Counting and Label-Free Approaches to Identify PTM Biomarkers of Cardiovascular Disease in a Mouse Model

Jean L. Spencer, Stephen A. Whelan, Christian F. Heckendorf, Deborah A. Siwik, Wilson Colucci, Markus M. Bachschmid, Richard A. Cohen, Catherine E. Costello, and Mark E. McComb

Boston University School of Medicine, Boston, MA 02118

Overview

- **Purpose:** Counting software and label-free proteomics were used to characterize differential PTMs and potential biomarkers of cardiovascular disease in a mouse model.
- **Methods:** Blood samples from an American diet mouse model were analyzed using in-house software (STRAP PTM) based on spectral counting and a novel scoring algorithm.
- **Results:** Trends in differential PTMs were readily detected and easily sorted for biologically relevant modifications induced by oxidative stress.

Introduction

There are many factors that contribute to cardiovascular disease (CVD), but unfavorable metabolic conditions associated with obesity, diabetes, and hyperlipidemia are major causes. Early detection and monitoring of these adverse effects on the heart and vasculature are difficult to achieve. Since plasma proteins are indicators of inflammation and oxidants, nonspecific changes in these components may reflect systemic metabolic disease. The appearance and change of post-translational modifications (PTMs) on circulating proteins may be critical in determining biomarkers and therapeutic targets for this disease. Here we investigate the power of counting software (STRAP PTM) and label-free proteomics to characterize differential PTMs and potential biomarkers of CVD using an American diet (high-fat, high-sucrose intake) mouse model.

Methods

Mouse Model
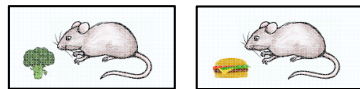
Control diet (Control)   High-fat, high-sucrose diet (HFHS)

Sample Processing

- Blood collection at 4 months from each group (n = 4)
- Protein digestion with trypsin to generate peptides
- Plasma proteomics

LC-MS/MS Analysis

- nanoACQUITY HPLC system (Waters)
- TriVersa NanoMate ESI source (Advion)
- LTQ-Orbitrap mass spectrometer (Thermo Scientific)
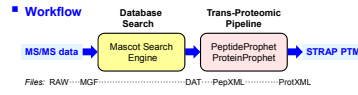- Data-dependent MS/MS acquisition

Label-Free Quantitative Analysis

- Progenesis LC-MS (v4.1; Nonlinear Dynamics)
- Mascot search engine (v2.3; Matrix Science)

Methods

STRAP PTM Analysis

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

MS/MS data → [Database Search: Mascot Search Engine] → [Trans-Proteomic Pipeline: PeptideProphet ProteinProphet] → STRAP PTM

Files: RAW ·· MGF ·········· DAT · PepXML ········ ProtXML

- **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)

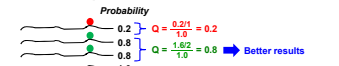$$S_{mip} = 100 \times Q_{mip} \times G_{mip} \times W_{mip} \times U_{mp}$$

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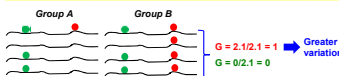
$$Q_{mip} = \frac{\langle P_{mip} \rangle}{\langle P^0_{ip} \rangle}$$

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

Probability

Q = 0.2/1 = 0.2
Q = 1.6/2 = 0.8 → Better results

Grouping (G): Variation of a specific PTM on a specific site of a specific protein across groups (max = 1)
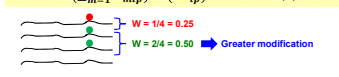
$$G_{mip} = \frac{\sigma_{mip}}{\max \sigma}$$

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

Group A   Group B

G = 2.1/2.1 = 1 → Greater variation
G = 0/2.1 = 0

Occupancy (W): Degree of modification of a specific site on a specific protein with a specific PTM (max = 1)
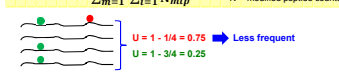
$$W_{mip} = \frac{N_{mip}}{\left(\sum_{m=1}^{M} N_{mip}\right) + \left(N^0_{ip}\right)}$$

