

STRAP PTM: Differential Characterization by PTM Counting and Much More

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

Cardiovascular Proteomics Center, Center for Biomedical Mass Spectrometry, Boston University School of Medicine, Boston, MA 02118

Overview

- Purpose:** Software tool required to perform fast and easy characterization of global PTM changes in large proteomics data sets from LC-MS/MS experiments.
- Methods:** In-house application (STRAP PTM) developed to use spectral counting and novel scoring algorithm to identify and rank differential PTMs.
- Results:** Differential PTMs readily detected in diverse data sets and easily sorted for biologically relevant modifications.

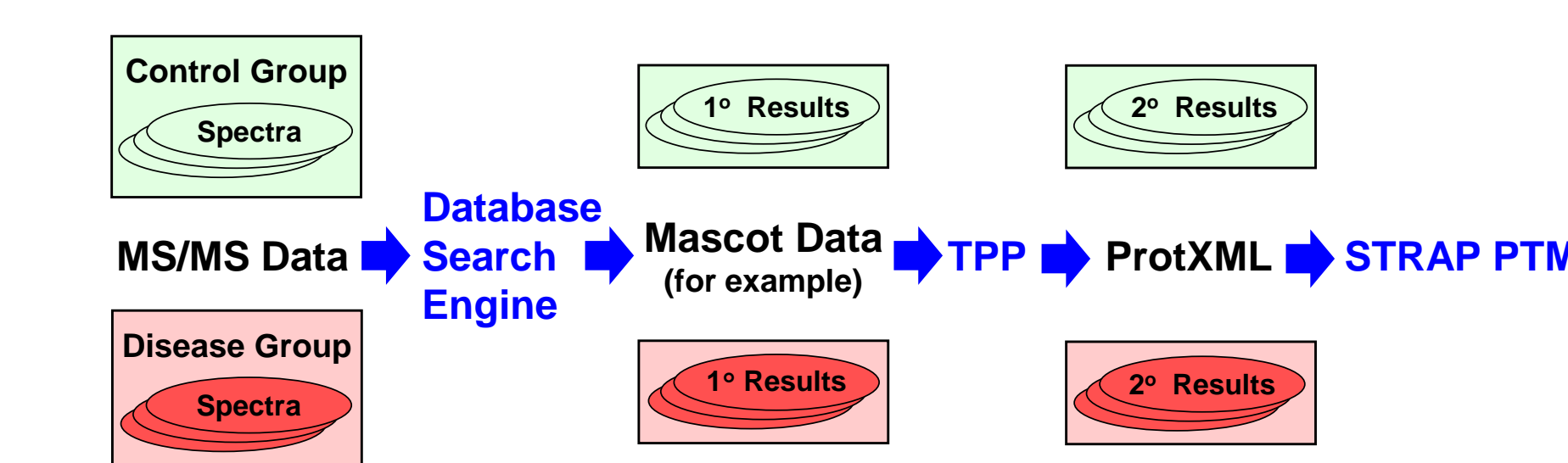
Introduction

Protein post-translational modifications (PTMs) play significant roles in disease pathology, and the identification of PTMs is an increasingly important component of proteomics and biomarker discovery. There are few tools for performing relatively fast and easy characterization of global PTM changes in large proteomics data sets and differential comparison of PTMs across groups. A software program called the **Software Tool for Rapid Annotation of Proteins: Post-Translational Modification edition (STRAP PTM)** was developed in response to this challenge. STRAP PTM uses a novel counting-based scoring algorithm that facilitates multi-sample PTM comparisons through collation and visualization. Here we explore the utility of STRAP PTM through various scoring factors and across different PTM/proteomics experiments.

