

OVERVIEW

- Over 118 target proteins in *Rhodopseudomonas palustris* have been expressed as fusions with affinity labels to enable isolation of protein complexes.
- Components of protein complexes are identified by mass spectrometry.
- A large and growing data set facilitates robust identification of authentic protein-protein interactions.

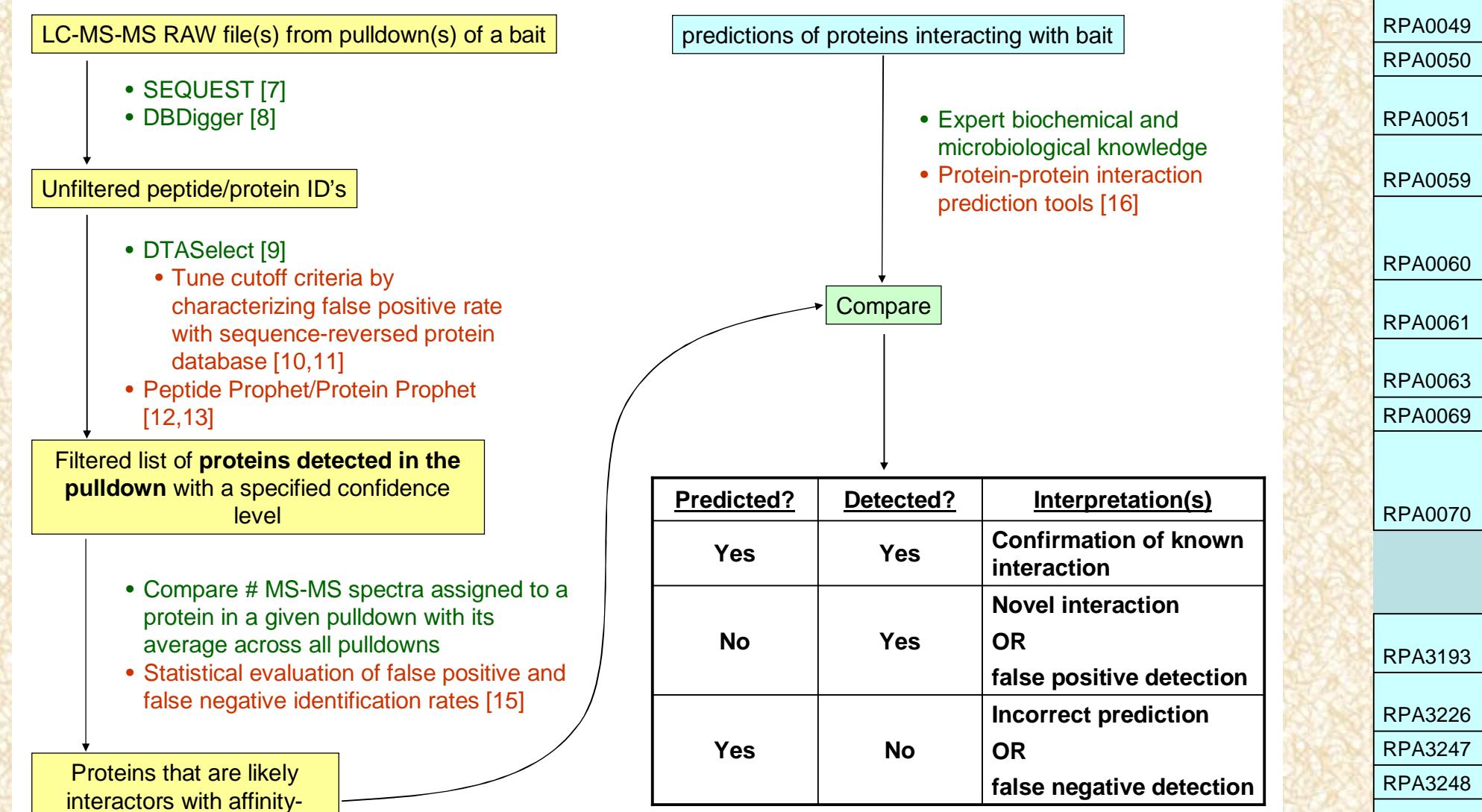
INTRODUCTION

- The bacterial species *Rhodopseudomonas palustris* [1]
 - occurs nearly ubiquitously in the environment
 - survives in a variety of conditions
 - light/dark
 - aerobic/anaerobic
 - is of interest for energy-related research
 - Produces H₂ as a byproduct of nitrogen fixation
 - secretes lignin monomers
- This species thus has the potential to express markedly different complements of proteins and protein complexes under different growth conditions.
- As part of the Center for Molecular and Cellular Systems[†] funded by the U.S. Department of Energy Genomes To Life Program [2,3], 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 [4].

EXPERIMENTAL

- Selected *R. palustris* genes were cloned with a dual (His₆/V5 epitope) affinity tag [5] at the C-terminus and expressed in *R. palustris* using modified pDEST vectors (Invitrogen).
- Isolation of fusion proteins
 - Affinity purification with Ni-NTA agarose beads, followed by anti-V5 antibody agarose beads.
 - Expression confirmed using 1-D PAGE and western blots.
- "Shotgun" analysis: analysis of isolated protein complexes by mass spectrometry without prior gel separation [6]
 - Trypsin digestion
 - Reverse-phase HPLC separation online with electrospray/quadrupole ion trap MS-MS.
 - Protein ID's: Sequest [7] comparison of MS-MS data with *R. palustris* database [1].

