Software for Differential Characterization of PTMs: Approaches in Data Acquisition and Processing

Jean L. Spencer, Vivek N. Bhatia, Amanuel Kehasse, Stephen A. Whelan, Christian F. Heckendorf, Catherine E. Costello and Mark E. McComb

Cardiovascular Proteomics Center, Boston University School of Medicine, Boston, MA 02118

Overview

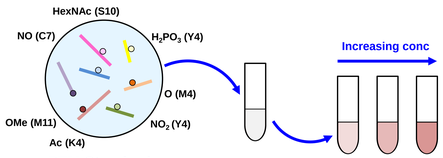
- **Purpose:** Investigate software that characterizes differential PTMs in complex data sets and determine the effects of data acquisition/processing parameters on its capability.
- **Methods:** In-house software (STRAP PTM) uses spectral counting and a novel scoring algorithm to collate and rank differential PTMs.
- **Results:** Trends in differential PTMs readily displayed and substantiated by quantitative analysis with results easily optimized by variation of acquisition/processing parameters.

Introduction

The identification of post-translational modifications (PTMs) across different states is critical for determining biomarkers and therapeutic targets in proteomics studies. Although software can identify peptides and by inference proteins, identification of PTMs is more challenging. Once PTMs are identified, few methods exist to analyze significant trends. We created **STRAP PTM (Software Tool for Rapid Annotation of Proteins: Post-Translational Modification edition)** to facilitate multi-sample comparison by collating and ranking PTMs. Here we explore the utility of STRAP PTM and the effects of data acquisition and processing parameters on its capability.

Materials

PTM Peptide Standards in Plasma



Methods

LC-MS/MS Analysis

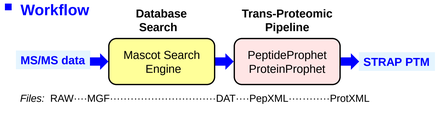
- nanoACQUITY UPLC system (Waters)
- TriVersa NanoMate ESI source (Advion)
- Q Exactive mass spectrometer (Thermo Scientific)

Label-Free Quantitative Analysis

- Progenesis LC-MS (v4.1; Nonlinear Dynamics)
- Mascot search engine (v1.0; Matrix Science)

STRAP PTM Analysis

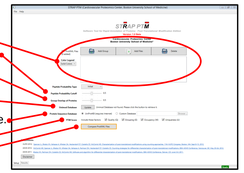
- **Software:** STRAP PTM (v1.0 beta) freely available at http://www.bumc.bu.edu/cardiovascularproteomics/cpctools
- **Workflow**



Methods

STRAP PTM Analysis

- **Setup Window**
  - Enter protXML data files.
  - Select peptide probability cutoff.
  - Choose protein group overlap.
  - Download PTM database.
  - Enter protein sequence database.
  - Select PTM score factors.
  - Execute analysis.



- **PTM Scoring**

PTM Score (S): Overall score for a specific PTM (m) on a specific site (i) of a specific protein (p) based on user-selectable factors relevant to the system (max = 100)

$$S_{mip} = 100 \times Q_{mip} \times G_{mip} \times W_{mip} \times U_{mip}$$

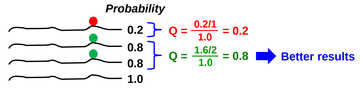
Quality · Grouping · Occupancy · Uniqueness

User-selectable factors

Quality (Q): Goodness of database search results for a specific PTM on a specific site of a specific protein (max = 1)

$$Q_{mip} = \frac{\langle P_{mip} \rangle}{\langle P^0_{ip} \rangle}$$

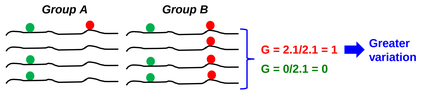
P = probability of modified peptides
P^0 = probability of unmodified peptides