Methods

STRAP PTM: PTM Counting

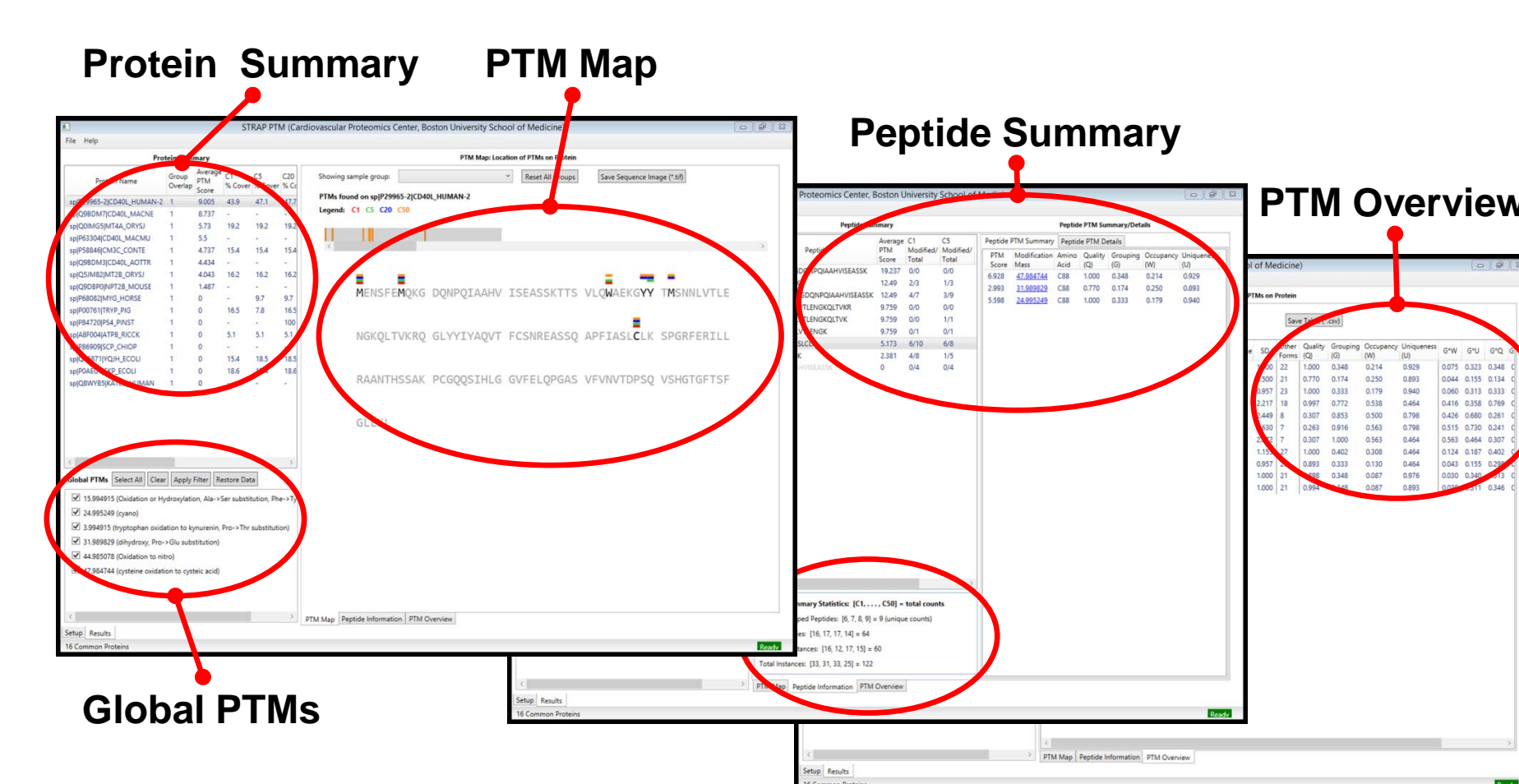
Workflow



Setup Window



Results Windows



Methods

STRAP PTM: 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 value = 100)

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

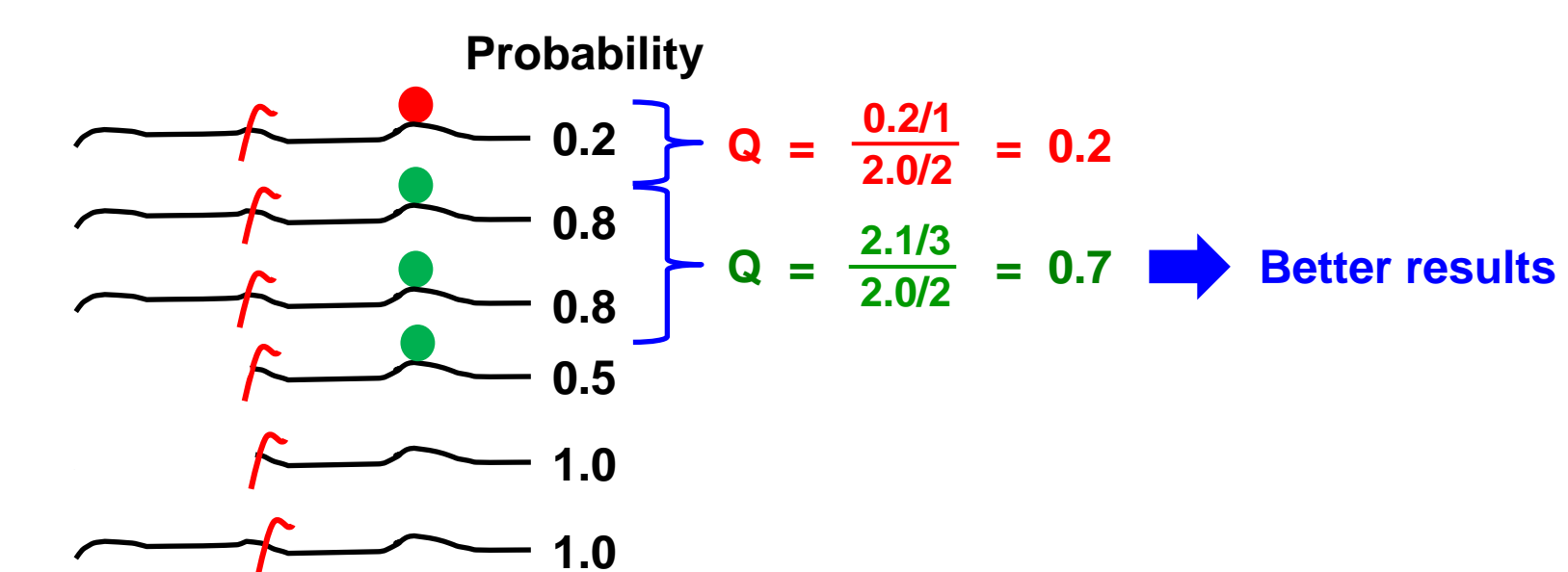
Quality Grouping Occupancy Uniqueness

User-selectable factors

- Quality (Q):** Goodness of database search results assigned to the MS/MS spectrum for a specific PTM on a specific site of a specific protein (max value = 1)

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

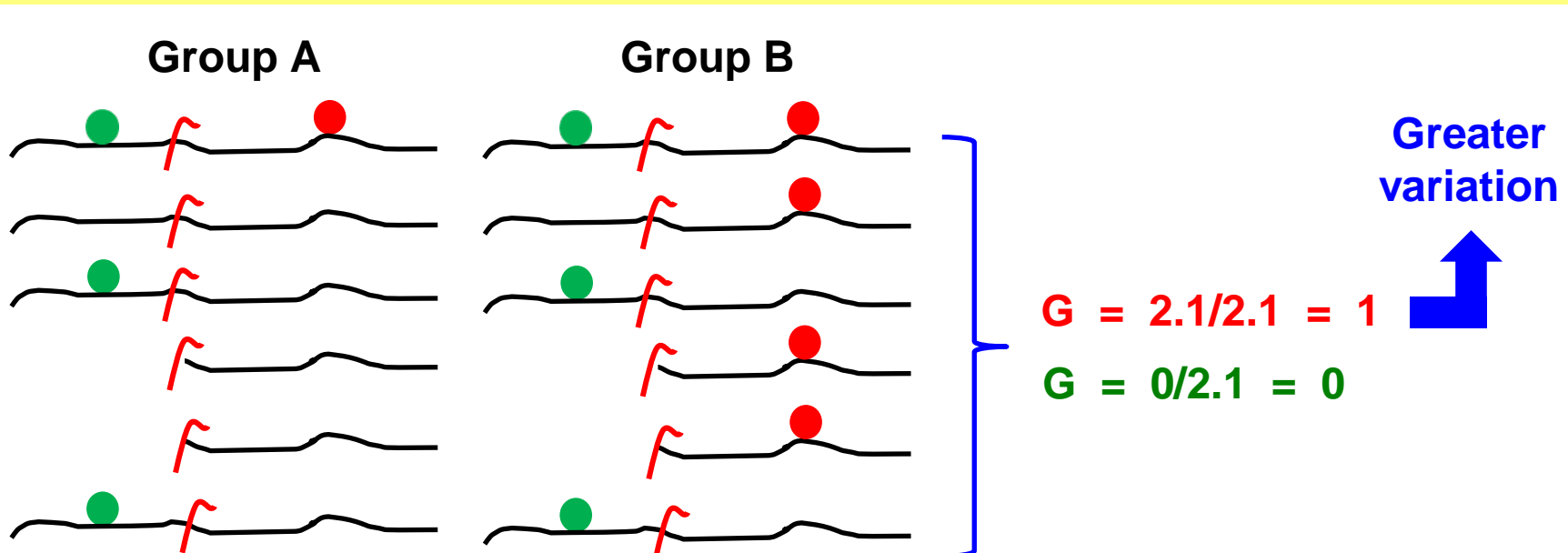
P = probability of modified peptides
P⁰ = probability of unmodified peptides



