

Characterizing the Secreted Proteome of *Mycobacterium tuberculosis*

Analysis of Sub-Cellular Proteomes using the AB SCIEX 3200 QTRAP® System with iTRAQ® Reagent Protein Quantification

Matthew M Champion¹, Patricia A DiGiuseppe Champion²
¹AB SCIEX, USA; ²University of California, San Francisco, USA

A common strategy of MS-based proteome analysis and identification is a partial analysis of sub-cellular proteomes to detect more of the lower abundant proteins. Depending on the sample, the data quality and validity of the resultant identifications are important factors to monitor when determining true sub-proteome identifications vs. coincidental detection.

In addition to general secretion apparatuses, mycobacteria utilize an alternative secretion system called ESX-1 (Figure 1, left). This system is required for virulence and has very few characterized substrates. Here, we employed iTRAQ® reagent labeling coupled with LC/MS/MS analysis to identify, quantify and characterize the secreted proteome from *M. tuberculosis* (Erdman) (Figure 1, right).



Key Features for Detecting and Quantifying Secreted Proteins

The goal of this research was to characterize proteins secreted specifically by the ESX-1 secretion apparatus into the culture supernatant. The analytical challenge lies in the fact that there are many other proteins present in the culture supernatant. In addition, a major problem is differentiating changes related to the biological question of interest from changes that arise from sample preparation, such as cellular lysis. For this data, the following features were essential to accurately characterize secreted proteins.

- Robust, flexible quantification strategy – The iTRAQ Reagent workflow with its ability to enable sample bias correction provides a unique and essential capability to normalize data and correct for sample preparation inconsistencies and ensure accurate quantification of the sub-proteome proteins.
- High quality MS/MS data with good low-mass ion statistics - Due to the unique hybrid nature of the 3200 QTRAP® system, mass isolation and fragmentation occur outside of the ion trap, producing high quality fragmentation patterns for accurate peptide identification and good low mass ion statistics for accurate quantification of iTRAQ Reagent reporter ions.¹
- Powerful protein identification and quantification software - The Paragon™ Database Search Algorithm in ProteinPilot™ Software identifies more peptides with its unique ability to search for many peptide modifications and non-conformant enzyme cleavages simultaneously. The ProGroup™ Algorithm also enables isoform specific iTRAQ Reagent quantification, important for quantifying closely related substrate isoforms.

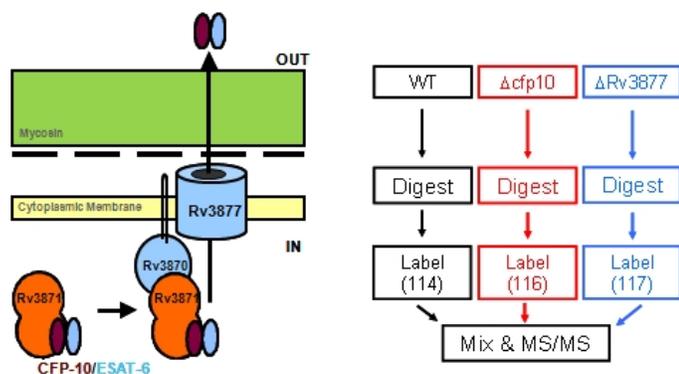


Figure 1. Experimental Plan. (Left) The general model for protein secretion via the ESX-1 secretion machine is illustrated. The two known substrates ESAT-6 and Cfp-10 are both dependant on one another for secretion. They interact with components of the machine to be secreted out to the extra-cellular medium. Removal of any component of this machine or either substrate results in the absence of both substrates from the media. (Yellow=Cytoplasmic membrane, Green=Mycosin) (Right) Analytical strategy. WT tuberculosis culture filtrates were labeled with iTRAQ Reagent 114. Culture filtrates from strains deleted for Cfp-10 substrate were labeled with 116 reagent and strains deleted for the secretion machine protein Rv3877 were labeled with 117 reagent.

Methods

Sample: Short-term culture filtrates from *Mycobacterium tuberculosis* were obtained from 50 mL cultures and 10 mL cell equivalents of culture supernatant were used. 50 µg of each sample was denatured, reduced and acetylated (MMTS), then digested with trypsin overnight at 37 °C. Labeling with iTRAQ Reagent was performed according to the standard Applied Biosystems kit protocol. iTRAQ Reagent labeling added approximately 65 minutes to the general procedure for preparing proteins for LC/MS/MS analysis. The 50 µg used for the digestion constituted about 25% of the total amount from each filtrate, or about 2.5 mL worth of cells. The samples utilized were WT cells (114), cells in which the substrate Cfp10 was deleted (116), and cells in which a secretion machine protein Rv3887 was deleted (117). A portion of each sample was mixed 1:1:1 prior to LC/MS/MS analysis.