Grouping (G): Variation of a specific PTM on a specific site of a specific protein across groups (max = 1)

$$G_{mip} = \frac{\sigma_{mip}}{\max \sigma}$$

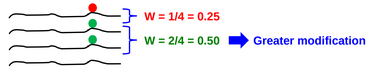
σ = std dev of counts
max σ = max σ of all proteins



Occupancy (W): Degree of modification of a specific site on a specific protein with a specific PTM (max = 1)

$$W_{mip} = \frac{N_{mip}}{\left(\sum_{m=1}^{M} N_{mip}\right) + \left(N^0_{ip}\right)}$$

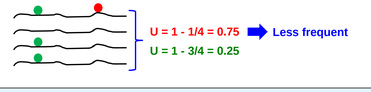
M = total PTMs
N = modified peptide counts
N^0 = unmodified peptide counts



Uniqueness (U): Rarity of a specific PTM on a specific protein (max = 1)

$$U_{mp} = 1 - \frac{\sum_{i=1}^{I} N_{mip}}{\sum_{m=1}^{M} \sum_{i=1}^{I} N_{mip}}$$

I = total sites
M = total PTMs
N = modified peptide counts



Results

STRAP PTM Results

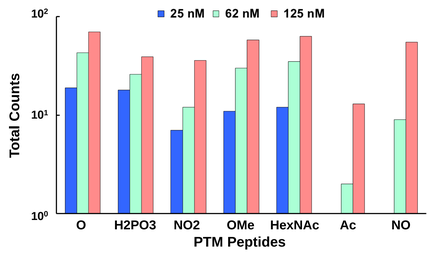
PTM Scores (top 7 protein hits)

Peptide	PTM	PTM Score QOW x 100	Total Counts 25 nM	62 nM	125 nM	Modified Forms	Other Forms	Scoring Factors Q	G	W	U
PS-280	O	77.9	19	43	70	132	0	0.901	0.865	1.000	0
PS-500	HexNAc	73.0	12	35	63	110	0	0.844	0.866	1.000	0
PS-580	NO	71.1	0	9	55	64	21	0.944	1.000	0.753	0
PS-412	OMe	67.6	11	30	58	99	0	0.844	0.801	1.000	0
PS-321	NO2	39.9	7	12	36	55	0	0.759	0.525	1.000	0
PS-301	H2PO3	35.5	18	26	39	83	1	1.000	0.359	0.988	0
PS-532	Ac	20.3	0	2	13	15	0	0.855	0.237	1.000	0

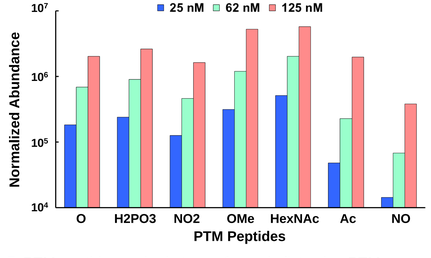
PTM Map (location of counts at 25 nM, 62 nM, 125 nM)

Peptide	PTM (Site)	Sequence
PS-280	O (M4)	SYSMEHFRWG
PS-301	H2PO3 (Y4)	DRVYIHPF
PS-321	NO2 (Y4)	DRVYIHPFHL
PS-412	OMe (M11)	RPKPQQFFGLM
PS-500	HexNAc (S10)	EAISPPDAASAAPLR
PS-532	Ac (K4)	DFNKFHTFPQTAIGV
PS-580	NO (C7)	EMFTYICNHIK

PTM Counts



Comparison with Progenesis LC-MS

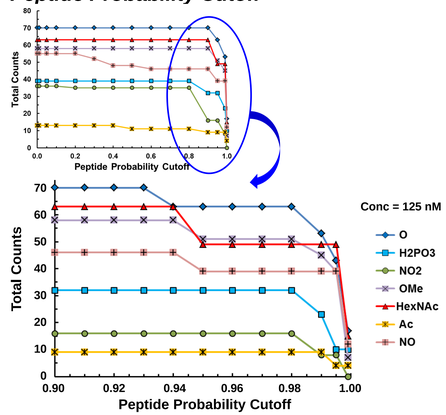


- PTM peptides ranked at top of protein list using PTM score.
- Trend of increasing counts with increasing concentration observed for all PTM peptides.
- Counting results compared favorably with quantitative results.