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

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

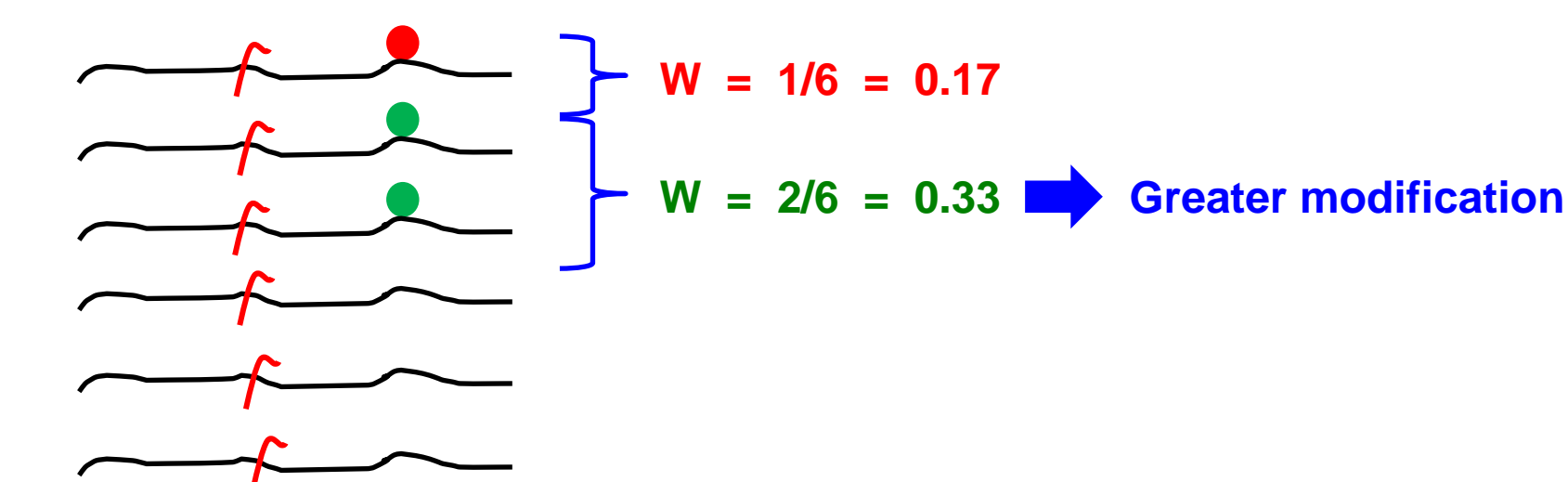
σ = std dev of counts
max σ = max σ of all proteins



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

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

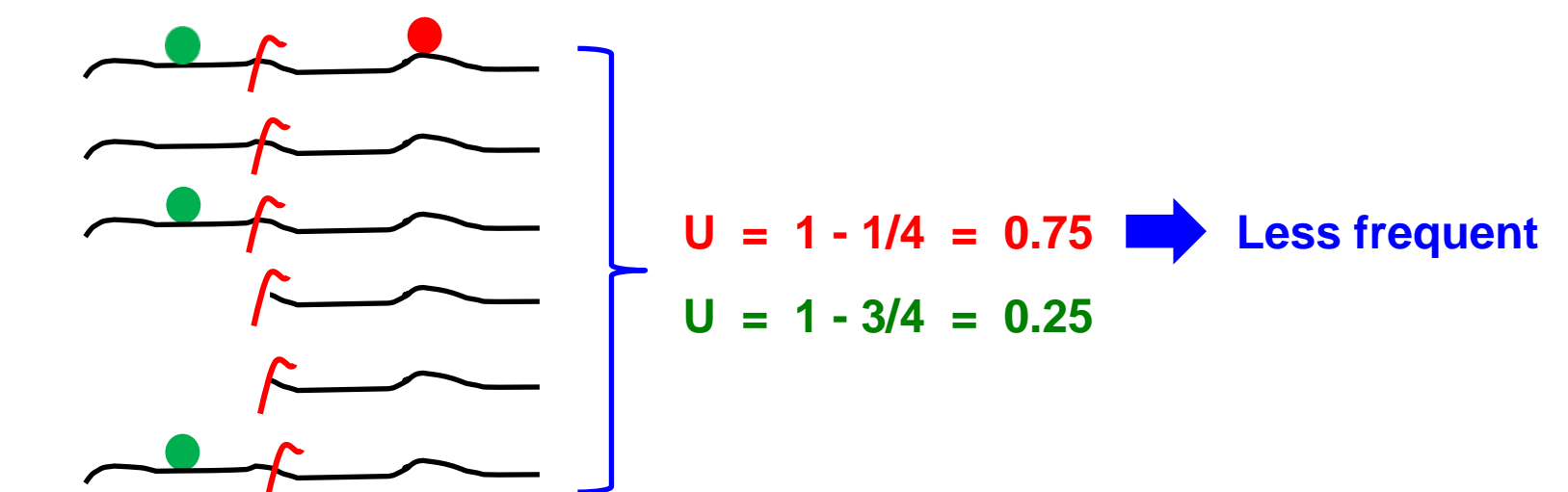
M = total PTMs
N = modified peptide counts
N⁰ = unmodified peptide counts



