# QUALITATIVE AND QUANTITATIVE PROTEIN PROFILING FOR TARGET IDENTIFICATION STUDIES WITH DRUG-INDUCED CHANGES IN MYCOBACTERIA

Jeffrey C. Silva<sup>1,3</sup>, Minerva A. Hughes<sup>2</sup>, Craig A. Dorschel<sup>1</sup>, Scott J. Geromanos<sup>1</sup>, Johannes P. C. Vissers<sup>1</sup>, Craig A. Townsend<sup>2</sup> Waters Corporation, Milford, MA; Johns Hopkins University, Department of Chemistry, Baltimore, MD; Current address: Cell Signaling Technology, Danvers, MA

# **OVERVIEW**

- A qualitative and quantitive study of a microbial proteome from a single LC-MS method.
- A label-free method to conduct both relative and absolute quantitation in the same experiment.
- Utilize simple, traditional protein fraction methods to increase the total number of securely identified proteins
- The LC-MS informatics can be used to characterize proteins in future or related studies.

# INTRODUCTION

Tuberculosis is a major infectious disease causing more than 3 million deaths annually. Isoniazid (INH) has been the most-effective therapeutic agent against TB. However, the rise of multi-drug resistant strains presents a significant health burden and a need for new preventative and treatment regimens. Using proteomics to elucidate drug-induced alterations, affected metabolic pathways, mechanisms for drug resistance or the identification of novel enzymes is possible. Classical approaches using 2D-PAGE and mass spectrometry have been applied to Mycobacteria for this purpose but are not amenable to high-throughput studies. The Waters<sup>®</sup> Identity<sup>E</sup> High Definition Proteomics system was used to analyze M. bovis BCG after exposure to INH and a synthetic antimicrobial candidate, DSA, to utilize protein profiling as a means to determine differences between the two bioactive molecules.

# **METHODS**

### Sample Preparation

*M. bovis* BCG was grown to mid-exponential phase  $(OD_{600} = 0.4-0.5)$  in standard Middlebrook 7H9/ADC-Tween 80 medium. Cultures were treated with INH (isoniazid, 1 µg/mL final), DSA (synthesized antimicrobial, 10 µg/mL final) or DMSO for 3 h, harvested, washed and lysed with 50 mM NH<sub>4</sub>HCO<sub>3</sub>, 5 mM EDTA and 1M urea to a final protein concentration of 5 - 10 mg/mL. Protein extracts were fractionated sequentially into five fractions by ammonium sulfate precipitation. The ammonium sulfate concentrations were chosen to distribute the protein equally among the five fractions. Both crude (unfractionated) and fractionated samples were reduced (dithiothreitol), alkylated (iodoacetamide), and treated with trypsin. After digestion, predigested bovine serum albumin was added to each sample at 250 fmol/µL as a qualitative and quantitative standard for analysis by LC-MS. Triplicate injections were performed for the six fractions from each sample and subsequent protein identification and expression profiling analysis was performed for the three conditions in this study as described in previous studies.<sup>1-3</sup>

## Data Acquisition and Processing

Chromatography was carried out on a Waters nanoACQUITY UPLC<sup>®</sup> system using  $3 \mu C_{18}$  columns (300  $\mu$  ID x 15 cm). Peptides were eluted using an acetonitrile/0.1% formic acid gradient (5-40%, 90 minutes) and the column was reequilibrated at initial conditions for 30 minutes before the next injection.

Data were acquired on a Q-Tof Premier<sup>™</sup> hybrid guadrupole-time of flight mass spectrometer in alternating low and elevated collision energy scanning mode (MS<sup>E</sup>), utilizing a reference spray of [Glu]<sup>1</sup>-fibrinopeptide B and erythromycin.

Mass/retention time peak detection, charge state reduction, deisotoping, time alignment, databank searching and quantitation were carried out using Identity<sup>E</sup> software. Additional data analysis was performed with Microsoft Excel and Spotfire DecisionSite.

# **RESULTS**

### Intersection of Identified Proteins from *M. bovis*



Figure 1. Reproducibility of Protein Identifications.

Protein identifications from triplicate injections of Control, DSA and INH treated *M. bovis* (~3µg loaded on column for each condition) were independently evaluated using only those proteins reported in 2 out of 3 injections for each condition. The resulting protein identifications from each condition were compared with the others to illustrate the intersection of proteins among the three conditions (**Figure 1**). The extent of intersection is reported showing the **NUMBER** of proteins in common and unique to the three conditions as well as the **fmole PERCENT** of proteins in common and unique to the three conditions. A union of the protein identifications from the three conditions indicates a total of 434 identified proteins. A total of 315 proteins were found in at least 2 out of 3 conditions. These proteins accounted for an estimated 96% of the total mass of material loaded on the analytical column (in fmoles) as determined by the relationship of the top 3 most intense peptides for each characterized protein and the top 3 most intense peptides to a known quantity of spiked internal standard (bovine serum albumin).<sup>4</sup> The ability to estimate the absolute quantity of protein is a feature of the Identity<sup>E</sup> software.

# TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS



Figure 2. Differentiation of Microbial Treatments by PCA.

