

Characterization of the Proteome of Tuberculin Using a Combination of LC-MALDI-MS/MS and LC-ESI-MS/MS

Prenni J.E., Dobos K.M., Cho Y.S., Ben-Hur A.A., Ryan P.W., Belisle J.T
Colorado State University, Fort Collins, CO

Introduction

Tuberculin, also known as purified protein derivative (PPD), is widely used as a diagnostic test for tuberculosis despite the fact that the protein composition is unknown. In addition, there are multiple preparations of PPD in use throughout the world. The work presented here represents two specific goals of an effort to characterize PPD: 1) Evaluate the use of MALDI-MS/MS as a complement to a previous ESI-MS/MS investigation of the PPD proteome and 2) Develop an assay to establish and monitor variability among the PPD preparations currently in use in different parts of the world.

Complementary Proteome Analysis

Sample Preparation

100 μ g of standard PPD (S2) from the Food and Drug Administration was reduced with dithiothreitol and alkylated with iodoacetamide by standard methods. Trypsin digestion was performed in 100 mM ammonium bicarbonate with incubation at 37°C for 16 hr.

ESI-MS/MS

Off-line strong cation exchange chromatography (SCX) was performed on the digested PPD with KPO_4 as the starting buffer and 0.5M KCl for the salt cuts. Six fractions were collected. Each of these fractions was further separated by reverse-phase HPLC using a polysulfoethyl column (The Nest Group, Inc., Southborough, MA) with an Alliance HPLC system (Waters Co., Milford, MA). Individual peaks were collected; on average 25 fractions from each SCX fraction. Each of these were injected onto an additional reverse-phase column for MS analysis on a Finnigan LCQ electrospray ionization (ESI) mass spectrometer (Thermo Electron Co., West Palm Beach, FL).

MALDI-MS/MS

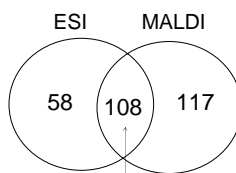
Off-line SCX separation was performed as indicated for the ESI-MS/MS experiments. The six salt cuts were further separated by reverse-phase HPLC using a C18 PepMap 100 (LC-Packings, Sunnyvale, CA). The effluent from each reverse phase separation was spotted robotically onto a MALDI target with collection every 15 seconds. The data represent a combination of PPD SCX fractions as well as direct analysis of digested PPD. MALDI-MS/MS experiments were conducted with a Bruker Ultraflex TOF/TOF (Billerica, MA).

Data Analysis

The data from both ESI-MS/MS and MALDI-MS/MS experiments were searched against the mycobacterium tuberculosis protein sequence database using the Mascot, SEQUEST and X! Tandem search engines. The results from the three database searches were combined in Scaffold (Proteome Software, Portland, OR). Additional statistical analysis was performed in Scaffold to further increase the confidence of the reported protein identifications¹⁻². For example, the ESI data set consisted of over 20,000 MS/MS spectra and the initial SEQUEST search yielded ~3000 protein identifications. After subsequent database searches by Mascot and X! Tandem coupled with the statistical analysis in Scaffold this number was reduced to 166 high confidence identifications.

Results

The combination of the results from MALDI-MS/MS and ESI-MS/MS analysis yielded a total of 283 proteins identified with high confidence. Specifically, 108 proteins were identified by both ionization techniques, 58 proteins were unique to the ESI experiments and 117 proteins were unique to the MALDI experiments.



Proteins Identified by both ionization techniques

Figure 1. Results of the complementary analysis of PPD by ESI-MS/MS and MALDI-MS/MS.

In addition to the identification of a larger number of proteins in the sample, the use of complementary ionization techniques resulted in the ionization of unique peptides which generated an increase in the sequence coverage of proteins identified by both techniques. On average, we found an increase in surface coverage of approximately ~10%.

Rv0129c
MTFFEQVRRLL RSAATTLPR LALAAANDAVI VYGLVDTFGG PATAGAFSRP
GLPVEYLQVF SASMRDILV DFDDGGPHAV YLLDGLDAD DINGWDINTP
AFEEYDQGL SVIMPVGDG SFYTDWYQPS DSRRDRTYR WELFKLRMP
AWLGANKGVE PTGNAAVGLS MGGGALILA AYPGQFPYA ASLSGFLNPS
EGWVPTLIGL AMNDGGYNA NENWGFSDP AWKRDPMVG IPRLVANNTR
HWVYGGRTF SDLGGDNIQA KLEGLTLRL RGTFTDYAA DGGRGVYFV
PFGNTHSWFY WNEGLVANCA DIGHVLNGAT PPAAPAAFAA

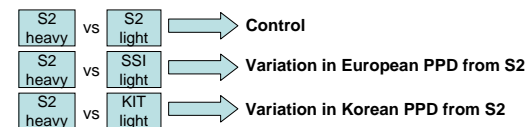
Figure 2. PPD protein that was identified by both ESI-MS/MS and MALDI-MS/MS. The red boxes indicate the peptides found by ESI and the yellow highlighted amino acids found by MALDI. The confidence of the identification greatly increased because both the sequence coverage and the number of unique peptides are doubled when the results from the two experiments are combined.

Determining Variation in PPD Preparation

Sample Preparation and MALDI-MS/MS

100 μ g of standard PPD (S2) from the Food and Drug Administration was labeled with the ICAT reagent according to manufacturer's protocol (Applied Biosystems, Foster City, CA). Additional PPD preparations included the European PPD (SSI) and the Korean PPD (KIT). All ICAT experiments were conducted using the Bruker Ultraflex TOF/TOF. The samples were separated on a reverse-phase column and the effluent was robotically spotted on a MALDI target plate with collection every 20 seconds of a 40 min linear gradient.

Experimental Design



Data Analysis and Results

The MS spectra for each of the sample sets was mined for the peak masses and ratios of all ICAT pairs. A pair was defined by the mass of the light ICAT peptide and the absolute value of the log of the ICAT ratio. The peaks were clustered and a statistical analysis was performed to determine the variation of SSI and KIT from the standard S2 preparation. The control experiment was used to determine any variation due to experimental factors and all other data sets were compared against the control. The analysis resulted in p-values of 1×10^{-6} for the Korean preparation and 5×10^{-6} for the European preparation.



Figure 3. Lane B represents the control experiment, S2 vs S2. Lane A and C represent S2 vs SSI and S2 vs KIT, respectively. All ratios are plotted as the absolute value of the log. Therefore, any variation from a ratio = 0 shows a variation between PPD preparations. In the color scale, red areas represent ratios near 0 and an increasing yellow color indicates variation away from a ratio = 0.

Conclusions

The results of this study illustrate the advantage of using complementary ionization sources in the shotgun analysis of complex mixtures. A significant number of unique proteins were identified by each technique which increased the overall number of proteins identified. In addition, for proteins identified by both ESI and MALDI the average sequence coverage increase by ~10%. Also, the use of multiple search engines and the statistical analysis available in the Scaffold software generates results of high confidence. The second portion of the study resulted in the development of an assay that can determine if there is a qualitative difference PPD preparations. Significant variation was found in both the European and Korean PPD preparations as compared to the FDA standard. Elucidation of the PPD proteome and the ability to assay PPD preparation will provide insights into the mechanistic behavior of PPD and aid in the development and standardization of a new, more specific, diagnostic reagent for tuberculosis.

References

- Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem.* 2002 Oct 15;74(20):5383-92.
- Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem.* 2003 Sep 1;75(17):4646-58.

Acknowledgments

NIH contract HHSN266200400091c