Evaluation of Periplasmic Proteome Enrichment by Differential Isotope Labeling and Mass Spectrometry

W. Judson Hervey, IV[1,2], Adam M. Tebbe[1,3], Patricia K. Lankford[2,4], Dale A. Pelletier[4], and Gregory B. Hurst[2]

[1]UT-ORNL Graduate School of Genome Science and Technology, Oak Ridge, TN 37830, [2]Organic and Biological Mass Spectrometry Group (OBMS), Oak Ridge National Laboratory, Oak Ridge, TN 37831, [3]Transform Pharmaceuticals, Lexington, MA 02421, [4]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

OVERVIEW

- **Isolation of selected cellular fractions is a prerequisite to detailed proteomic characterization of microbial cells.**
- **Often, selected protein isolations are difficult to characterize by existing qualitative proteomics techniques due to the presence of contaminant proteins released as an artifact of enrichment.**
- **Here, we describe a quantitative methodology that employs differential isotopic labeling of periplasm-enriched versus whole cell lysates from *E. coli* K12.**
- **The measurement of isotopic ratios for proteins in isotopologous mixtures of disparate cellular fractions assists in distinguishing *bona fide* periplasmic proteins from "contaminant" cytoplasmic proteins.**
- **This approach highlighted the identification and quantification of:**
  - **60 periplasmic proteins isolated by the H_2O cold osmotic shock isolation[1], and**
  - **61 periplasmic proteins isolated by the $GdCl_3$ cold osmotic shock isolation[2]; however,**
  - **A few abundant non-periplasmic proteins were also quantified in both periplasmic isolations.**

INTRODUCTION

- Periplasmic proteins play important roles in electron transport, binding of nutrients, cell wall biosynthesis, and modification of molecules that will eventually enter the cytoplasm[3] (see Figure 1 below).
- Cold osmotic shock-based periplasmic enrichment[1] often results in a residual amount of cellular lysis[4-5], complicating the biological interpretation of results.
- The measurement of the relative abundance of a protein in a cold-osmotic shock periplasmic enrichment versus a whole cell lysate in differentially-labeled protein mixtures provides a measurement that may be used to distinguish *bona fide* periplasmic proteins.
- This measurement is expressed as the periplasmic-to-whole cell lysate ratio (PP:WCL).
  - *Bona fide* periplasmic proteins should have a high PP:WCL ratio in the differentially-labeled protein mixtures, while
  - cytoplasmic proteins originating from cellular lysis during the cold-osmotic shock should have a low PP:WCL ratio.

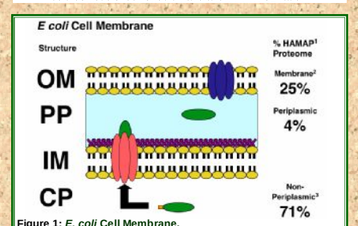


Figure 1: *E. coli* Cell Membrane.
The four structures of the gram-negative cell membrane are shown above: the outer membrane (OM), the periplasm (PP), the inner membrane (IM), and the cytoplasm (CP). [1]Percentages of the total 4,337 proteins in the predicted *E. coli* high-quality automated and manually annotated proteome (HAMAP) based on SwissProt keyword annotation. [2]Membrane is a combination of OM- and IM-annotated proteins. [3]Non-periplasmic includes proteins annotated as cytoplasmic, hypothetical, and unknown.

EXPERIMENTAL

Cellular Growth & Metabolic Labeling

- Identical *E. coli* K12 MG1655 cultures were grown in ^{14}N- and ^{15}N-enriched M9 medium with ammonium chloride as the sole nitrogen source, as described previously[6] .

Protein Isolation & Sample Preparation

- Two types of cold osmotic shock-based periplasmic enrichments were isolated from both ^{14}N-enriched ("light") and ^{15}N-enriched ("heavy") *E. coli* cultures:
  - Periplasmic isolation with H_2O as the hypotonic shock buffer[1]; and,
  - Periplasmic isolation with gadolinium chloride ($GdCl_3$) as the hypotonic shock buffer[2].
- The soluble portion of *E. coli* cell lysate was isolated from a "light" culture.
- The protein isolates described above were quantified by the MicroBCA Assay (Pierce) and mixed at equal protein concentrations as illustrated in Figure 2.

