

## Development of a Web-based Top-Down Protein Identification Tool

Christian Heckendorf; Roger Theberge; Catherine Costello; Mark McComb

Boston University School of Medicine, Boston, MA

### Abstract

Top-down mass spectrometry-based proteomics has emerged as an approach for protein characterization which preserves labile post-translational modifications and offers full protein sequence coverage. However, effective utilization of top-down MS/MS is hampered by the lack of availability of readily accessible data interpretation tools. Here we describe the development of a web-based extension of BUPID Top-Down (Boston University Protein Identifier Top-Down) to identify proteins in top-down MS/MS spectra.

BUPID Top-Down uses a sequence tag approach to identify potential protein candidates from a protein sequence database. Tags are generated by comparing the mass difference between pairs of experimental masses obtained from MS/MS. These tags are then used to identify candidate identifications in a protein sequence database. The algorithm is implemented as a standalone program written in C and is designed to be run on Linux/Unix servers. A web front-end allows it to be accessible through standard HTTP web browsing. The software was also designed to take advantage of server grade hardware to efficiently process data. A client/server model was chosen for the back-end to reduce overhead and system strain associated with simultaneous searches.

An overview of the algorithm, representative results and a comparison to other tools will be presented. The ultimate goal for this work is to develop a flexible top-down protein identification tool that can correctly identify proteins from a database search and assign the presence of post-translational modifications.

Acknowledgements: This project was funded by NIH/NHLBI N01 HV-28178, NO1 HV-00239, NIH/NCRR P41 RR10888/GM104603, S10 RR020946 and S10 RR025082.