Deeper proteome coverage of musculoskeletal samples

Emma Doud, Ph.D.[1,2], Guihong Qi[1,2], Aruna Wijeratne, Ph.D.[1,2], Xiaoling Zhong, M.D., Ph.D.[3], Joseph Rupert[3], Matthew Willetts, Ph.D.[5], Shourjo Ghose, Ph.D.[5], Teresa Zimmers, Ph.D.[1,3,4], Amber Mosley, Ph.D.[1,2] •

[1]Biochemistry and Molecular Biology, [2]Proteomics Core Facility, [3]General Surgery, Indiana University School of Medicine, Indianapolis, IN [4] Richard Roudebrush VA Medical Center, Indianapolis, IN [5] Bruker Scientific, Billerica, MA

ABSTRACT

Bottom-up proteomic analysis of skeletal muscle samples is hindered by the presence of abundant and large proteins such as titin and myosin, which create hundreds to thousands of unique tryptic peptides. As a result, despite offline fractionation, significant MS time in data dependent acquisition (DDA)-based analysis is spent cataloging a small number highly abundant peptides rather than increasing the depth of proteome coverage. In this work, we have compared our standard global proteomics workflow to the PASEF acquisition method on the timsTOF Pro (Bruker). PASEF acquisition allows a significant increase in sequencing speed by pre-fractionating peptides based on their collision cross section (CCS). These studies found that peptide-spectral matches were doubled, and protein identifications increased by one third using PASEF acquisition. Although these results were striking, we will continue to explore additional prefractionation methods to improve the overall depth of coverage of the musculoskeletal proteome for discovery proteomics of mouse and human skeletal muscle disease states.

BACKGROUND

Myosin, actin, and titin comprise up to 25 % of the total mass of muscle tissue. Together, these proteins interact to create muscle contraction. There are over 40 myosin genes in the human genome which encode 12 distinct protein classes. At over 35,000 amino acids or 3 mDa, titin is the largest known protein; trypsin digestion results in over 3000 titin peptides.[1]

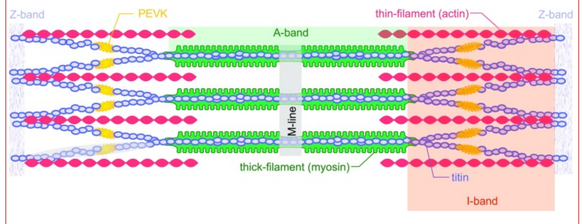

Figure 1. Two-dimensional illustration of a sarcomere; showing myosin, actin, and titin filaments.[1]

PROTEOMICS CORE WORKFLOW

After lysis in 8 M Urea, proteins are reduced with TCEP, alkylated with chloroacetamide (CAM), digested overnight with trypsin/LysC. For multiplexing, samples are TMT-labeled and mixed. Peptides are then fractionated with high pH reverse-phase offline chromatography prior to low pH reverse-phase LC-MS/MS.

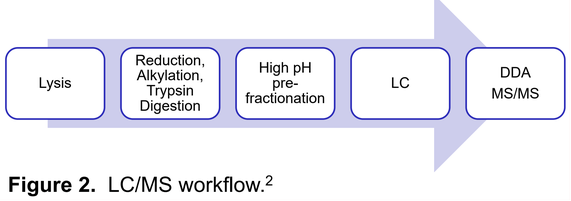

Figure 2. LC/MS workflow.[2]

MUSCLE PROTEOME COVERAGE

Application of the typical proteomics workflow pipeline resulted in 2190 protein group IDs and 19986 peptide IDs. However, over 50 % of the data comes from just the top 15 most abundant proteins, mostly titin, actin, and myosin isoforms (Table 1).

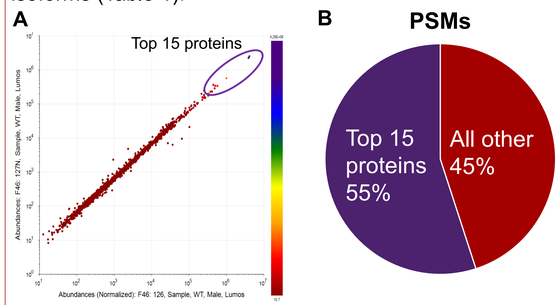

Figure 3. Coverage of musculoskeletal samples 2190 protein group identifications and 19986 peptides. A. Normalized abundances of proteins in two TMT labeled LC/MS/MS replicates of mouse muscle tissue. **B.** Peptide spectral matches from the top 15 proteins versus the entire experiment.

Table 1. Top 15 most abundant proteins in muscle tissue.

Accession	Description	# PSMs	# Peptides
A2ASS6	Titin	71355	2192
Q5SX39	Myosin-4	86501	261
Q5SX40	Myosin-1	55832	242
G3UW82	MCG140437, isoform CRA_d	49184	231
P13542	Myosin-8	35516	181
Q91Z83	Myosin-7	17500	190
E9Q1W3	Protein Neb	16308	476
P13541	Myosin-3	15776	101
Q02566	Myosin-6	15287	132
P07310	Creatine kinase M-type	15234	36
Q8R429	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	13923	67
P68134	Actin, alpha skeletal muscle	13055	29
B1AR69	Protein Myh13	11415	90
P68033	Actin, alpha cardiac muscle 1	9693	29
F8WID5	Tropomyosin alpha-1 chain	8573	60

HIGH THROUGHPUT FRACTIONS

Instead of 12 high pH basic fractions, a representative muscle tissue digest was fractionated into 48 samples which were run in one day via Evosep-LC/MS/MS technology. The microflow chromatography resulted in fewer IDs (2223 vs 1511 Protein groups).