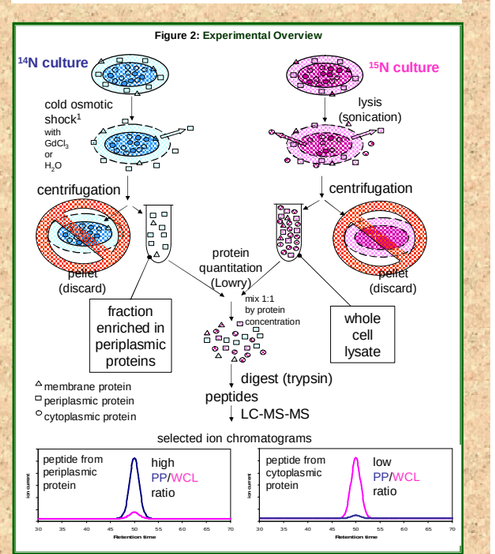
Two-dimensional nanoHPLC & Tandem Mass Spectrometry

- Proteolytic peptides were separated using the multiple dimensional protein identification technology (MudPIT)[7] coupled to a linear ion trap mass spectrometer (ThermoFinnigan LTQ).
  - The "3-phase" MudPIT separation was performed as described previously[2], over a course of 22 hours.
  - The LTQ mass spectrometer was operated in data dependent mode with dynamic exclusion enabled.

Proteome Informatics

- The SEQUEST[8] algorithm matched tandem mass spectra of peptide ions to the predicted proteome of *E. coli* K12 without enzyme specificity.
- Perl scripting was used to:
  - identify 166 known periplasmic proteins from the HAMAP SwissProt proteome,
  - prepare two separate sequences for each periplasmic protein, corresponding to presence or cleavage of an amino-terminal sequence (i.e. signal peptides, and
  - concatenate the resulting 332 proteins, along with common contaminant proteins (e. g. keratins), to the predicted HAMAP proteome for a total of 4,708 proteins.
- DTASelect[9] filtered and organized peptide identifications. Peptide and protein abundance ratios were calculated by ProRata[10].

Figure 2: Experimental Overview

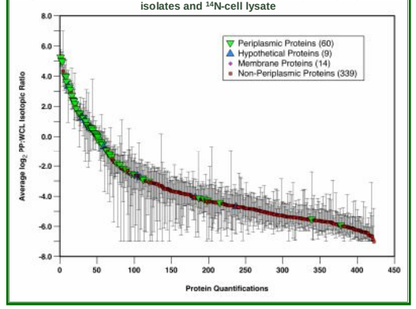


RESULTS AND DISCUSSION

Differentially-Labeled Mixtures of Periplasmic Isolations and Cellular Lysates

- ^{15}N H_2O Periplasmic Isolation mixed with a ^{14}N Cellular Lysate (Figure 3)
- ^{15}N $GdCl_3$ Periplasmic Isolation mixed with a ^{14}N Cellular Lysate (Figure 4)

Figure 3: Distribution of Protein Quantifications from H_2O ^{15}N-periplasmic isolates and ^{14}N-cell lysate

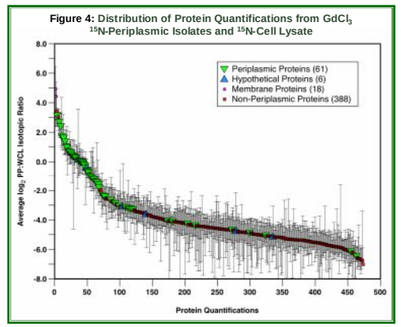


A total of 421 protein quantifications were calculated by ProRata in at least 2 of 4 total LC-MS/MS experiments.

Most of the largest average log_2 PP:WCL ratios (-3 to 5.2) were annotated as periplasmic proteins. Of 18 other proteins with log_2 ratios >0, 10 were annotated as non-periplasmic; 4 as membrane; and, 4 as hypothetical.

Most non-periplasmic proteins were quantified with PP:WCL ratios < -3.0. However, flagellin, an extracellular protein, was calculated with an average PP:WCL ratio of 3.4. It is plausible that this extracellular protein co-enriches with the periplasmic proteins during isolation. Abundant cellular proteins, such as ribosomal protein L20 and ribosome binding factor A, were also quantified, with PP:WCL ratios of -3.5 and -1.8, respectively.

Figure 4: Distribution of Protein Quantifications from $GdCl_3$ ^{15}N-Periplasmic Isolates and ^{15}N-Cell Lysate



A total of 473 protein quantifications were calculated by ProRata in at least 2 of 4 total LC-MS/MS experiments.

