24	4
5426	Chaperonin GroEL 1 (HSP60 family)	23	6	16	3	63	24	17	4	10	2		
457	Phosphoribosylpyrophosphate synthetase	32	4			16	2			24	3	8	1
201	hypothetical protein	52	3			41	2	39	2			41	2
539	Peroxiredoxin	19	2	16	2			16	2			39	5
2445	two-component response regulator	19	2	19	2			19	2			19	2
6969	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	17	2	17	2			21	2			21	2
3895	Ribosomal protein L17	14	1	14	1	14	1					22	2
6131	F0F1-type ATP synthase alpha subunit	12	3	13	3			8	2	8	2		
5042	Mg-protoporphyrin IX monomethyl ester oxidative	11	3	8	2	5	1			8	2		
678	Prephenate dehydrogenase	22	3					9	3			14	3
2665	IMP dehydrogenase/GMP reductase	16	4			10	2					18	5
456	Thymidine phosphorylase	13	4	8	2							10	3
1392	putative phosphoenolpyruvate-protein	4	2	4	2							8	4
6079	3-isopropylmalate dehydratase large subunit	10	2	16	4							9	2
3876	Ribosomal protein S3	15	2	15	2	15	2						
6703	conserved hypothetical protein			37	3							6	1
1858	CBS domains / conserved hypothetical protein	25	4									13	2
7145	conserved hypothetical protein	4	2									8	4
6433	Fe2+ transport system / conserved hypothetical protein	12	1					22	2				
3887	Ribosomal protein S5			15	2	15	2						
2520	acetolactate synthase (large subunit)	11	3			20	5						
5325	Carbamoylphosphate synthase small subunit	14	2			14	2						
2182	phosphoribosylformylglycinamide synthetase											8	3
3099	3-oxoacyl-[acyl-carrier-protein] synthase III											9	2
3263	hypothetical protein											17	2
4175	Enolase											12	2
4579	Acyl CoA:acetate/3-ketoacid CoA transferase, beta subunit											5	2
7144	Transglutaminase-like domain											4	2
2138	H subunit of photosynthetic rxn center complex									7	1		
3147	nitrile hydratase alpha subunit							16	2				
2358	hypothetical protein					8	1						
3839	GTPases - translation elongation factors					17	0						
3866	Ribosomal protein S7					19	2						
3873	Ribosomal protein L2			21	5								
401	Transcriptional regulator, lysR family	13	3										
2537	Isopropylmalate/homocitrate/citramalate synthases	13	3										
3599	Metal-dependent amidase/aminoacylase/carboxypeptidase	13	2										
5620	7-keto-8-aminopelargonate synthetase and related enzymes	10	2										

Results: Isolation of GroELS complex

- Most fusion proteins expressed to date in *R. palustris* are not normally produced under the growth conditions used
- GroEL is an exception.
- C-terminally His6 tagged GroEL-2 was expressed and isolated using an ATP-containing buffer in order to stabilize the GroEL-GroES interaction.
- LC-MS-MS analysis of the digested isolate allowed identification of components of both versions of the GroELS chaperonin complex (**Table 4**).
- ESI-FTICR-MS shows the His6 tagged GroEL-2 and the native GroEL-1 (**Figure 4**).