Verification of MS-Based Protein Identifications: Current and Future



*<http://www.ornl.gov/sci/GenomestLife/>
http://www.ornl.gov/sci/csd/Research_areas/obms_group.html

Affinity Isolation and Mass Spectrometric Analysis of Protein Complexes from *Rhodopseudomonas palustris*

Gregory B. Hurst¹, Dale A. Pelletier², Stephen J. Kennel², Frank W. Larimer², W. Hayes McDonald¹, Trish K. Lankford², Manesh B. Shah², Denise D. Schmoyer³, Tse-Yuan S. Lu², Linda J. Foote², Cathy K. McKeown², W. Judson Hervey IV⁴, Michael B. Strader¹, Robert L. Hettich¹, Heather M. Connelly⁴, Michelle V. Buchanan¹
 Chemical Sciences Division¹, Life Sciences Division², Computer Science and Mathematics Division³, Graduate School of Genome Science and Technology⁴
 Oak Ridge National Laboratory

RESULTS AND DISCUSSION

- Proteins identified in each affinity isolation experiment are collected in a database, and displayed to an internal web page.
 - See Tables 1, 2, and 3.
 - A summary is publicly available at http://maple.lsd.ornl.gov/cgi-bin/gtl_demo/public_target_status.cgi

- Detailed information for each tagged gene is accessible internally from the summary table:
 - Links to *R. palustris* genome web site
 - Links to lists of proteins detected in each pulldown experiment – See Table 4 for one example (RNA polymerase)

- For approximately 40 target proteins, our MS measurements have identified interacting proteins based on known complexes in other organisms (data not shown), with an average of 3.6 interacting proteins per target protein (minimum of 1; maximum of 11 interactions)

Table 3. Summary list of progress in analyzing protein complexes isolated from affinity-tagged genes (as of 10:34:02 AM Friday May 27 2005)

