

Differentiation of Specific and Non-Specific Protein Interactions in Bacteria by Isotopic Labeling and Mass Spectrometry

W. Judson Hervey, IV^{1,3}, Chongle Pan^{1,3}, Patricia K. Lankford^{2,3}, Dale A. Pelletier², and Gregory B. Hurst³

¹UT-ORNL Graduate School of Genome Science and Technology (GST), ²Life Sciences Division (LSD), Oak Ridge National Laboratory, ³Organic and Biological Mass Spectrometry (OBMS), Chemical Sciences Division (CSD), Oak Ridge National Laboratory

OVERVIEW

- Individual members of the proteome work synergistically to accomplish biological functions within protein complexes.
- The affinity isolation of protein complex components coupled with peptide identification by mass spectrometry is a powerful, robust methodology for the characterization of biological systems.
- Disruption and removal of complex components by affinity isolation often results in non-specific, artifactual interactions among components.
- This complicates biological interpretation of interaction results, as non-specific interactions are difficult to distinguish from unknown interactions.
- This challenge has become apparent in the ongoing characterization of protein complexes of *Rhodospseudomonas palustris* by the Center for Molecular and Cellular Systems (CMCS). <http://www.ornl.gov/sci/Genomestolife/index.shtml>
- To address the specificity of interactions, we have modified existing relative quantitative proteomics techniques and applied them to the RNA polymerase complex in two different bacterial systems utilizing two alternative methods of protein expression.

INTRODUCTION

- Elucidation of the interactions among components of protein complexes is essential for an improved understanding of cellular function. Identification of the subunit proteins of protein complexes is often performed using affinity purification followed by mass spectrometric identification of the enriched proteins^{1,2}.
- However, affinity purification is laden with interference due to artifactual, non-specific interactions. Recently, quantitative proteomics techniques have been applied to the strategy in order to differentiate between specific and non-specific interactions³.
- This technique utilizes the relative quantification of protein abundance between two different cell types to differentiate between specific and non-specific protein interactions.
- To assist in our goal of characterizing the protein complexes of *Rhodospseudomonas palustris* at the CMCS, we have adapted this strategy to distinguish between authentic versus artifactual protein interactions of the affinity-purified RNA polymerase complex.
- RNA polymerase is a macromolecular protein complex whose primary function is to transcribe the genetic information encoded within the genome into RNA^{4,5}.
- In prokaryotes, a single DNA-dependent RNA polymerase protein complex transcribes all classes of RNA: mRNA, rRNA, and tRNA^{6,7}.
- The core RNA polymerase protein complex is composed of four different protein subunits: α , β , β' , and ω with a stoichiometric ratio of 2:1:1:1, respectively. The core complex associates with various sigma factors in order to transcribe genes into mRNA⁸.
- A system-wide study of the protein complexes of *R. palustris* with plasmid-expressed affinity-tagged proteins is underway to elucidate the nature of this bacterium's metabolic versatility.
- In prokaryotic systems, protein-protein interactions have been studied through the model system of *E. coli* by the isolation of affinity-tagged proteins from chromosomal insertions combined with LC-MS/MS⁹.
- In this study, relative quantitative proteomics techniques have been applied to affinity isolations of the RNA polymerase complex from *E. coli* and *R. palustris* to distinguish specific protein interactions from non-specific protein interactions. To expand the use of isotopic differentiation of protein interactions³ to prokaryotic systems, *E. coli* is employed as benchmark system to examine the protein-protein interactions of bacterial RNA polymerase.