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

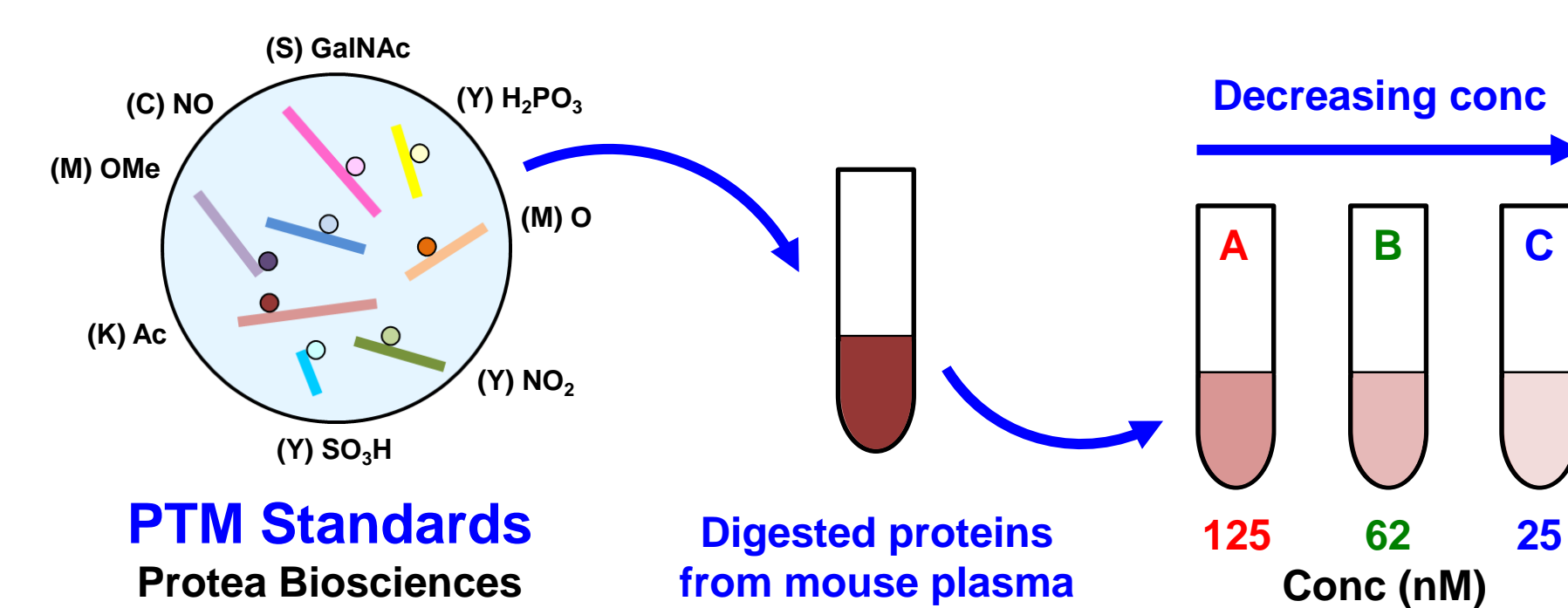
$$U_{mp} = 1 - \frac{\sum_{i=1}^I N_{mip}}{\sum_{m=1}^M \sum_{i=1}^I N_{mip}}$$

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



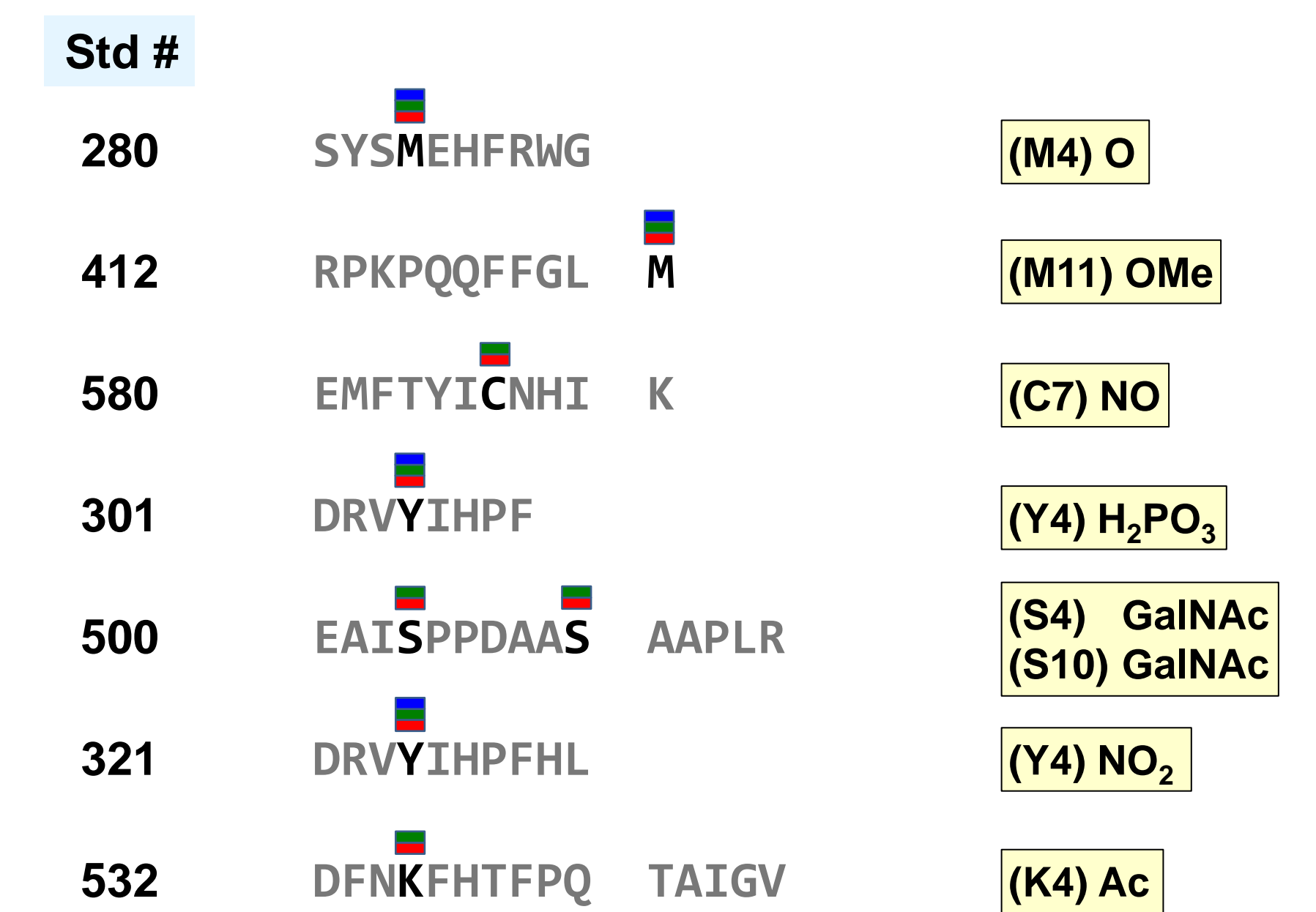
Results

Example 1: PTM Standards in Plasma



PTM Map (modified sites on peptide standards)