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

W = 1/4 = 0.25
W = 2/4 = 0.50 → Greater modification

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

U = 1 - 1/4 = 0.75 → Less frequent
U = 1 - 3/4 = 0.25

Results

STRAP PTM Results

PTM Scores (6 modified protein hits)

Protein Summary

Protein Name	Group Overlap	Average PTM Score	Control % Cover	HFHS % Cover	
sp|Q91X72|HEMO_MOUSE	1	9.954	19.68	22.2	Hemopexin
sp|P07724|ALBU_MOUSE	1	7.028	50.34	50.34	Serum albumin
sp|Q00623|APOA1_MOUSE	1	1.21	46.25	41	Apolipoprote
sp|Q8K0E8|FIBB_MOUSE	1	0.265	44.06	48.66	Fibrinogen b
sp|Q61129|CFAI_MOUSE	1	0.187	14.19	13.68	Complement
sp|Q8VCM7|FIBG_MOUSE	1	0.173	48.42	45.5	Fibrinogen g
sp|Q8K182|CO8A_MOUSE	0		12.39	12.21	Complement
sp|P06909|CFAH_MOUSE	0		4.85	4.85	Complement
sp|P13020|GELS_MOUSE	0		21.56	33.33	Gelsolin OS-
sp|P01872|IGHM_MOUSE	0		13.66	16.74	Ig mu chain

PTM Search Space

PTM	AA	Δ Mass	Composition
Oxidation*	(M)	[+15.99]	O
Cyano*	(C)	[+25.00]	H(-1) C N
Dioxidation*	(C)	[+31.99]	O(2)
Methylthio	(C)	[+45.99]	H(2) C S
Trioxidation*	(C)	[+47.98]	O(3)
Glutathione	(C)	[+305.07]	H(15) C(10) N(3) O(6) S

Differential Observation of PTMs

Diff = HFHS counts – Control counts

*Oxidation (M): ALBU M264, APOA1 M218, APOA1 M216
*Cyano (C): ALBU C392, ALBU C34, ALBU C567, ALBU C487
*Dioxidation (C): ALBU C265
*Trioxidation (C): ALBU C279, ALBU C278
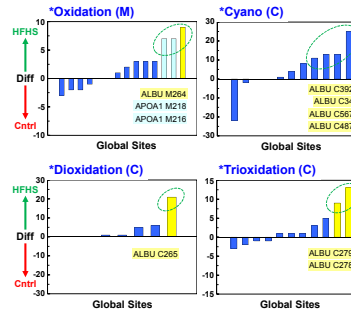
- Proteins with significant differential PTMs ranked high among other proteins based on average PTM scores.
- Distributions of differential counts highlighted trends in global protein sites with PTMs of non-random positive changes.

