Target Identification Studies: Application of a New LC-MS and Data Analysis Software System to Identify Drug-Induced Changes in Mycobacteria

Minerva A. Hughes[1], Jeffrey C. Silva[2], Craig A. Dorschel[2], Scott J. Geromanos[2], Craig A. Townsend[1]

[1] Johns Hopkins University, Baltimore, MD; [2]Waters Corporation, Milford, MA

Introduction

Tuberculosis remains a major infectious disease world-wide with more that 2 million deaths annually.[1] Isoniazid (INH), Figure 1, was first introduced in the mid 1950's and currently represents a key first-line therapeutic for antituberculosis therapy. INH is a pro-drug requiring *in vitro* activation by the catalase-peroxidase KatG to its biologically active form.[2a] *Mycobacterium tuberculosis*, the causative agent for TB, is exquisitely sensitive to INH in part due to a dysregulation in the peroxidative stress response.[2b] Though widely accepted as a cell wall biosynthesis inhibitor targeting enzymes of the fatty acid synthase type II system, the complete cellular dynamics surrounding INH toxicity are still unclear.

In recent years, clinical isolates resistant to INH, as well as other first-line therapeutics (MDR-TB), have emerged.[2c] Inadequate treatments for MDR-TB and the apparent synergy in individuals co-infected with HIV necessitate efforts to develop new preventative and treatment regimens. Current approaches are largely mediated by improving existing therapeutics through a comprehensive analysis of affected metabolic pathways and conclusive identification of the target molecule(s). Using proteomics as a platform for understanding drug-induced alterations, insights into affected metabolic pathways, mechanisms for drug resistance or identification of novel enzymes are possible. This poster focuses on the application of a new LC-MS[3,4] based strategy for qualitative and quantitative proteomic analysis to monitor protein expression patterns in *M. tuberculosis* var. BCG following exposure to INH. Cross-validation of the results was made by comparisons with existing genomic and proteomic data.[5a-c]

Figure 1. *Chemical structure of isoniazid*

Protein Preparation

Cell Culture

M. bovis BCG was grown to mid-log phase (OD_{600} 0.4-0.5) in Middlebrook 7H9 medium supplemented with 10% ADC and 0.025% Tween80. Cultures were treated with INH (1 µg/mL final) or DMSO for 6 h, harvested, washed and lysed with 50 mM NH_4HCO_3, 5 mM EDTA and 0.25% RapiGest to a final protein concentration of 5-10 mg/mL.

Ammonium Sulfate Fractionation

Approximately 50 mg of total protein was fractionated into five portions using finely ground NH_4SO_4 according to standard protocols: 0-30%, 30-40%, 40-45%, 45-50%, 50-90%. Pellets were re-solubilized in minimal buffer (50 mM NH_4HCO_3, 5 mM EDTA) and dialyzed overnight against the same buffer.

Digestion and LC-MS analysis

Digests

• Reduction with 5 mM DTT at 60 °C for 30 min followed by alkylation with 15 mM iodoacetamide for 30 min in the dark at room temperature.

•Initiated with trypsin (1:75 w/w trypsin:total protein) and incubated at 37 °C overnight. The mixture was diluted with an equal volume of 50 mM NH_4HCO_3 and analyzed in 10 µL injections (approx. 6 µg total protein).

LC-MS

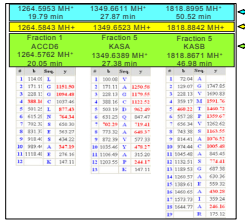
• Waters CapLC/Micromass Q-Tof Ultima API Mass Spectrometer equipped with a NanoLockSpray source and operated at approximately 12,000 resolving power. Used [Glu¹]-Fibrinopeptide B] as the accurate mass standard.

• Waters NanoEase Atlantis C_{18} column, 300 µm x 15 cm; mobile phase: A = 1% acetonitrile in water, 0.1 % formic acid, B= 80% acetonitrile in water, 0.1% formic acid. Gradient: 6% to 40% B over 100 min at 4.4 µL/min, followed by 10 min rinse (99%B) and 20 min re-equilibration at initial conditions.

Qualitative and Quantitative Proteomic Analysis

Protein identifications were made using the peptide mass fingerprint (PMF) in conjunction with the relative quantitation as well as the MS^E data for a given precursor peptide. Additional identifications were made possible by pre-fractionation of the soluble protein mixture with ammonium sulfate to reduce the complexity and dynamic range of proteins. A protein database was generated to correlate those ions with additional MS^E data in the fractionated samples with those in the un-fractionated sample with PMF data only.

Figure 4 *illustrates examples of a peptide ion identified in the unfractionated sample based on PMF and the corresponding relative quantitation for AccD6, KasA, and KasB. These identifications were sequence validated using the MS^E data obtained from the fractionation sample.*



Data Acquisition

The mass spectrometer is configured to acquire data in alternating elevated (28 – 35 eV) and low (10 eV) energy modes to allow sufficient sampling across all detected peptides to preserve the chromatographic integrity of the detected precursors and associated fragment ions. Each ion detection represents a three dimensional arrangement of accurate-mass, retention-time, and signal intensity that are de-isotoped and charge-state reduced using Water's Protein Expression System Informatics software. The time-resolved LC-MS data is organized to provide an inventory of every detected precursor with their associated fragment ions.