Legend: A (125 nM) B (62 nM) C (25 nM)

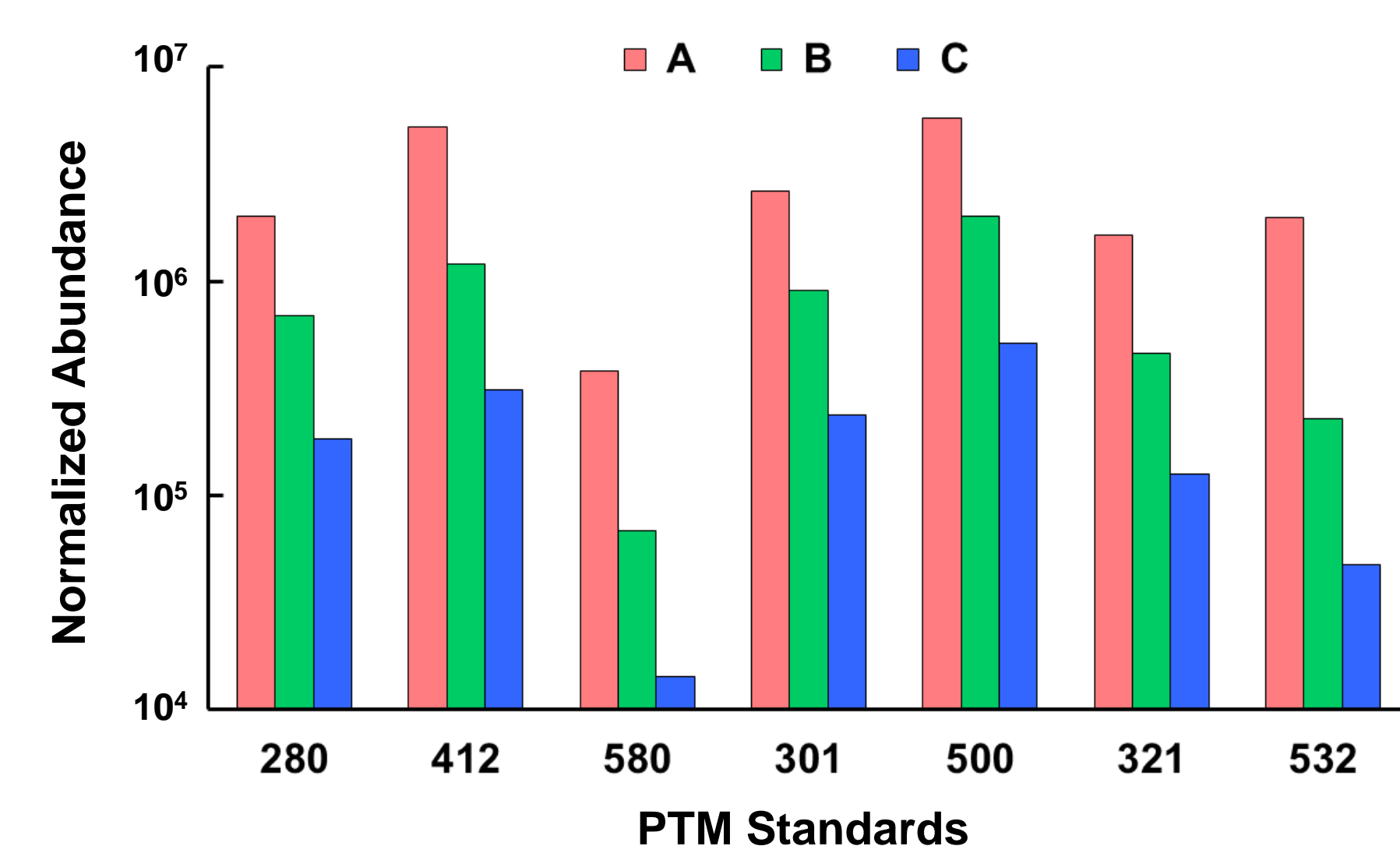


PTM Scoring (top 7 PTMs)

Std #	PTM Score	Mod Mass	Amino Acid	Counts			Mod Forms	Avg	SD	Other Forms	Scoring Factors			
				A	B	C					Q	G	W	U
280	85.3	15.994915	M4	67	49	18	134	44.67	24.79	0	0.963	0.885	1.000	0
412	77.5	14.01566	M11	51	31	4	86	26.67	23.59	0	0.920	0.842	1.000	0
580	68.8	28.990164	C7	52	8	0	60	20.00	28.00	11	0.815	1.000	0.845	0
301	36.5	9.966331	Y4	31	21	9	61	20.33	11.02	0	0.928	0.393	1.000	0
500*	58.5	243.079373	S4	40	18	0	58	19.33	20.03	12	0.987	0.715	0.829	0
500*	3.0	203.079373	S10	10	2	0	12	4.00	5.29	58	0.915	0.189	0.171	0
321	23.1	44.985078	Y4	23	15	1	39	13.00	11.14	0	0.580	0.398	1.000	0
532	17.5	42.010565	K4	11	2	0	13	4.33	5.86	0	0.835	0.209	1.000	0