METHODS

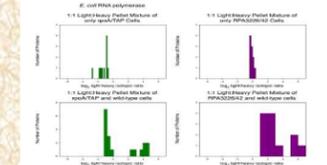
- Cellular Growth**
 - The RpoA, RpoB, and RpoC proteins of the RNA polymerase complex were expressed bearing C-terminal affinity epitopes in two different bacterial species, *E. coli* K12 and *R. palustris* CGA010.
 - The *E. coli* proteins were expressed in the tandem affinity purification (TAP)⁷ vector integrated into the *E. coli* chromosome. In the TAP vector, the native protein sequence is flanked from the C-terminus by a calmodulin binding peptide (CBP) epitope, a tobacco etch virus (TEV) cleavage site, and an immunoglobulin G (IgG) epitope.
 - The *R. palustris* proteins were expressed in the pDEST42 plasmid, which encodes the His₆ and V5 affinity epitopes onto the C-terminus of the native protein.
 - Cell cultures expressing the RpoA/TAP, RpoB/TAP, and RpoC/TAP proteins of *E. coli* were grown in "light" M9 minimal media with ¹⁴NH₄SO₄ as the sole nitrogen source. Wild-type cultures of *E. coli* and one culture of RpoA/TAP were grown in "heavy" M9 minimal medium with ¹⁵NH₄SO₄ as the sole nitrogen source. All cell cultures were harvested in the mid-logarithmic phase of growth. Each of the three cultures expressing an affinity-tagged protein grown in "light" M9 was mixed 1:1 by cell pellet mass with wild-type *E. coli* cells grown in "heavy" M9. The "heavy" RpoA/TAP culture was mixed 1:1 by cell pellet mass with "light" RpoA/TAP cells.
 - The affinity-tagged RpoA, RpoB, and RpoC proteins of *R. palustris* were expressed from the pDEST42 vector in the RPA3226/42, RPA3267/42, and RPA3268/42 strains, respectively. Each of these cultures were grown in "light" PMS-10 medium with ¹⁴NH₄SO₄ as the sole nitrogen source. Wild-type cultures of *R. palustris* CGA010 and one culture of RPA3226/42 were grown in "heavy" PMS-10 medium with ¹⁵NH₄SO₄ as the sole nitrogen source. All cell cultures were harvested in the mid-logarithmic phase of growth. Each of the three cultures expressing an affinity-tagged protein grown in "light" PMS-10 was mixed 1:1 by cell pellet mass with wild-type *R. palustris* cells grown in "heavy" PMS-10. The "heavy" RPA3226/42 culture was mixed 1:1 by cell pellet mass with "light" RPA3226/42 cells.
- Isolation of Affinity-Tagged Proteins from Mixed Cell Pellets of *E. coli***
 - Each *E. coli* cell pellet mixture was lysed with BugBuster reagent (Novagen). The cellular lysate was incubated with IgG (Amersham/GE) affinity resin for 1 hour. Elution from the IgG resin was performed by treatment with ActTEV protease (Invitrogen) for 1 hour. Protein eluates from the NINTA resin were further enriched for affinity-tagged proteins by incubation with the XV5 affinity resin (Sigma) for 1 hour, as described previously⁹. The XV5 eluate was digested with trypsin for 16 hours. The resulting proteolytic peptides were extracted from the mixture using 100 μ l reverse phase OMIX pipette tips (Varian) and stored at -80°C.
- Isolation of Affinity-Tagged Proteins from Mixed Cell Pellets of *R. palustris***
 - Each *R. palustris* cell pellet mixture was lysed with BugBuster reagent (Novagen). The cellular lysate was incubated with NINTA affinity resin (Invitrogen) for 1 hour. Protein eluates from the NINTA resin were further enriched for affinity-tagged proteins by incubation with the XV5 affinity resin (Sigma) for 1 hour, as described previously⁹. The XV5 eluate was digested with trypsin for 16 hours. The resulting proteolytic peptides were extracted from the mixture using 100 μ l reverse phase OMIX pipette tips (Varian) and stored at -80°C.
- Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)**
 - Proteolytic peptides from all affinity isolations were separated by a one-dimensional reverse phase liquid chromatography system consisting of an Ultimate HPLC pump, Switchos, and Famos Autosampler (LC Packings) coupled to a quadrupole ion trap mass spectrometer (ThermoFinnigan LCQ DECA XP+). Mass spectral data were acquired in data dependent mode with dynamic exclusion enabled (repeat count 2) during a 190 minute LC gradient. Tandem mass spectra were collected from the four most abundant ions in the full mass scan.
- Data Analysis**
 - Tandem mass spectra were searched with the SEQUEST³ algorithm against the predicted proteomes of each species^{10,11} in two separate iterations. One SEQUEST iteration searched for the identification of light peptides, while the second SEQUEST searched iteration for the identification of heavy peptides. Search results were merged and identification results were filtered and sorted by default criteria by the DTASelect¹² program.
 - Light-heavy (¹⁴N/¹⁵N) isotopic ratios for each protein identification were estimated from full mass spectra by the ProRata^{13,14} program.
 - Tab-delimited data files from the DTASelect and ProRata programs were imported into relational database software for analysis.
 - Differentiation between specific and non-specific protein interactions were based upon the light:heavy isotopic ratio estimated by the ProRata program. The proteins that were observed in more than 1 affinity isolation with a log₂ light:heavy isotopic ratio > 0 were designated as specific protein interactions with the RNA polymerase complex.