Chromatography: Nanoflow chromatography was performed using the Tempo™ nanoLC System using a C18 reverse phase column (PepMap, 75 µm ID, 250 nL/min).

Mass Spectrometry: LC/MS/MS analysis was performed using the NanoSpray® Source and Heated Interface on the 3200 QTRAP® LC/MS/MS System using Analyst® Software.

Database Searching: Protein identification and iTRAQ® Reagent quantification was performed using the Paragon™ Algorithm Thorough search mode in ProteinPilot™ Software.

Unique Normalization Enabled by the iTRAQ® Reagent Labeling Strategy

Changes in proteins identified in the culture supernatant can be expected to arise from three sources: changes in secretion, improper sample mixing and cell lysis. Because iTRAQ Reagent labels all peptides, normalizing the three samples using a non-secreted protein corrects for both quantification due to lysis and mixing, enabling true changes in secretion to be accurately measured. In this case, the chaperones, GroEL, DnaK (Hsp70) and the elongation factor EFtu were used to normalize the quantification.

The protein identified in Figure 2, ESAT-6, is a known substrate of the machine and therefore functions as a positive control for the sample preparation and the experiment. ESAT-6 is identified in media from WT cells, but no signature ions are detected for the same peptide from the media from the ΔCfp10 (116) or ΔRv3877 (117) strains. Mutations in Cfp-10 are known to destabilize ESAT-6 and mutations in the secretion machine are expected to impact secretion².

This quantification pattern provides a signature for substrates of the secretion machine, which we can use to further probe for dependent and non-dependent substrates. Since the three samples are mixed together, similar labeled peptides are isobaric and co-elute, the failure to detect a peptide in the clear presence of the same peptide allows us to confidently say the protein is no longer secreted within our detection limits.

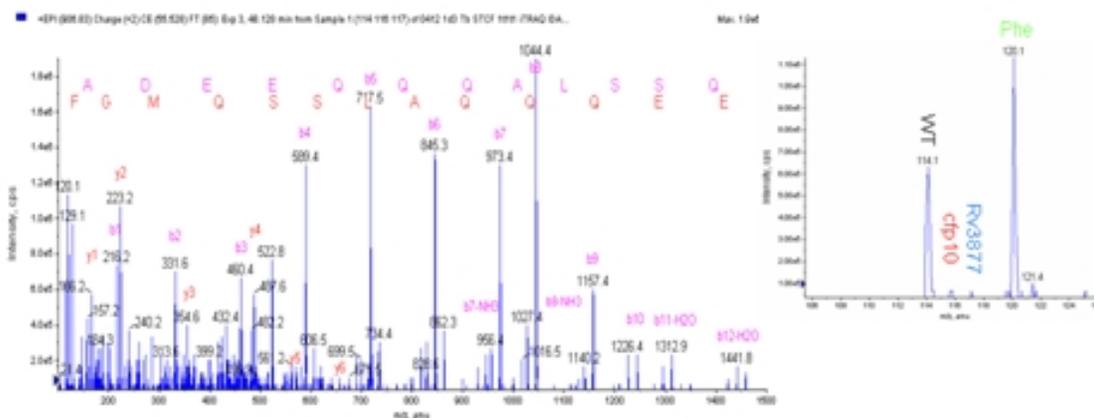


Figure 2. Example Spectra from a Secreted Protein. (Left) Full scan MS/MS spectrum (Enhanced Product Ion, EPI) from the 3200 QTRAP® System of a proteotypic peptide from a secreted substrate of *M. tuberculosis*. Note the extremely high spectral quality of both b- and y-ion series in addition to the excellent ion-statistics in the low mass region. (Right) iTRAQ® Reagent reporter ion region from this spectrum clearly denotes this protein as being secreted in wild-type ET cells (114), but not secreted at all from cells with mutations in Cfp10 (a dependent substrate, 116) or RV3877 (the secretion machine, 117). The immonium ion from phenylalanine is labeled for reference.

Table 1. Sequence Alignment of Two ESAT-6 Homologs from *M. tuberculosis*. In the top section of the table, the amino acid sequence is aligned in 10 amino acid increments. The bottom section shows the theoretical tryptic fragments that arise from digestion. Shown in red are amino acids that discriminate these substrates from each other and others within the family of homologs (other homologs not shown).