Results

Effect of STRAP PTM Settings

Peptide Probability Cutoff



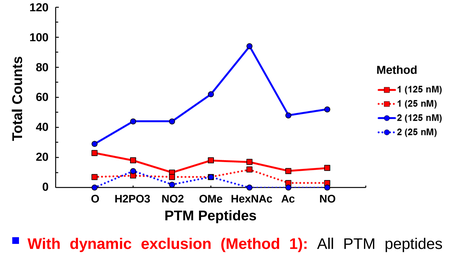
- Data quality improved by filtering out poorer search results with higher peptide probability cutoff.
- Better quality data balanced against falloff in counts ($p > 0.99$).

Effect of Instrument Settings

Method	Mode	Description	DE (s)	IT (ms)	List	MSX	Intensity	AGC
1	DD-MS2	Standard settings: DE on	4	60	--	--	8.3×10^4	5.0×10^5
2	DD-MS2	DE off	--	60	--	--	8.3×10^4	5.0×10^5
3	DD-MS2	List on; DE on	4	60	Yes	--	8.3×10^4	5.0×10^5
4	DD-MS2	List on; DE off	--	60	Yes	--	4.2×10^5	5.0×10^5
5	T-MS2	List on	--	60	Yes	--	--	5.0×10^5
6	T-MS2	List on; IT short	--	20	Yes	--	--	3.0×10^6
7	T-MS2	List on; IT short; multiplexing	--	20	Yes	7	--	3.0×10^6

DD-MS2 = data-dependent MS/MS; T-MS2 = targeted MS/MS
DE = dynamic exclusion; IT = injection time; List = inclusion list (PTM peptides)
MSX = multiplexing; AGC = automatic gain control

DD-MS2 and Dynamic Exclusion

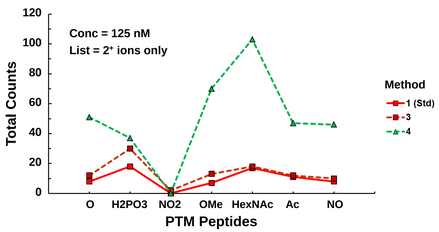


- **With dynamic exclusion (Method 1):** All PTM peptides detected at each concentration.
- **Without dynamic exclusion (Method 2):** More counts for all PTM peptides at highest concentration; 4 PTM peptides lost at lowest concentration (masked by higher peaks).

Results

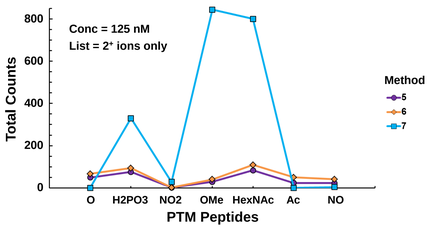
Effect of Instrument Settings

DD-MS2 and Inclusion Lists



- **With list and dynamic exclusion (Method 3):** Minor increase in counts over dynamic exclusion alone.
- **With list and without dynamic exclusion (Method 4):** More counts for all PTM peptides.
- Note: 3^+ precursor dominant ion for NO2 peptide.

T-MS2 and Multiplexing



- **With shorter injection time (Method 6):** More counts for all PTM peptides (higher duty cycle).
- **With multiplexing and shorter injection time (Method 7):** Inconsistent results from interference with other precursor ions.

Conclusions

STRAP PTM is a powerful counting approach that:

- Collates and ranks differential PTMs in complex data sets.
- Generates results with trends substantiated by label-free, quantitative analysis.
- Allows removal of poorer quality data for improved analysis.
- Provides a rapid means to optimize instrument settings for best results.
- Represents a fast and easy tool for semi-quantitative results.

Acknowledgments

- NIH-NHLBI contract HHSN268201000031C
- NIH grants P41 GM104603, R21 HL107993, S10 RR020946