RESULTS

Overview of Results

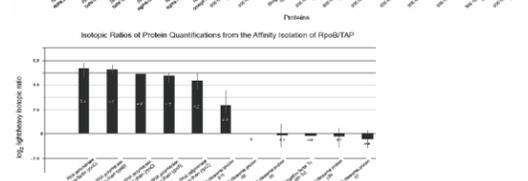
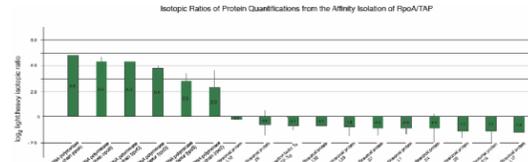
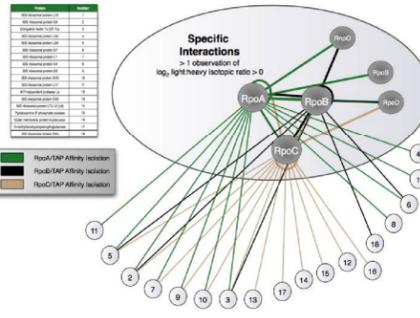
- Comparisons among the distributions of estimated light:heavy isotopic ratios in differentially-labeled cell pellet mixtures.
 - "Distributions of Light:Heavy Isotopic Ratios from Differentially-Labeled Cellular Pellets."
- LCQ Measurements of the Interactions of the RNA polymerase complex of *E. coli* by the Affinity Isolations of rpoA/TAP, rpoB/TAP, and rpoC/TAP from ¹⁴N-labeled *E. coli* cells mixed with ¹⁵N-labeled wild-type *E. coli* cells at 1:1 ratio by cell pellet mass.
 - "Differentiation of the Protein Interactions of *E. coli* RNA polymerase."
- LCQ Measurements of the Interactions of the RNA polymerase complex of *R. palustris* by the Affinity Isolations of RPA3226/42, RPA3267/42, and RPA3268/42 from ¹⁴N-labeled *R. palustris* cells mixed with ¹⁵N-labeled wild-type *R. palustris* cells at 1:1 ratio by cell pellet mass.
 - "Differentiation of the Protein Interactions of *R. palustris* RNA polymerase."

Distributions of Light:Heavy Isotopic Ratios from Differentially-Labeled Cellular Pellets



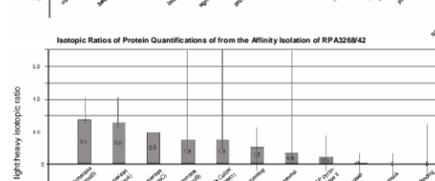
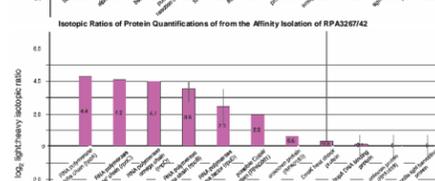
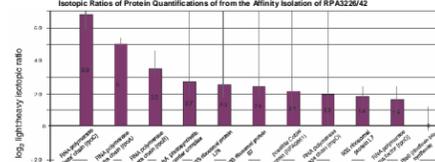
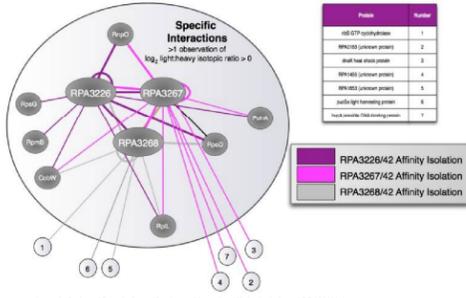
Differentiation of the Protein Interactions of *E. coli* RNA Polymerase

