Software and Algorithm for Differential Characterization of Post-Translational Modifications

Vivek N. Bhatia, David H. Periman, Catherine E. Costello, Mark E. McComb
Cardiovascular Proteomics Center, Boston University School of Medicine, Boston, MA

PROJECT AIMS

Protein post-translational modifications (PTMs) play a critical role in normal cellular function. Changes in PTMs are known to affect the mechanisms of disease. Thus, studying PTM identification and localization is a significant problem in contemporary biomedical research. To facilitate differential PTM characterization, we have developed STRAP PTM (Software Tool for Rapid Annotation of PTMs), for automated analysis of global changes in PTMs on proteins.

METHODS

STRAP PTM

- utilizes data from large-scale proteomics experiments
- performs multi-way differential analyses by scoring and ranking proteins, peptides and PTM identifications
- provides a visual means to identify important PTMs that are differentially observed between multiple sample groups
- uses a unique algorithm to find differential PTMs

BACKGROUND

MS-based proteomics typically yields thousands of protein and post-translational modification (PTM) identifications in an experiment. Post-identification, information about PTMs must be interpreted to gain insight into their biological roles in disease. While there is a great interest in PTMs due to their significant roles in disease, there are few tools for differentially comparing PTMs across multiple different sample groups. An example workflow is shown in Figure 1.

THE PROTEOMIC WORKFLOWS

RESULTS

PTM COUNTING VERSUS. LABEL-FREE QUANTIFICATION

STRAP PTM was used to analyze data from a labelfree, LC-MS-based oxidant sensitivity study, in which a target protein containing several suspected redox-sensitive residues was subjected to treatment with increasing concentrations of paranitrite (PN). In this example, we report on the ligand fragment CD40L_HUMAN, which has been shown to be sensitive to oxidation. The identification of the resulting oxidized species was determined by trypan digestion, LC-MS/MS and PTM- tolerant database searching. Relative abundances were quantified by peptide ion abundances or by PTM counting. Among the different modifications detected, we observed various forms of cysteine oxidation (Figure 2) and nitration. Nitration showed distinct and differing degrees of formation with respect to oxidation, as depicted in the label-free analysis shown in Figure 3. We observed increases in both nitrosotyrosine and dinitrotyrosine formation as PN concentration was increased.

Nitration and other PTMs were rapidly elucidated using STRAP PTM, as shown in Figure 4. The findings were comparable to the label-free LCMS results.

RESULTS II

Figure 3: PTM expression trend for nitrination and dinitration of the CD40L peptide (numbered) THPMYKTVNLGK. The amount of the unmodified peptide decreased with increasing PN concentration. The singly nitrated version of this peptide increased as CD40L was exposed to lower concentrations of PN, but decreased in abundance at higher concentrations of PN as it was converted to the dinitrotyrosine form. The dinitrotyrosine bearing peptide increased with increasing PN. The abundance values, obtained from Progress LCMS, were normalized to the maximum value in each series.

RESULTS III

Following LC-MS/MS, RAW files were processed through Proteome Discoverer (Thermo) to obtain MGF files, which were used in the Mascot database search. Resulting .dat files were processed through the T revert to pepXML analysis of PTMs, which are then imported into STRAP PTM for differential PTM analysis.

The most striking feature of this processing, shown in Figure 6, was that in under 5 minutes, STRAP PTM afforded a list of overlapping proteins observed between conditions, PTM maps of these proteins, and PTM scores to rank these proteins based on their differential expression of PTMs. This experimental data originated from 3 experimental groups, having 3 samples each. The LC-MS/MS analysis yielded a total of 6034 spectra, and 313 PTM assignments overall.

CONCLUSIONS

STRAP PTM performs differential comparison between files and/or sample groups, yielding differentially observed PTMs. One of the products of this comparison is a PTM map that overlays PTMs on the primary sequences of the parent proteins. This visualization aids in data interpretation, as it conveniently illustrates experimentally observed differences in modifications among several sample groups. Moreover, the scoring function built into STRAP PTM emphasizes PTMs of interest. The use of PTM maps and the scoring function together allow for global proteome-wide PTM analysis in a manner more systematic and faster than manual analysis by an expert. STRAP PTM accelerates differential PTM analysis that is otherwise extremely laborious when performed manually on complex data sets. By using MS- search results obtained from the open standard protXML format, the user is not restricted to a specific search strategy or type of instrumentation.

Here, we demonstrated that PTMs detected by STRAP PTM via differential analysis of CD40L oxidation were comparable to those observed via label-free LCMS. By illustrating an application to a proteome-wide analysis of a complex data set, we also demonstrated that differential PTMs may be mapped with relative ease in a substantially shortened time scale.

CURRENT AND FUTURE WORK

- Review and validate scoring algorithm to better mimic manual analysis by an expert
- Improve visualization of data, with a more informative PTM map and alternative forms of data representation
- Improve user interface to make data manipulation and interrogation more flexible

ACKNOWLEDGMENTS

- NIH-NCTR grants P41 RR010888, S10 RR020946, S10 RR025082
- NIH-NCRR contracts N01-HV-28175 and N01-HV-28239
- Special thanks to the members of the CPCR and CBMS