Principal component analysis (PCA) was performed on the proteins identified in each replicate injection for each condition and their corresponding estimated fmole quantities. Figure 2 illustrates the **distinction** among the three different samples by the separation of the three clusters (Control, INH and DSA treatment) as well as the **reproducibility** of the analysis by the close proximity of the replicate injections.

The average fmole quantity of identified proteins has been organized in descending order for each treatment condition and is illustrated in Figure 3 and is shown to exhibit a characteristic trend. A few proteins of interest are shown in the right panel to illustrate change in protein abundance among the different conditions.

Absolute Quantitation of Identified Proteins



Figure 3. Descending Abundance of Identified Proteins.

Heirarchical Clustering of Unfractionated Samples



Figure 4. Unique Proteins to DSA Condition.



Figure 5. Normalization of Matched Intensity Ratios.

**Figure 4** highlights two proteins that are uniquely associated to the DSA condition, heat-shock protein HSP and a benzoquinone methylase. The absence of these proteins in both Control and INH suggests a high degree of induction resulting from the DSA treatment.

Autonormalization of the identified peptides from the three conditions (Figure 5) was performed before determining protein fold changes among the different treatments. The foldchanges associated with a subset of identified proteins are shown in **Figure 6**, along with their corresponding 95% confidence intervals. These proteins are also listed in **Table 1** with their corresponding index number and associated foldchange. Of particular interest is ACPM, known to be upregulated in INH treated *Mycobacterium*, which does not exhibit any significant change in DSA treated cells (Figure 7). A number of these highlighted proteins exhibit differential changes between INH and DSA treatment.



Figure 6. Relative Quantitation across Multiple Conditions.



Figure 7. Protein Quantitation from Identified Tryptic Peptides.

Protein fractionation was carried on the protein extracts from the three treatments (Control, INH and DSA) using traditional ammonium sulfate precipitation methods Unfractionated (fraction 0) and fractionated samples (fractions 1-5) were analyzed by LC-MS<sup>E</sup> to identify and quantify the constituent proteins.

PCA was performed on the identified proteins for each replicate injection of each fraction and condition along with their corresponding estimated fmole quantity. Figure 8 illustrates the **distinction** among the six sets of samples by their separation using two of the principal components. It also suggests a reasonable degree of **reproducibility** with the protein fractionation by the close proximity of the replicate injections and corresponding treatments. The reproducibility of the protein fractionation is also demonstrated from heirarchical clustering results illustrated in **Figure 9**. The outcome of the fractionation resulted in the characterization of 1110 proteins in which 851 (~77%) were found in at least two out of the three conditions (Figure 10).



Table 1. Highlighted Identified Proteins and Relative Quantitation.





**Reproduciblity of Tryptic Peptide Ion Maps** 

The data in Figure 7 illustrates the reproducibility of peak intensity measurements among the replicate injections and the mass accuracy and retention time reproducibility of characterized peptides typically obtained in this study.



**PCA of Fractionated Protein Samples** 

Figure 8. Differentiation of Fractionated Samples by PCA.

Heirarchical Clustering of Fractionated Samples



Figure 9. Reproducible Protein Fractionation.

### **Improved Characterization of Proteome**



Figure 10. Intersection of Identified Proteins.

# CONCLUSIONS

- The Waters Identity<sup>E</sup> High Definition Proteomics system was used to securely identify and quantify proteins to establish differences between separate drug treatments in Mycobacterium.
- PCA is a useful strategy for assessing LC-MS<sup>E</sup> data.
- Coupling protein fractionation methods with LC-MS<sup>E</sup> substantially increased coverage of confidently identified proteins in *Mycobacterium*.

### References

- 1. J. C. Silva, R. Denny, C. Dorschel, M. V. Gorenstein, G.-Z. Li, K. Richardson, D. Wall and S. J. Geromanos Simultaneous Qualitative and Quantitative Analysis of the E. coli Proteome: A Sweet Tale, (2006) Molecular and Cellular Proteomics, 5:589-607.
- 2. M. Hughes, J. C. Silva, S. J. Geromanos and C. A. Townsend Quantitative Proteomic Analysis of Drug-Induced Changes in Mycobacteria, (2006) Journal of Proteome Research, 5:54-63.
- 3. J. C. Silva, R. Denny, C. Dorschel, M. Gorenstein, I. J. Kass, G.Z. Li, T. McKenna, M. J. Nold, K. Richardson, P. Young and S. Geromanos Quantitative Proteomic Analysis by Accurate Mass Retention Time Pairs (2005) Analytical Chemistry, April 1, 77(7): 2187-2200.
- 4. J. C. Silva, M. V. Gorenstein, G.-Z. Li, J. P. C. Johannes and S. J. Geromanos Absolute Quantification of Proteins by LCMSE: A virtue of parallel MS acquisition, (2006) Molecular and Cellular Proteomics, 5:144-156
- 5. Johannes P. C. Vissers, James I. Langridge, and Johannes M. F. G. Aerts, Analysis and Quantification of Diagnostic Serum Markers and Protein Signatures for Gaucher Disease (2007) Mol. Cell. Proteomics, 6: 755 - 766

720002234EN ©2007 Waters Corporation v1