Peptide Database Search Strategies to Improve Peptide Identifications Using High Resolution Mass Spectrometry

Edward J. Hsieh, Michael R. Hoopmann, Brendan MacLean, Michael J. MacCoss
Department of Genome Sciences, University of Washington, Seattle, WA 98195-5065

Overview

In a previous proteomics, the analysis of tandem mass spectrometry data from peptides can benefit greatly from high mass accuracy measurements. We have developed a software program, Bullseye, that uses high resolution precursor scans to assign precursor monoisotopic masses to tandem spectra more accurately than what is determined by the instrument’s onboard computer. We examined two different search strategies to determine the most effective manner to utilize this high mass accuracy information for peptide identifications. The first method constrained the database search to those peptides in a narrow mass window around the precursor mass. The second method expands the search to a wide mass window, but then discards peptide identifications that fall outside a narrow mass tolerance.

Results

We present results from the evaluation of two different database search strategies to determine the most effective manner to utilize high mass accuracy measurements. We developed a software tool, Bullseye, which uses the result from the high resolution precursor mass to generate an approximate mass and charge state of the peptide. Bullseye was used to determine the mass of the precursor ion in MS/MS fragmentations. Bullseye uses this result to determine the best approach to assign peptide identifications to high resolution MS/MS spectra. For each MS/MS spectrum all possible matching precursor PSMs can be determined. We present results from the evaluation of two different database search methods to determine the most effective approach for the use of high mass accuracy. The first method limits the precursor mass in the database search to only those ions that are within a very narrow mass window surrounding precursor monoisotopic mass. We have previously used this approach to search FT-ICR followed by a data processor to assign peptide identifications with a narrow mass tolerance. We have also examined the potential database search speed increases obtained by reducing the number of peptide searches.

Methods

Sample Preparation

- Tryptic digestion of ammonium bicarbonate soluble component of cell lysate from yeast
- Data processing scheme for Bullseye

Instrumentation

- High performance liquid chromatography (HPLC) system
- FT IC-TOF Ultra
- FT Ultra

Data processing with Bullseye

- Bullseye uses the result from the high resolution precursor mass to generate an approximate mass and charge state of the peptide.
- Database searching is constrained to a narrow mass tolerance window.
- MS/MS spectra were searched against a S. cerevisiae protein database.

Data processing with Bullseye

- Bullseye uses the result from the high resolution precursor mass to generate an approximate mass and charge state of the peptide.
- Database searching is constrained to a narrow mass tolerance window.
- MS/MS spectra were searched against a S. cerevisiae protein database.

Acknowledgements

Support for this work was provided in part by National Institutes of Health grants R01 HL068123, P41 RR011823, R01 GM082683, and R01 GM082683. Edward J. Hsieh was supported by National Institute of Health grant T32 HG00035.