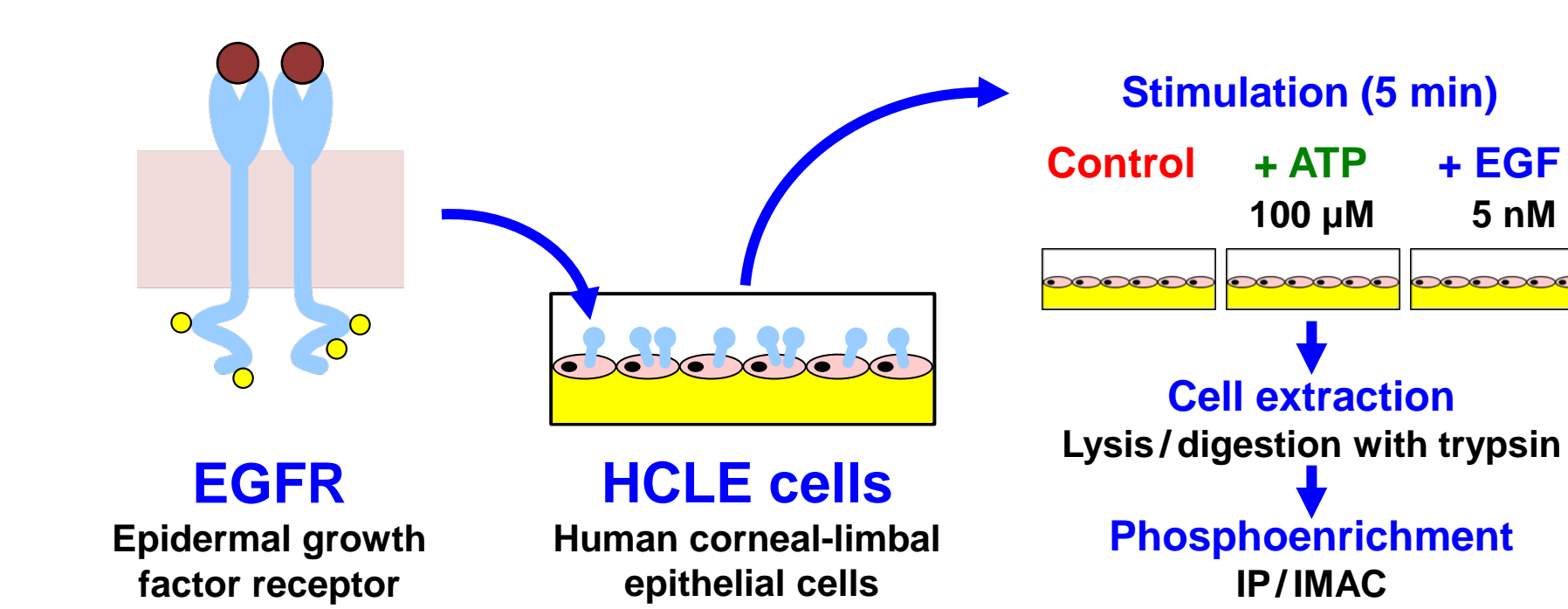
- All PTM standards ranked at top of list.
- Correct trend of decreasing counts observed for all PTMs.
- Two potential sites (S4, S10) indicated for glycosylation on Std #500; ranking determined by average PTM score (30.7).

Label-Free Comparison (Progenesis LC-MS)



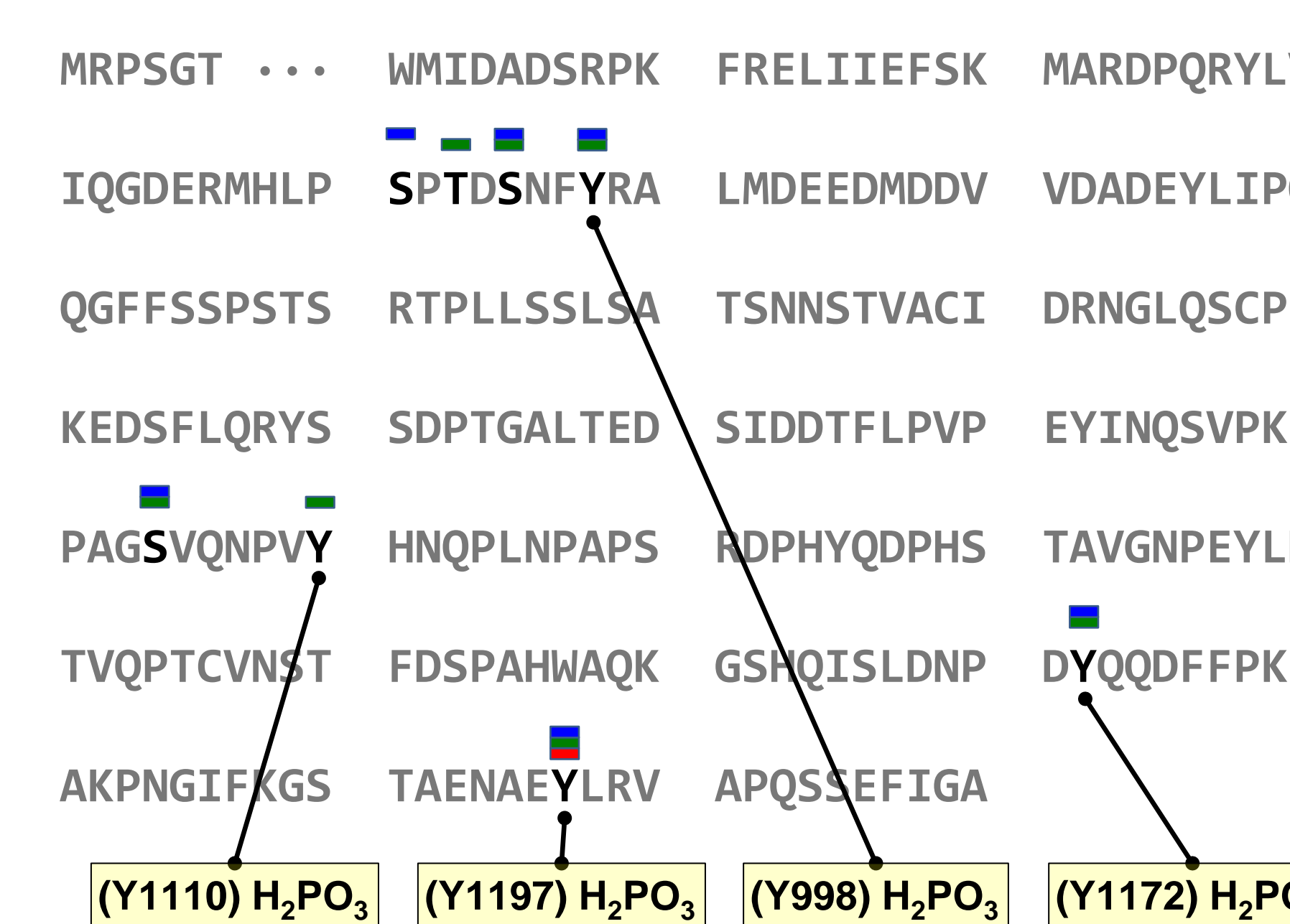
Results

Example 2: EGFR Phosphorylation



PTM Map (phosphorylated sites on EGFR)