Rv1793	EsxN	MTINYQFGDV DAHGAMIRAQ AAS LEAEHQ A IVRDVL AAGD FWGGAG SVAC QEF ITQLGRN FQVIYEQANA HGQK
Rv2346c	EsxO	MTINYQFGDV DAHGAMIRAQ AGL LEAEHQ A IVRDVL AAGD FWGGAG SVAC QEF ITQLGRN FQVIYEQANA HGQK
Tryptics		
Rv1793	EsxN	MTINYQFGD V DAHGAMIR AQAAS LEAEHQ AIVR DVL AAGD FWGGAG SVACQEF ITQLGR NFQVIYEQANA HGQK
Rv2346c	EsxO	MTINYQFGD V DAHGAMIR AQA GL LEAEHQ AIVR DVL AAGD FWGGAG SVACQEF ITQLGR NFQVIYEQANA HGQK

Isoform Specific Quantification from iTRAQ® Reagent Labeled Peptides

In *M. tuberculosis*, five near-duplicates of Cfp-10 and ESAT-6 are known to exist within the chromosome. Although the orientation and existence of these genes is known, little characterization has been performed on their secretion, and the dependence of these duplicates on the ESX-1 machine (Rv3877) is unknown. Table 1 presents an alignment of the amino acid sequences from two such Cfp-10 paralogs, called EsxN and EsxO. A high degree of sequence identity (72/74 amino acids, 97%) exists between the two homologs, however a single proteotypic tryptic fragment is present that is capable of differentiating these isoforms.

With a small number of differentiating peptides for these substrates, quantification of the ESX-1 paralogs sometimes relied on single peptide quantification. In Figure 3, quantification and annotation of two such substrates is shown. Some of the specific y and b-series ions are highlighted on the spectra as they are responsible for distinguishing between the two peptide sequences. Interestingly, the iTRAQ Reagent quantification revealed that both of these ESX-1 like substrates, in addition to at least three other ESAT-6/Cfp-10 like pairs were secreted to WT levels in the absence of a functional secretion apparatus. Since these are not substrates of the ESX-1 secretion machine, there must be other, uncharacterized and unidentified secretion machines present in this organism.²

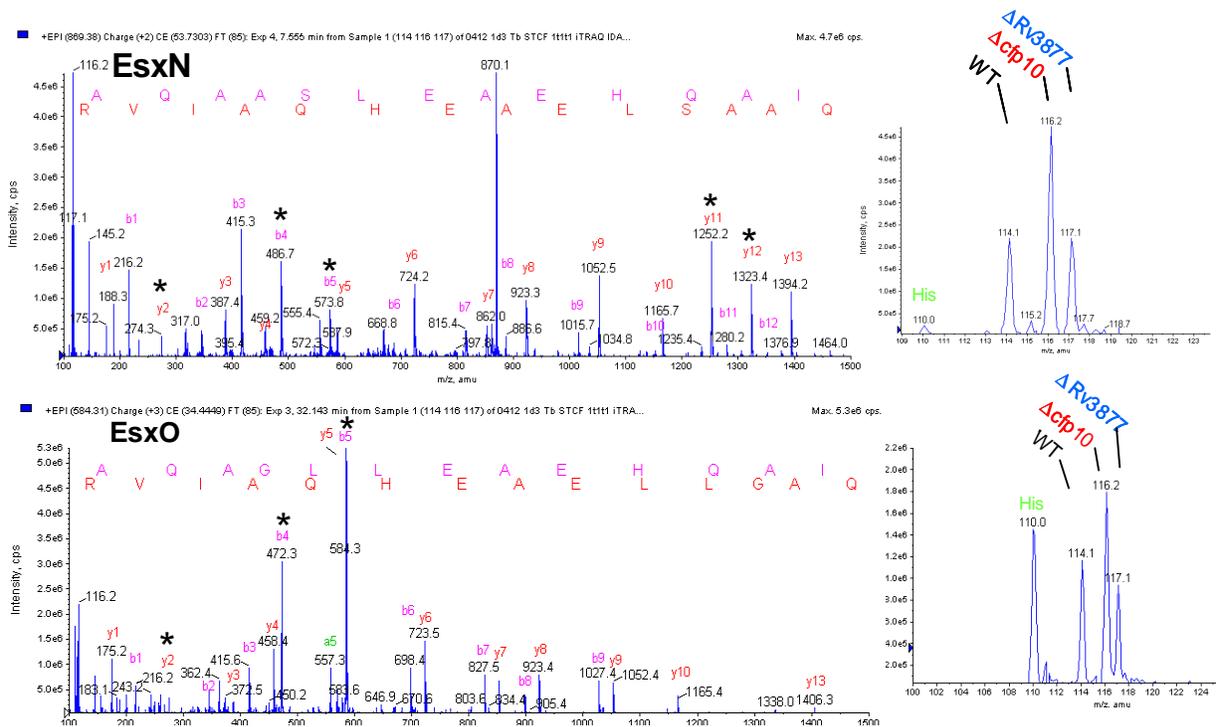


Figure 3. Peptide Identification and Quantification for Closely Related Secreted Proteins. Shown above are the differentiating tryptic peptides between EsxN and EsxO. Labeled with * are the relevant y and b-series ions that establish specific identity. Shown in the inset is the specific quantification from each isoform. Note each of these substrates is secreted independent of either Cfp10 or Rv3877.