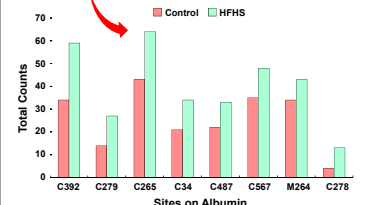## Results

STRAP PTM Results

PTM Map: Albumin

Legend: Control HFHS

```
EAHKSEIAHR YNDLGEQHFK GLVLIAFSQY LQKCSYDEHA KLVQEVTDFA
KTCVADESAA NCDKSLHTLF GDKLCAIPNL RENYGELADC CTKQEPERNE
CFLQHKDDNP SLPPFERPEA EAMCTSFKEN PTTFMGHYLH EVARRHPYFY
APELLYYAEQ YNEILTQCCA EADKESCLTP KLDGVKEKAL VSSVRQRMKC
SSMQKFGERA FKAWAVARLS QTFPNADFAE ITKLATDLTK VNKECCHGDL
LECADDRAEL AKYMCENQAT ISSKLQTCCD KPLLKKAHCL SEVEHDTMPA
DLPAIAADFV EDQEVCKNYA EAKDVFLGTF LYEYSRRHPD YSVSLLLRLA
KKYEATLEKC CAEANPPACY GTVLAEFQPL VEEPKNLVKT NCDLYEKLGE
YGFQNAILVR YTQKAPQVST PTLVEAARNL GRVGTKCCTL PEDQRLPCVE
DYLSAILNRV CLLHEKTPVS EHVTKCCSGS LVERRPCFSA LTVDETYVPK
EFKAETFTFH SDICTLPEKE KQIKKQTALA ELVKHKPKAT AEQLKTVMDD
FAQFLDTCCK AADKDTCFST EGPNLVTRCK DALA
```

17 key sites with PTMs: 15/36 Cys, 2/7 Met

PTM Scoring: Albumin

PTM Score	Mod Mass	Site	Total Counts Control	Total Counts HFHS	Diff	Modified Forms	Other Forms	Q	G	W	U
46.5	25.00	C392	34	59	25	93	27	1.000	1.000	0.775	0.600
30.6	47.98	C279	14	27	13	41	12	1.000	0.520	0.774	0.761
30.1	31.99	C265	43	64	21	107	66	0.752	0.840	0.618	0.777
21.8	25.00	C34	22	33	11	55	1	0.710	0.520	0.982	0.600
20.7	25.00	C487	22	33	11	55	6	1.000	0.440	0.786	0.600
12.2	25.00	C567	35	48	13	83	96	0.843	0.520	0.464	0.600
11.7	15.99	M264	34	43	9	77	96	0.820	0.360	0.445	0.890
8.8	47.98	C278	4	13	9	17	36	1.000	0.360	0.321	0.761
6.0	47.98	C51	6	12	6	18	45	0.535	0.200	0.739	0.761
3.5	25.00	C265	12	20	8	32	141	1.000	0.320	0.185	0.600

Control HFHS

Sites on Albumin: C392, C279, C265, C34, C487, C567, M264, C278
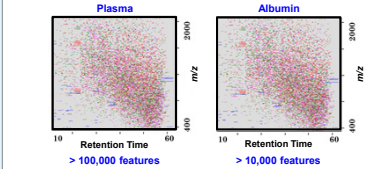
- PTM sites sensitive to HFHS diet were easily identified on albumin by high PTM scores.
- Top-8 redox-specific sites on albumin were selected and mapped for further investigation (e.g., C392).
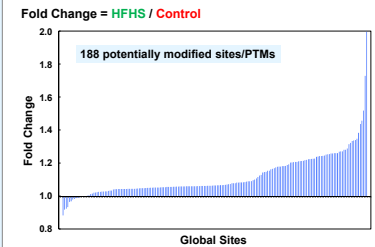
Results

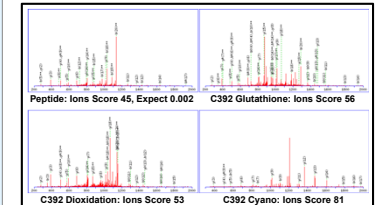### Label-Free Results

Plasma Proteins vs. Albumin PTMs

Plasma   Albumin

> 100,000 features   > 10,000 features

Normalized Abundances: Albumin - PTMs

Fold Change = HFHS / Control

188 potentially modified sites/PTMs

Confirmation of Results

MS/MS Fragmentation of Albumin Peptide +3 PTMs

Peptide: Ions Score 45, Expect 0.002   C392 Glutathione: Ions Score 56
C392 Dioxidation: Ions Score 53   C392 Cyano: Ions Score 81

Conclusions

- Select HFHS-induced PTMs were successfully mapped onto albumin by counting and label-free approaches.
- STRAP PTM counting provided an easy and quick overview of modified sites on protein sequences.
- Subsequent quantitative analysis with label-free methods was facilitated by preliminary analysis with STRAP PTM.

Acknowledgments