Gene ID	Gene Name	Protein Description	PCR Amp	Entry Clone	Express Clone	Host Transform	MS Analysis
RPA0001	dnaA	chromosomal replication initiator protein DnaA	96%	94%	99%	96%	81%
RPA0002	dnaN	DNA polymerase III beta subunit					
RPA0004	gyrB	DNA gyrase subunit B					
RPA0008	kaiB	circadian clock protein					
RPA0009	kaiC	circadian clock protein					
RPA0010	CDS	transcriptional regulator, probable glutamate synthase					
RPA0015	fixG, rdxB	4Fe-4S ferredoxin, iron-sulfur binding domain					
RPA0016	ccoP, fixP	cytochrome-c oxidase fixP chain					
RPA0018	ccoO, fixO	cytochrome-c oxidase fixO chain					
RPA0028	purH	bifunctional purine biosynthesis protein					
RPA0036	CDS	conserved unknown protein					
RPA0041	CDS	conserved unknown protein					
RPA0049	CDS	possible ABC transporter ATP-binding protein (in)					
RPA0050	rpoN	sigma factor (54) RpoN					
RPA0051	CDS	putative sigma-54 modulation protein					
RPA0059	CDS	L-carnitine dehydratase/bile acid-inducible					
RPA0060	CDS	RPA0060 conserved unknown protein 66656:67057 reverse MW:14461					
RPA0061	aroA	3-phosphoshikimate 1-carboxyvinyltransferase					
RPA0063	cmk	cytidylate monophosphate kinase					
RPA0069	trpB	tryptophan synthase beta chain					
RPA0070	trpA	trpA tryptophan synthase alpha subunit 76054:76890 forward MW:18607					

Table 1. Summary Statistics for LC-MS-MS Analysis of Protein Complexes from *R. palustris*

	Distinct Targets	Total (with controls, reps, QC)
Production Cultures	121	233
Total Proteins (redundant)	7700	679
Pulldowns	118	235
LC-MS-MS Analyses	118	609

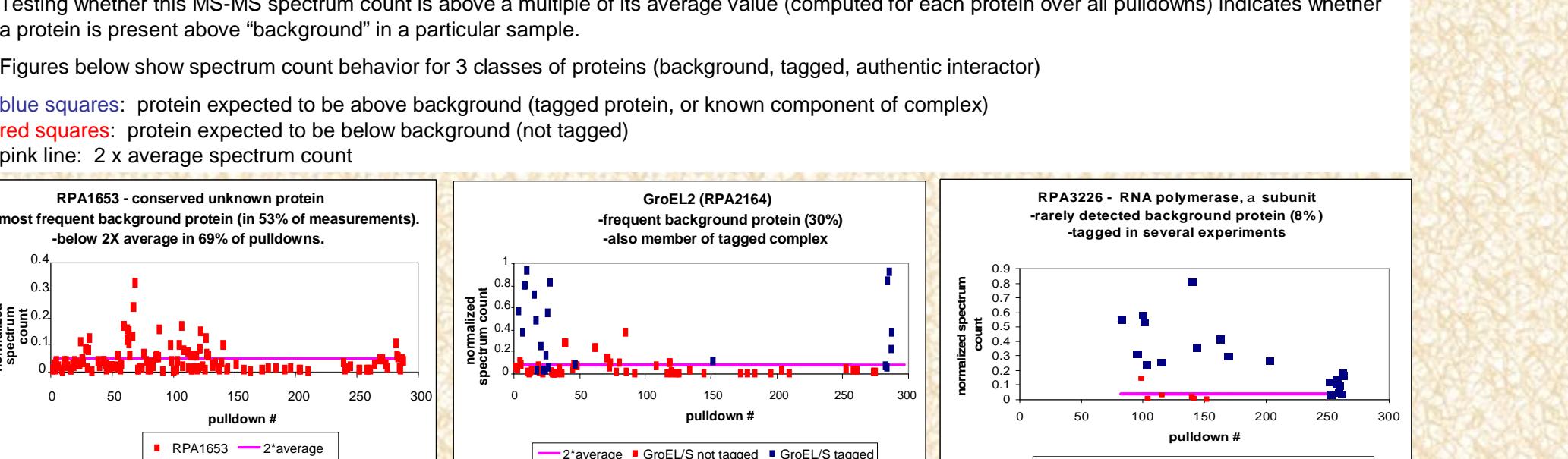
Table 2. Protein Identification Summary

Identifications	In Tagged Pulldowns	In Wild Type Controls	Total (with QC samples)
Distinct Proteins	995	270	1103
Total Proteins	995	270	1103
Pulldowns	118	235	353
Proteins above Background	5147	405	5552
Peptides passing criteria [9]	18478	1264	26595

