

## Top-Down Identification of Bacterial Intact Protein Expression Profile Markers

Melinda A. McFarland; John H. Callahan; Denis Andrzejewski; Rebecca Bell; Steven M. Musser  
*FDA/CFSAN, College Park, MD*

### Novel Aspect:

The addition of a top-down identification platform to intact protein profiling for bacterial strain and toxin differentiation.

### Introduction:

Faster and more effective methods for bacterial assessment are becoming increasingly relevant. Intact protein expression profiling by LCMS has proven to be a powerful tool for bacterial strain differentiation, virulence assessment, and antibiotic resistance monitoring. However, transfer to field usable assays or monitoring of drift in these species requires identification of marker proteins. Here we expand our protein profiling platform to include semi-automated top-down identification of bacterial marker proteins and apply these methods to thermally regulated *Shigella* virulence factors.

### Methods:

*Shigella* proteins are extracted in 50:45:5 acetonitrile, water, formic acid using the **Barocycler pressure cycler (Pressure Biosciences)**. On-line separation of proteins is performed on an Agilent 1100 HPLC with a Prosphere P-HR column (150mm x 2.1mm). Online fraction collection and LC eluent and fraction infusion into the Orbitrap mass spectrometer (ThermoFisher) are carried out via the TriVersa Nanomate (Advion). Automated deconvolution, and extraction of molecular weight, retention time, and intensity from LCMS data is performed with ProteinTrawler (BioAnalyte). ETD and CID spectra from infused fractions are searched via ProSight PTM ([prosigthptm.scs.uiuc.edu](http://prosigthptm.scs.uiuc.edu)).

### Preliminary Results:

Previous qTOF data shows good protein level separation and reliable extraction of protein molecular weight and retention time profiles. *Shigella* grown above and below 37° show markedly different protein expression profiles. Transfer of this method to the Orbitrap shows a truncated molecular weight range, with proteins larger than 30 kDa being inefficiently detected, poorly resolved or not detected. Detection in the LTQ region of the instrument offers a wider mass range but lower resolving power. Coupling of the LCMS to online fraction collection allows for decoupling of MSMS from intact protein molecular weight profiling. Infusion of collected fractions provides for collection of many more scans resulting in higher mass accuracy on larger proteins not well resolved on an LC time scale. Top-down MSMS of these fractions detected at high resolution offers more reliable sequencing and greater sequence coverage. CID and ETD coupled with the low sample consumption of chip-based nanospray facilitates extensive scan summing which improves detection of low abundance fragment ions for improved coverage. This is particularly beneficial for differentiation of post-translational modifications and new genetic mutations distinguished by mass shifts on previously unidentified proteins. Data will be used to populate a library of identifications for bacterial accurate mass-retention time pairs to narrow the search space for distinction of relevant marker targets.