Interaction Map Based upon Relative Protein Quantification



Differentiation of the Protein Interactions of *R. palustris* RNA Polymerase

Interaction Map Based upon Relative Protein Quantification



DISCUSSION

- Differences in the distributions of isotopic ratios between 1:1 Mixtures
 - In both bacterial species, the distributions of isotopic ratios are different in the differentially-labeled cell pellet mixtures.
 - The distributions of the affinity-tagged protein with the wild-type background are bimodal (as shown in the figure to the upper left), indicating that there are two groups of isotopic ratios:
 - One group of ratios encompass the known, experimentally determined specific interactions range from approximately 2.0 - 4.5 for mixtures of *E. coli* and from 4.5 - 7.0 for mixtures of *R. palustris*.
 - The other group consists of ratios for non-specific protein interactions. The values of the light:heavy ratios in this group differ between the two bacterial species. In *E. coli*, these ratios range from -1 - 0, while in *R. palustris* these values range from 1 - 4.
 - These differences could be attributed to the method of expression of the affinity-tagged bait protein between bacterial species.
- Protein Interactions in *E. coli* RNA polymerase (middle left)
 - The differentiation between specific and non-specific protein interactions was straightforward based upon observed log₂ light:heavy isotopic ratios > 0 (see Methods).
 - The isolation of each of the affinity-tagged proteins resulted in the identification and relative quantification of the members of the RNA polymerase core complex. This suggests that the method of isotopic differentiation of interactions is effective from any of the tested entry points into the RNA polymerase complex.
 - The affinity isolation of RNA polymerase complex from the RpoA/TAP protein yielded log₂ light:heavy isotopic ratios of 4.8, 4.3, 4.3, and 2.3 for the α , β , β' , and ω subunits respectively.
 - The affinity isolation of RNA polymerase complex from the RpoB/TAP protein yielded log₂ light:heavy isotopic ratios of 4.7, 5.2, 4.8, and 4.3 for the α , β , β' , and ω subunits respectively.
 - The affinity isolation of RNA polymerase complex from the RpoC/TAP protein yielded log₂ light:heavy isotopic ratios of 2.1, 4.1, and 3.4 for the α , β , and β' subunits respectively.
 - The σ factor RpoD was isolated with the RNA polymerase complex with log₂ light:heavy isotopic ratios of 3.8, 5.3, and 6.4 for the RpoA/TAP, RpoB/TAP, and RpoC/TAP proteins, respectively.
 - Only one protein, that is not a member of the RNA polymerase complex, the 30S ribosomal protein S15, was identified with a log₂ light:heavy isotopic ratio > 0 in the affinity isolation of the RpoC/TAP protein.
- Protein Interactions in *R. palustris* RNA polymerase (bottom left)
 - The differentiation between specific and non-specific protein interactions required a more rigorous investigation of the use of log₂ light:heavy isotopic ratio as the sole discriminating factor.
 - On the basis of light:heavy isotopic ratio alone, specific interactions were not discernible from non-specific interactions.
 - In addition to the ratio, the number of observations in multiple affinity isolations is also an important discriminating factor between specific and non-specific interactions. Thus, specific interactions were required to have a log₂ light:heavy isotopic ratio > 0 in more than one affinity isolation.
 - The affinity isolation of RNA polymerase complex from the RPA3226/42 protein yielded log₂ light:heavy isotopic ratios of 5.0, 3.5, 6.8, and 1.9 for the α , β , β' , and ω subunits respectively.
 - The affinity isolation of RNA polymerase complex from the RPA3267/42 protein yielded log₂ light:heavy isotopic ratios of 4.4, 3.6, 4.2, and 4.1 for the α , β , β' , and ω subunits respectively.
 - The affinity isolation of RNA polymerase complex from the RPA3268/42 protein yielded log₂ light:heavy isotopic ratios of 3.3, 3.5, and 2.5 for the α , β , and β' subunits respectively.
 - The σ factor RpoD was isolated with the RNA polymerase complex with log₂ light:heavy isotopic ratios of 1.6 and 2.5 for the RPA3226/42 and RPA3267/42 proteins, respectively.
 - Five proteins were identified with log₂ light:heavy isotopic ratios > 0 in the affinity isolation of the RNA polymerase complex from *R. palustris*: RpsG, CobW, RpmB, RplL, and PuhA. These proteins were identified and quantified in more than one affinity isolation.
 - It is likely that another important discriminating factor in the differentiation between specific and non-specific interactions lies in the confidence interval determined by the ProRata program. Many confidence intervals of known non-specific proteins displayed a large confidence interval width, indicating low confidence in the measurement of the isotopic ratio.
 - The largest estimated light:heavy isotopic ratio for the affinity isolations of the RPA3226/42 and RPA3267/42 proteins were the affinity-tagged proteins themselves. This observation could be attributable to the inducible plasmid-based expression of the tagged protein from the 42 plasmid above physiological expression levels.