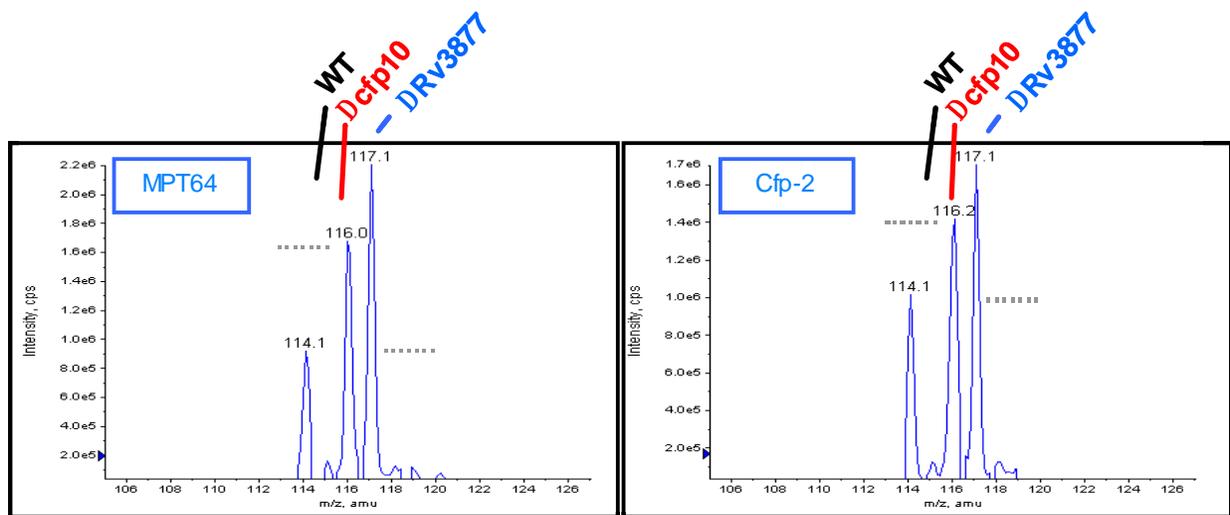


Figure 4. Sec Substrates Increase Secretion in Rv3877 Knockouts. This figure illustrates two examples of proteins that have increases in Snm2 mutants. Peptide quantification data from MPT64, a major secreted antigen, and Cfp-2, another secreted antigen, illustrate paradoxical secretion. The dotted lines mark the approximate peak intensity on the 116 and 117 signals which corresponds to a 1:1 peak area ratio after normalization. Both of these proteins contain N-terminal secretion signals, which means secretion should not be dependent on the ESX-1 system, but this data indicate that there is a relationship between Sec and ESX-1.

iTRAQ® Reagent Reveals Paradoxical Secretion from Sec-dependent Substrates.

- In addition to identification of secreted substrates specific to the ESX-1 alternate secretion machine, we were able to identify several dozen additional secreted proteins in the media that are secreted via the general secretion apparatus (Sec). For many of these Sec dependent substrates, we were able to observe a quantitative increase in their secretion in mutants which are deficient for ESX-1 secretion (Rv3877), but not the ESX-1 substrate mutants (Cfp-10) (Figure 4). Although the iTRAQ Reagent actually provides quantitative information on all labeled peptides from a protein, only one reporter region is illustrated from each protein for simplicity. These panels illustrate that for these two proteins, and eight others (data not shown), increases were seen in their secretion in strains deleted for Rv3877 (the ESX-1 machinery). Each of these two substrates has also been implicated in immunogenicity and possibly virulence of the organism in its host.

Conclusions

- iTRAQ reagent protein quantification strategy combined with the hybrid triple quadrupole linear ion trap technology in the 3200 QTRAP® system provided comprehensive and definitive data to normalize and characterize this secreted sub-proteome.
- This strategy enabled the identification of dozens of novel secreted proteins from *M. tuberculosis* and demonstrated the existence of novel secretion pathways
- ProteinPilot™ Software couples the highest level of protein identification with the most accurate MS/MS quantification for high integrity results.

References

- QTRAP® Systems for Differential Protein Expression using Multiplexed Isobaric Tagging Reagents. AB SCIEX Technical Note 2190211-01.
- C-terminal Signal Sequence Promotes Virulence Factor Secretion in *Mycobacterium tuberculosis*. DiGiuseppe PA et al. (2006) *Science* 313, 1632-1636.

For Research Use Only. Not for use in diagnostic procedures.

© 2011 AB SCIEX. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.-

Publication number: 3620211-01

Headquarters
110 Marsh Drive | Foster City CA 94404 USA
Phone 650-627-2600

International Sales
For our office locations please call the division
headquarters or refer to our website at
www.sciex.com