Again, most of the largest average log_2 PP:WCL ratios, ranging from -3.0 to 3.1, corresponded to periplasmic proteins. Of the other 22 proteins with log_2 ratios >0, 11 were annotated as non-periplasmic; 5 as membrane; and, 1 as hypothetical.

The non-periplasmic protein flagellin was again observed as co-enriching with the periplasmic proteins, with a PP:WCL ratio of 3.4.

Differentially-Labeled Mixtures of Periplasmic Isolations

- ^{14}N $GdCl_3$ Periplasmic Isolation mixed with a ^{15}N $GdCl_3$ Periplasmic Isolation (Figure 5A)
- ^{14}N H_2O Periplasmic Isolation mixed with a ^{15}N H_2O Periplasmic Isolation (Figure 5B)

Figure 5: Distributions of Protein Quantifications from Differentially-Labeled Periplasmic Enrichment Mixtures

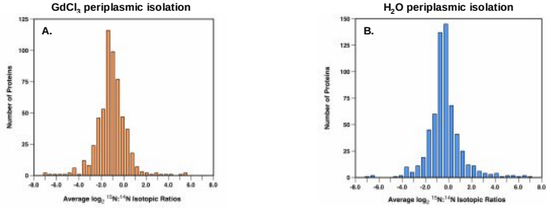


Figure 5A shows average ^{15}N:^{14}N isotopic ratios calculated for 573 proteins that represent the union of 2 LC-MS/MS experiments of the same $GdCl_3$ periplasmic isolation. 80% of the ratios ranged from log_2 2.0 to -2.0. ^{15}N:^{14}N isotopic ratios were calculated for 81 periplasmic proteins in this mixture, ranging from log_2 3.5 to -7, representative of ^{15}N:^{14}N abundance ratios of 11 to 0.01, respectively.

Figure 5B shows the average ^{15}N:^{14}N isotopic ratios calculated for 621 proteins that represent the union of 2 LC-MS/MS experiments of the same H_2O periplasmic isolation. 87% of the ratios ranged from log_2 2.0 to -2.0. ^{15}N:^{14}N isotopic ratios were estimated for 85 periplasmic proteins in this mixture, ranging from log_2 6.5 to -3.5, representative of ^{15}N:^{14}N abundance ratios of 91 to 0.09, respectively.

Table 1: 20 Largest PP:WCL Ratios of Proteins in the Differentially-Labeled Mixtures of Periplasmic Isolations and Cellular Lysate

Protein	Description	Protein Location	H_2O isolation PP:WCL log_2 Ratio	H_2O Isolation CI Width	$GdCl_3$ isolation PP:WCL log_2 Ratio	$GdCl_3$ Isolation CI Width	Signal Peptide Cleavage Identified	Spectrum Count of N-terminus
ProX	Glycine betadine-binding periplasmic protein	Periplasmic	5.2	1.1	2.7	2.1	Yes	194
PotF	Putrescine-binding periplasmic protein	Periplasmic	5.0	1.4	3.1	1.6		
GlnH	Glutamine-binding periplasmic protein	Periplasmic	5.0	1.3	1.6	1.1		
YdgH	Protein ydgH	Non-Periplasmic	4.1	1.0	2.5	1.6		
FiU	Catecholate siderophore receptor fiu	Membrane	4.0	2.1	2.9	3.4		
PtrA	Protease 3	Periplasmic	4.0	1.4	-0.8	3.9		
SubI	Sulfate-binding protein	Periplasmic	3.6	1.5	0.3	1.4	Yes	10
FliC	Flagellin	Non-Periplasmic	3.4	0.8	3.4	0.9		
ModA	Molybdate-binding periplasmic protein	Periplasmic	3.3	0.8	0.4	0.8		
DsbC	Thiol:disulfide interchange protein dsbC	Periplasmic	3.1	1.3	0.6	1.5	Yes	32
YcgK	Protein ycgK	Non-Periplasmic	2.9	1.6	3.1	1.6		
TolB	Protein tolB	Periplasmic	2.7	2.7	0.5	1.6		
SurA	Chaperone surA	Periplasmic	2.4	1.7	0.4	1.1	Yes	65
ArtJ	Arginine-binding periplasmic protein 2	Periplasmic	2.4	1.3	1.6	1.5		
PotD	Putrescine-binding periplasmic protein	Periplasmic	2.3	2.6	1.6	2.3	Yes	79
YodA	Metal-binding protein yodA	Non-Periplasmic	2.0	0.6	0.3	0.7	Yes	17
FkbA	FKBP-type peptidyl prolyl cis-trans isomerase	Periplasmic	1.8	1.2	-2.9	1.3	Yes	78
YifE	Protein yife	Non-Periplasmic	1.7	1.2	-4.3	2.5		
OsmE	Osmotically-inducible lipoprotein E	Membrane	1.7	1.2	0.4	1.6		
FliY	Cystine-binding periplasmic protein	Periplasmic	1.6	1.0	1.8	1.3	Yes	35