CONCLUSION

- We have adapted a relative quantitative proteomics methodology to differentiate between specific and non-specific interactions in a prokaryotic system.
- In the affinity isolations of the RNA polymerase complex from two different bacterial species, the core components of the complex were identified and could be distinguished as specific on the basis of light:heavy isotopic ratio in *E. coli*.
- Differentiation of the interactions of the RNA polymerase complex of *R. palustris* will require more rigorous methods of data analysis.
- Observable differences in the interactions of RNA polymerase of *E. coli* and *R. palustris* may be attributable to:
 - the method of expression of the "bait" protein,
 - the affinity epitopes used for the affinity isolation, and/or
 - the buffering conditions used during the affinity isolation.
- Future work will focus on the examination of these experimental parameters and the application of this technique to more protein complexes.

REFERENCES

- Gavin, AC et al *Nature* 415: 141-147. 2002
- Ho, Y et al *Nature* 415: 180-183. 2002
- Butland, G et al *Nature* 433: 531-537. 2005
- Tackett, G et al *Proteome Res* 4: 1752-1756. 2005
- Hiratsuki, K et al *J Bacteriol* 177: 2918. 1995
- Mah, TF et al *Genes Dev* 14: 2664. 2000
- Rigaut, G et al *Nature Biotechnol* 17: 1030-1032. 1999
- Strader, MB et al *Anal Chem* 78: 125-134. 2006
- Yates JR III et al *Analytical Chemistry* 67: 1426-1436. 1995
- Blattner FR et al *Science* 277: 1453-1474. 1997
- Larimer, FW et al *Nature Biotechnol* 22: 55-61. 2004
- Tabb DL et al *J Proteome Res* 1: 21-26. 2002
- Pan C et al "Novel Algorithms for Determination of Peptide Abundance Ratios in Quantitative Proteomics with Stable Isotope Labeling." Oral Presentation, 53rd ASMS Conference on Mass Spectrometry and Allied Topics. June 5, 2005, San Antonio, TX.
- Pan C et al *Anal Chem* (SUBMITTED) 2006.

ACKNOWLEDGMENTS

- The Laboratory of Dr. Andrew Emill, University of Toronto, for the gracious donation of the *E. coli* TAP clones used in this study.
- JH and CP acknowledge support from the UT-ORNL Graduate School of Genome Science and Technology.
- Research sponsored by the ORNL Laboratory Director's Research and Development Program, and the 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.