Legend: Control ATP EGF

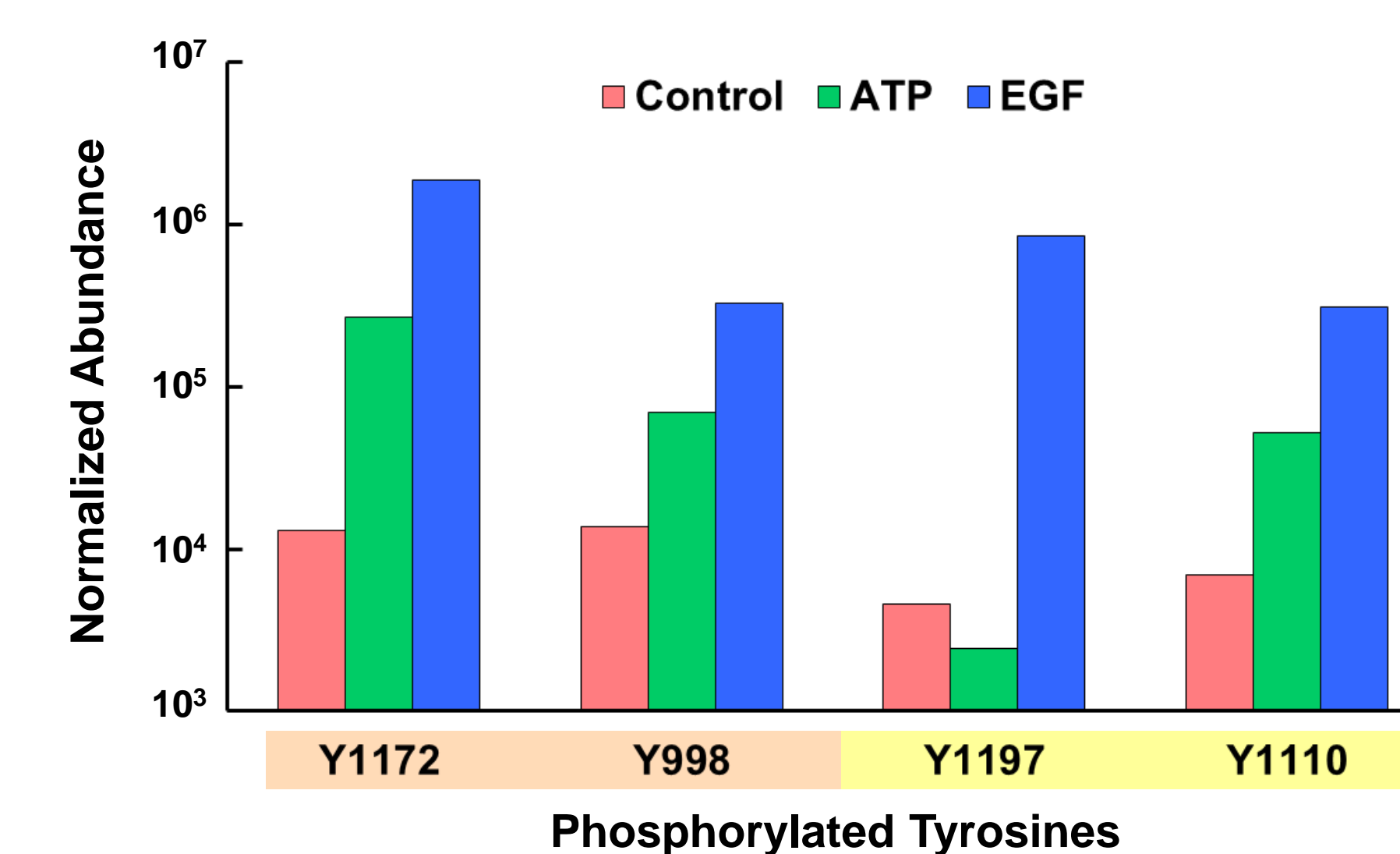


PTM Scoring (all PTMs)

PTM Score	Mod Mass	Amino Acid	Counts				Mod Forms	Avg	SD	Other Forms	Scoring Factors			
			Ctrl	ATP	EGF	Q					G	W	U	
67.4	79.966331	Y1172	0	1	7	8	2.667	3.788	0	0.878	0.767	1.000	0	
28.3	79.966331	Y998	0	7	9	16	5.333	4.726	4	0.369	0.958	0.800	0	
17.5	79.966331	Y1197	1	3	2	6	2.000	1.000	0	0.862	0.203	1.000	0	
3.8	79.966331	Y1110	0	1	0	1	0.333	0.577	2	0.983	0.117	0.333	0	

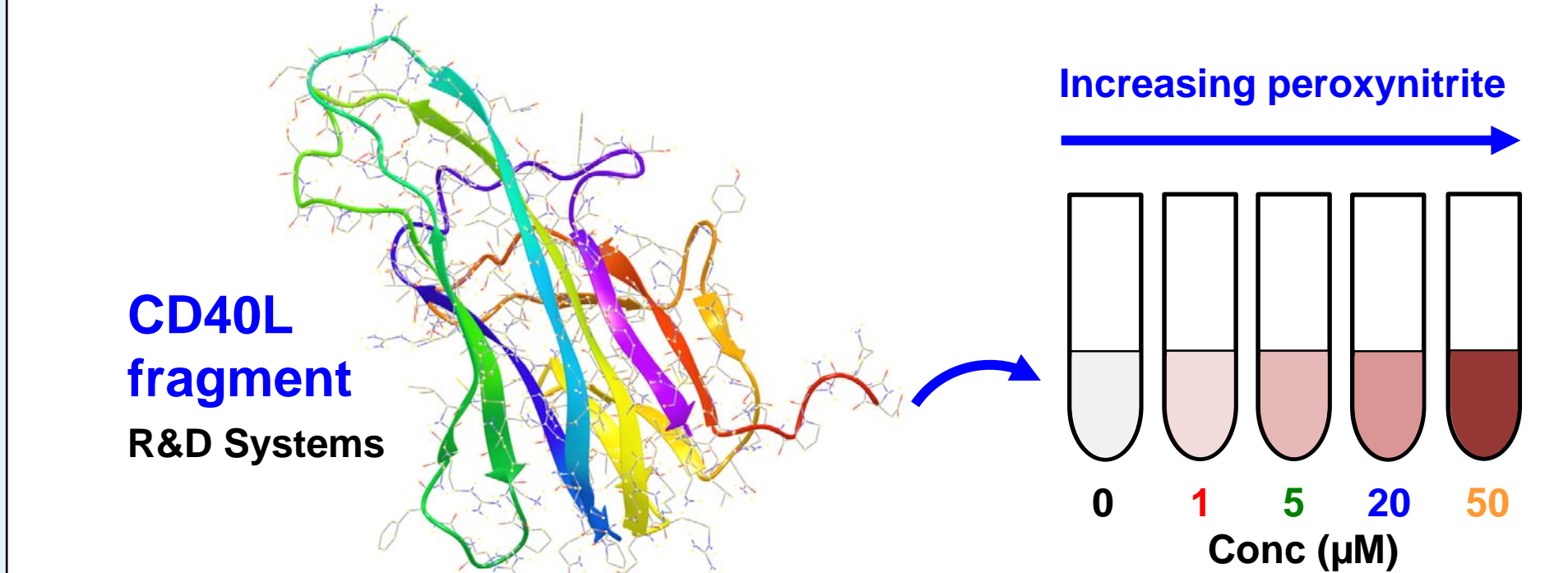
- Sites involved in cellular signaling processes (Y998, Y1110, Y1172, Y1197) ranked highest.
- Distinct patterns observed at two sites (Y998, Y1172) for ATP and EGF stimulation.