Figure 2. *Illustration for alternating scanning mode of data acquisition*

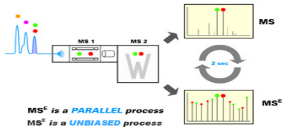


Figure 3. *(A) The LC base peak intensity chromatograms for the elevated and low energy channels illustrating the complexity of the samples. Once clustered, the peptide components are normalized and filtered to remove those peptides replicating only once across both conditions. The level of reproducibility is illustrated in (B) for the control sample. The INH sample exhibited a similar behavior. Although only 70% of the detected AMRTs replicated in at least 2 out of 3 injections, these peptides account for 98% of the total intensity. Additional data presented in (C) demonstrates the level of chromatographic reproducibility, mass precision, and intensity variation that allows cross-condition correlation.*

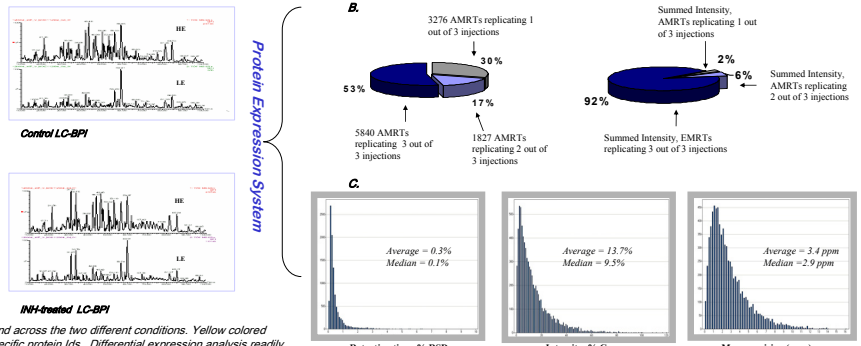
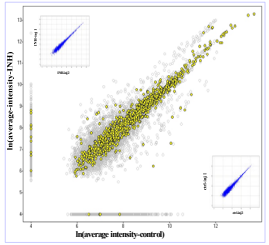
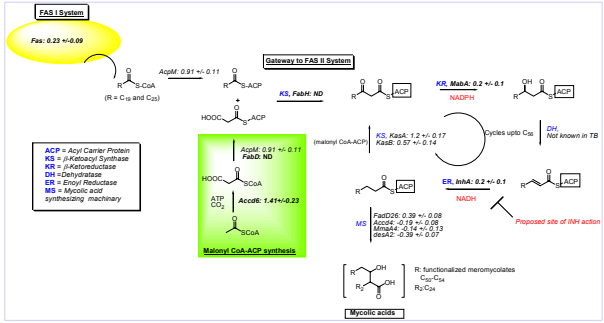


Figure 5. *Off diagonal plots illustrating the sample variation within replicates and across the two different conditions. Yellow colored AMRTs (accurate mass-retention time components) have been assigned to specific protein Ids. Differential expression analysis readily identified key enzymes involved in mycolic acid biosynthesis and the toxicity associated with INH. Illustrated is a representation of the fatty acid biosynthesis pathway and the effects on protein expression observed in this study. Values are listed as the ln (INH-treatment/ Control) +/- 95% confidence interval. For the purpose of this study, a ln(ratio) of < -0.4 or > 0.4 is considered differentially expressed.*





Conclusions

Analysis of a complex biological sample readily identified key proteins involved in the mechanism of action for INH. This new LC-MS method does not employ any labeling strategy and rapidly affords a comprehensive view of the proteome in a single experiment. Simultaneous acquisition of the precursor and fragment ions allows one to digitally preserve the data for further analysis. The mass accuracy of the precursors and associated fragment ions allows subsequent peptide identifications using stringent mass accuracy searching parameters.

References

1. World Health Organization: http://www.who.int/tb

2. (a) Johnsson, K.; Schultz, P.; *J. Am. Chem. Soc.* 1994, 116, 7425. (b) Zhang, Y.; Dhandayuthapani, S.; Deretic, V.; *Proc. Natl. Acad. Sci.* 1996, 93, 13212.

3.Bateman, R. H; Hoyes, J. B., U. K. Patent 2, 364, 168 A, 2002.

4. Silva, J.C et al. *Anal. Chem.* 2005, 77, 2187

5. (a) Waddell, Simon J. et al. *Tuberculosis.* 2004, 84, 263. (b) Boshoff, Helena I.M. et al. *J. Biol. Chem.* 2004, 279, 40174. (c) Betts, Joanna C. et al. *Antimicrob. Agen. Chem.* 2003, 47, 2903

Acknowledgements
National Institutes for Health
(AI 054842)

720001236EN