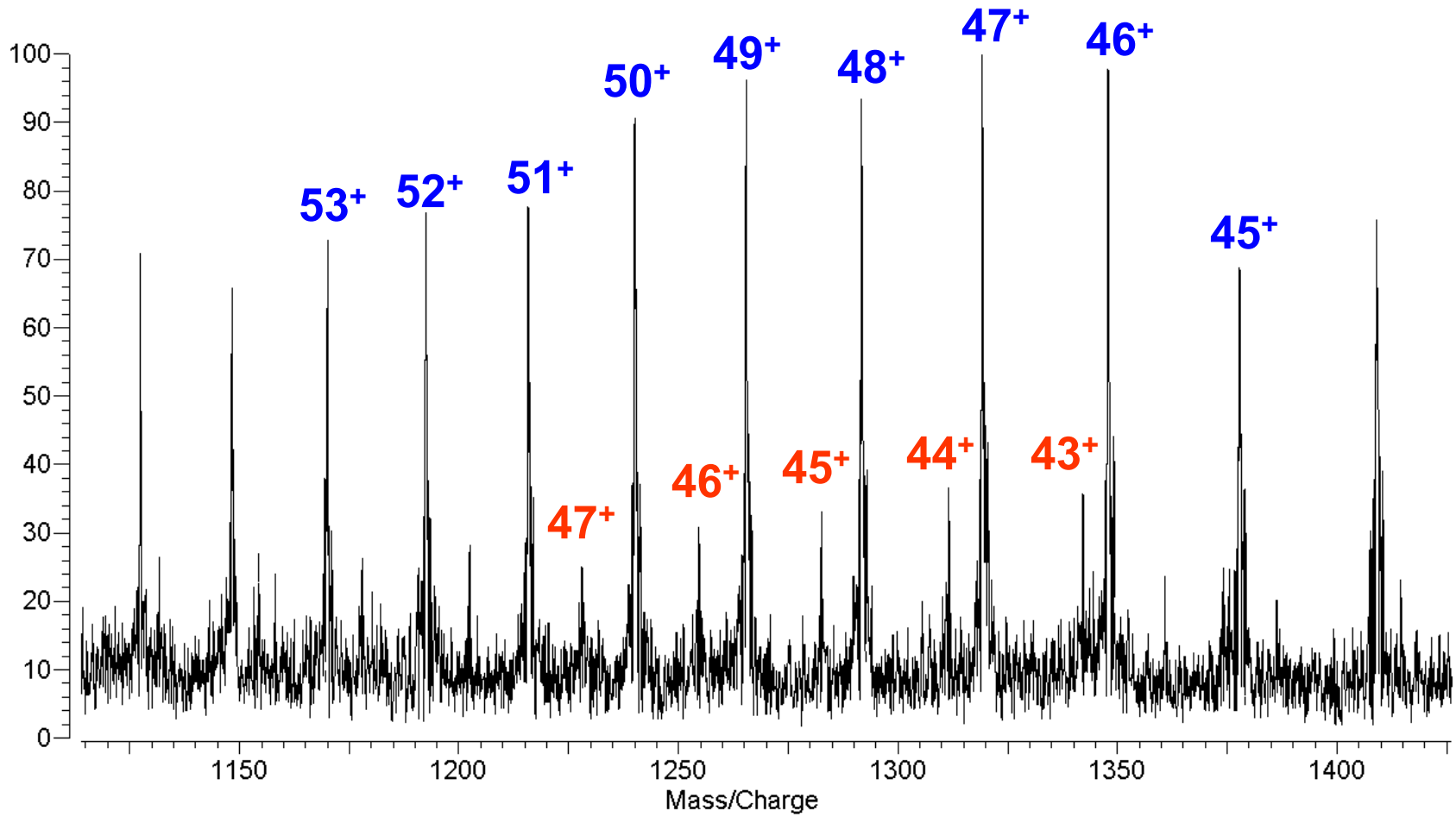
Table 4. GroELS Complex ID

Locus	Protein Name	Distinct Peptide ID's	Peptides Unique to Protein	Sequence Coverage
rpal_or4708	GroES2	7	7	58%
rpal_or4709	GroEL2 + HIS6	48	44	76%
rpal_or5425	GroES1	2	2	32%
rpal_or5426	GroEL1	29	25	64%

Figure 4. ESI-FTMS of GroEL-histag Pulldown

ORF 4709 GroEL2-histag (Mr = 62054.402 Da)

ORF 5426 GroEL1 (Mr = 57625.733 Da)



Conclusions

- Fusion proteins with affinity tags can be expressed in *R. palustris*
- The GroELS complex was isolated using a His_6-tagged GroEL as "bait"
- We are implementing higher-throughput production and analysis of fusion proteins in *R. palustris*.

References

1. http://www.genomestolife.org
2. (a) Gavin, A.-C. et al., "Functional organization of the yeast proteome by systematic analysis of protein complexes," *Nature* **2002**, *415,* 141-147.

   (b) Ho, Y. *et al*., "Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry," *Nature* **2002**, *415,* 180-183.
3. Link, A.J. *et al*., "Direct analysis of protein complexes using mass spectrometry," *Nature Biotech.* **1999**, *17,* 676-682.
4. Puig, O. *et al*., "The tandem affinity purification (TAP) method: a general procedure of protein complex purification," *Methods* **2001**, *24,* 218-229.

Acknowledgements

- This research sponsored by the Office of Biological and Environmental Research, U.S. Department of Energy. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.