Label-Free Comparison (Progenesis LC-MS)



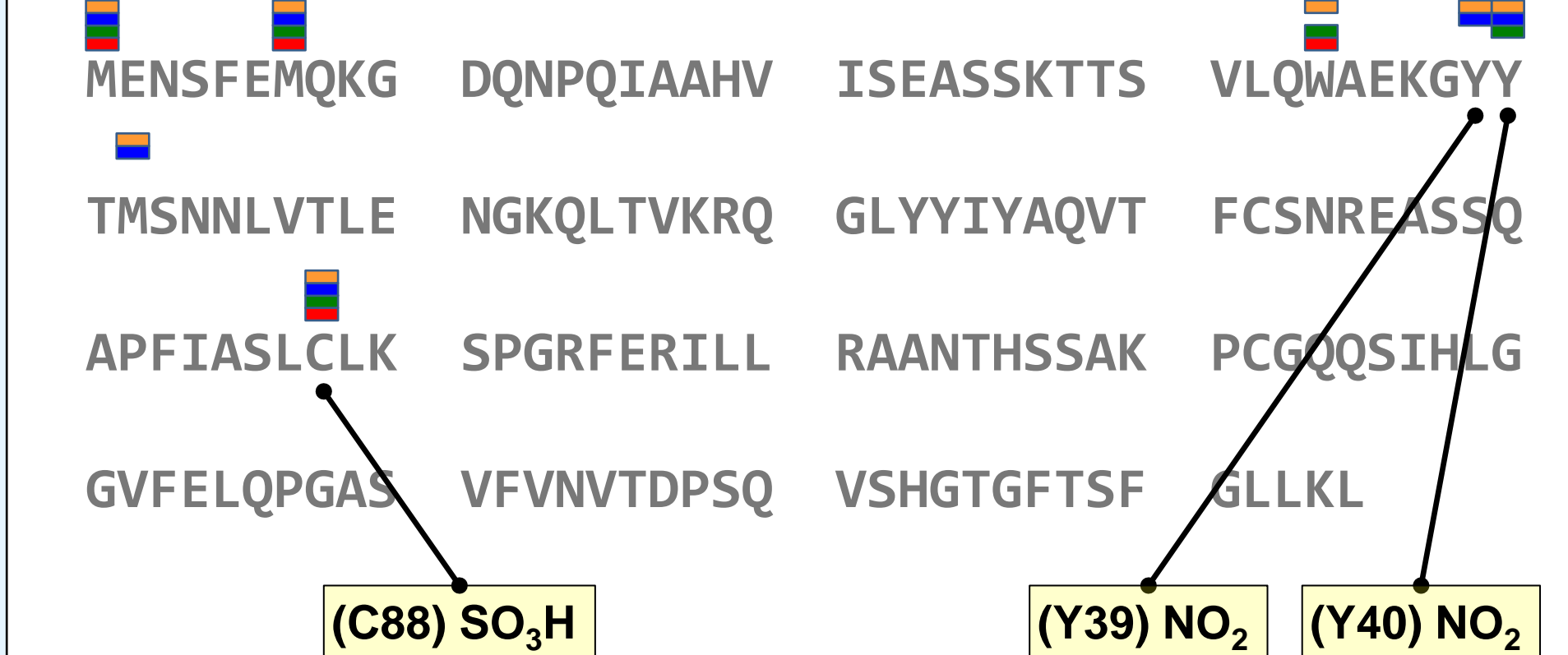
Results

Example 3: CD40L In Vitro Oxidation



PTM Map (all modified sites on CD40L)

Legend: 1 μM 5 μM 20 μM 50 μM



PTM Scoring (all PTMs)

PTM Score	Mod Mass	Amino Acid	Counts				Mod Forms	Avg	SD	Other Forms	Scoring Factors			
			1 μM	5 μM	20 μM	50 μM					Q	G	W	U
19.2	15.994915	M7	6	3	8	4	21	5.250	2.217	18	0.997	0.772	0.538	0.464
10.8	44.985078	Y40	0	1	2	6	9	2.250	2.630	7	0.263	0.916	0.563	0.798
10.4	44.985078	Y39	0	0	3	5	5	2.000	2.449	8	0.307	0.853	0.500	0.798
8.0	15.994915	M42	0	0	3	6	9	2.250	2.872	7	0.307	1.000	0.563	0.464
6.9	47.984744	C88	2	2	2	0	6	1.500	1.000	22	1.000	0.348	0.214	0.929

- Sites involved in CD40L structural functionality (Y39, Y40, C88) ranked high.
- Redox-sensitive residues observed for two nitrotyrosines (Y39, Y40) associated with receptor binding.

Conclusions

- STRAP PTM is a powerful counting approach for ranking differential PTMs in large MS-based proteomics data sets.
- STRAP PTM applies a novel multi-component score to PTMs with results clearly visualized in interactive tables and maps.
- STRAP PTM software is easy to implement on a PC and provides fast turnaround for large data sets.
- STRAP PTM produces semi-quantitative results with trends substantiated by label-free analysis.

Acknowledgments

- NIH-NHLBI contract HHSN268201000031C
- NIH grants P41 RR010888/GM104603 and S10 RR020946