Table 4: Link from Table 3 showing detailed MS results for one example affinity-tagged protein

Tagged Gene ID	Experiment ID	Growth conditions	Color scheme for boxes:	Legend for cells:	Legend for rows:	Legend for columns:	Protein Description
RPA3226_42	RPA3226_42	RPA3226_42	RPA3226_42	RPA3226_42	RPA3226_42	RPA3226_42	RPA3226_42
Gene	pipeline6	pipeline9	Pipeline12	Pipeline18	Pipeline21	Pipeline30	Pipeline47
	pHNH4	pHN2	pHN2	pHNH4	pHNH4	chNH4	pHNH4
RPA3226	37(25)-235-73.2--**	40(28)-165-87--**	42(27)-129-88.5--**	35(23)-124-77.3--**	25(18)-27-52.2--**	26(19)-21-54.9--**	48(31)-211-80.5--**
RPA3268	50(39)-96-41.6--**	97(70)-220-62.7--**	97(72)-225-64.2--**	63(50)-104-49.6--**	45(40)-46-36.2--**	64(54)-46-41.1--**	145(37)-313-67.9--**
RPA3267	29(26)-48-26.1--**	74(52)-156-59.1--**	54(42)-91-40.3--**	23(19)-40-18.3--**	25(25)-25-24.6--**	43(35)-20-27.9--**	86(59)-168-14.4--**
RPA2692	4(6)-8-53.1--**	10(6)-26-89.2--**	8(7)-17-89.2--**	6(4)-8-78.5--**	2(2)-34-6--**	5(4)-5-33.1--**	12(7)-23-78.5--**
RPA1288	4(4)-6-12.9--**	11(10)-17-27.7--**	3(3)-3-9--**	5(3)-8-14.2--**	3(3)-3-10--**	12(11)-12-19.6--**	14(10)-23-29.4--**
RPA0367	2(2)-3-5.7--**	8(7)-15-35.5--**	4(4)-5-17.4--**	9(8)-16-30.1--**	1(1)-1-7.4--**	8(7)-8-26.4--**	9(6)-11-28.4--**
RPA0060	2(2)-4-43.6--**	4(3)-10-45.9--**	6(5)-15-45.9--**	3(3)-4-45.1--**			5(5)-8-45.1--**
RPA0049							RPA0060 co
RPA0050							
RPA0051							
RPA0059							
RPA0060							
RPA0061							
RPA0063							
RPA0064							
RPA0065							
RPA0066							
RPA0067							
RPA0068							
RPA0069							
RPA0070							

Distinguishing Authentic Interactors from Background Proteins



CONCLUSION

- We are continuing our analysis of affinity-isolated protein complexes in *R. palustris*.
 - Analyses for 118 target proteins have been completed
 - Interacting proteins have been identified for 40 target proteins
- A database and web interface for tracking and viewing results has been implemented
- Protocols for distinguishing "background" from "authentic" interacting proteins have been implemented, and more sophisticated techniques are being evaluated.
- Related presentations at this meeting include:
 - MP31 #538: W.H. McDonald et al., "Characterization of sources of variability in LC-MS/MS analysis of protein complexes"
 - MP26 #429: H.M. Connolly et al., "Top-down MS Analysis of Microbial Growth States from *Rhodopseudomonas palustris* by Off-line FPLC Fractionation and Capillary HPLC Interfaced to FTICRMS"
 - WP24 #461: W.J. Hervey IV et al., "Proteomic Verification of Signal Peptide Prediction Algorithms in *Rhodopseudomonas palustris*"
 - WP12 #202: M.B. Strader et al., "Organic Solvents Improve Trypsin Digestion Efficiency and Specificity for Digesting Limited Protein Sample Amounts"

REFERENCES

- Larimer, F. W. et al. (2004) Complete genome sequence of the metabolically versatile phototrophic bacterium *Rhodopseudomonas palustr*