- **381 total protein quantifications were similar between the H_2O and $GdCl_3$ isolations. The 20 proteins quantified with the largest calculated PP:WCL ratios in BOTH mixtures of:**
  - **"heavy" H_2O periplasmic isolation and "light" cell lysate (shown in Figure 3) and**
  - **"heavy" $GdCl_3$ periplasmic isolation "light" cell lysate (shown in Figure 4) are shown above in Table 1.**
  - **The average log_2 values for PP:WCL ratio and confidence interval width are shown for each set of measurements.**
- **If the processed N-terminal peptide of a periplasmic protein was identified, it is denoted in the "Signal Peptide Cleavage Identified" column, followed by the number of MS/MS identification spectra for each processed N-terminus.**
- **Both the $GdCl_3$ and H_2O periplasmic isolation protocols enrich periplasmic proteins along with a few non-periplasmic contaminants and membrane proteins.**

CONCLUSIONS

- **Differentially-labeled mixtures of periplasmic isolates yielded similar distributions of ^{15}N:^{14}N isotopic ratios and protein quantifications for both of the periplasmic isolation methods tested.**
  - **$GdCl_3$ isolation – 81 periplasmic proteins quantified (Figure 5A).**
  - **H_2O isolation – 85 periplasmic proteins quantified (Figure 5B).**
  - **Periplasmic proteins represented the majority of high PP:WCL ratios for both methods.**
- **Differentially-labeled mixtures of periplasmic isolates and cell lysates quantified:**
  - **60 periplasmic proteins from the H_2O isolation (Figure 3).**
  - **61 periplasmic proteins from the $GdCl_3$ isolation (Figure 4).**
  - **Both methods resulted in reproducible isolation of approximately 60 periplasmic proteins in 2 or more of 4 LC-MS/MS experiments, suggesting that they are comparable methods of enrichment.**
- **Some proteins may be artifacts of either the periplasmic or cell lysate isolation, such as:**
  - **Non-periplasmic proteins with high PP:WCL ratios**
    - flagellin (Figures 3-4/Table 1)
  - **Periplasmic proteins with low PP:WCL ratios**
  - **spheroplast protein Y,**
    - expressed only in spheroplasts
    - rarely identified ^{15}N isotoplog
  - **periplasmic trehalase,**
    - expressed only during osmotic downshock; perhaps the osmotic shock procedure.
- **Thus, our strategy was able to identify a large portion of the periplasmic proteome with increased confidence based on measured abundances of enriched periplasmic proteins relative to the cell lysate.**

REFERENCES

1. Neu HC and Heppel LA *J Biol Chem* 240: 3685-3692. 1965
2. Ewis HE *et al* FEMS Microbiol Lett 253: 295-301. 2005
3. Ferguson SJ "The Periplasm" in *Prokaryotic Structure and Function: A New Perspective*, S Mohan, C Dow and JA Coles, eds. Cambridge University Press, 310-339. 1991
4. Link AJ *et al Int J Mass Spec and Ion Processes* 160: 303-316. 1997
5. Link AJ *et al Electrophoresis* 18: 1259-1313. 1997
6. Meselson M and Stahl FW *PNAS* 44: 671-682. 1958
7. McDonald WH *et al Int J Mass Spec and Ion Processes* 219: 245-251. 2002
8. Yates JR III *et al Anal Chem* 67: 1426-1436. 1995
9. Tabb DL *et al J Proteome Res* 1: 21-26. 2002
10. Pan C *et al Anal Chem* 78: 7121-7131. 2006

ACKNOWLEDGMENTS

- **Research sponsored by 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.**
- **WJH gratefully acknowledges the support of the UT-ORNL Graduate School of Genome Science, GBH, and fellow graduate students.**