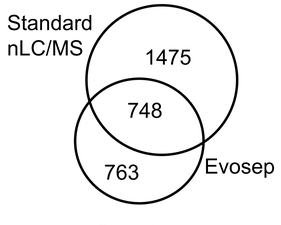

Figure 4. 48 fractions run on Evosep versus 12 on standard nanoflow LC/MS.

timsTOF PRO

Parallel accumulation:

Usage of nearly all ions (100% duty cycle)

Time focusing: Ion species are accumulated for 50-100 ms in the TIMS tunnel, however elution of each ion species from the TIMS tunnel happens within 2-5 ms

Space focusing: Ion packages eluted at a specific time point dependent on their mobility, resulting in cleaner MS/MS spectra

Parallel accumulation serial fragmentation (PASEF) cycle time: By synchronizing the quadrupole with the TIMS elution time, an average of 12 precursors can be fragmented within a 100 ms timescale. Intelligent software targets low-level precursors multiple times for PASEF MS/MS fragmentation.

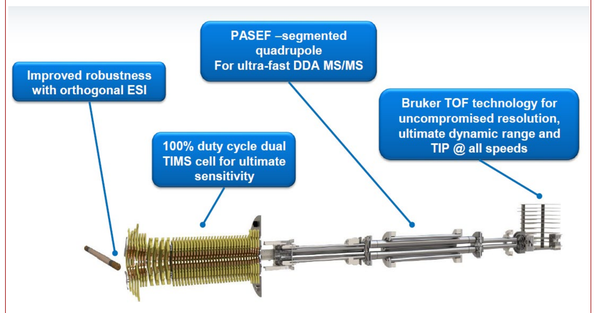

Figure 5. Schematic of timsTOF pro (Bruker Scientific.)[3]

timsTOF vs STANDARD LC/MS/MS

A representative muscle sample was prepared and fractionated into 9 high pH basic fractions. Each fraction was analyzed using:

(1) 180 min LC gradient with typical proteomics core instrumentation

(2) 90 min LC gradient onto a timsTOF Pro (Bruker Scientific)

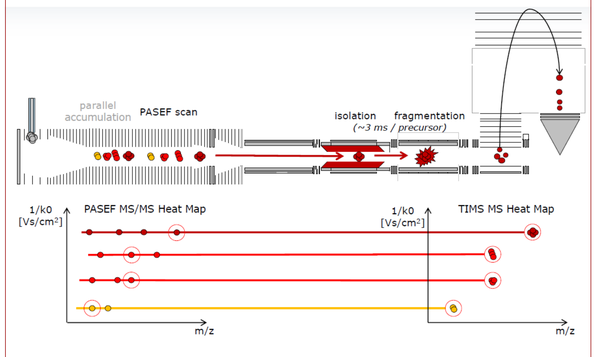

Figure 6. PASEF allows for speed as a result of time focusing and sensitivity as a result of space focusing.

SUMMARY

In half the total instrument time, the timsTOF Pro increased the number of protein and unique peptide identifications by one third, as well as more than doubling the number of peptide spectral matches. Both methods spent approximately a third of total acquisition time on the top 15 most abundant muscle proteins, including myosin and titin.

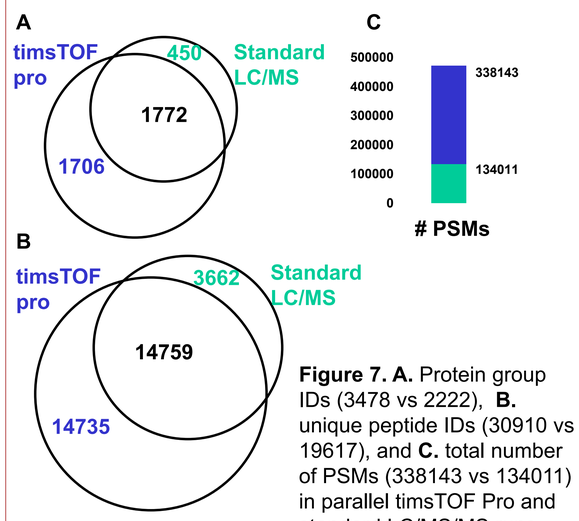

Figure 7. A. Protein group IDs (3478 vs 2222), **B.** unique peptide IDs (30910 vs 19617), and **C.** total number of PSMs (338143 vs 134011) in parallel timsTOF Pro and standard LC/MS/MS runs.

CONCLUSIONS

Additional pre-fractionation and orthogonal techniques for introduction of ions into the mass spectrometer are classic means of improving the number of protein and peptide IDs in a global proteome experiment. The sensitivity of trapped ion mobility combined with the speed of PASEF monitoring increased protein and peptide IDs by one third and doubled the total number of PSMs in our mouse muscle tissue system. We anticipate this technology will allow proteomics core workflows to expand proteome coverage in tissues where several highly abundant proteins preclude deeper proteome coverage.

REFERENCES

1 Herzog, Walter. “The multiple roles of titin in muscle contraction and force production.” *Biophysical reviews* vol. 10,4 (2018): 1187-1199.

2 Mosley, Amber et al. “A label free quantitative proteomic analysis of the Saccharomyces cerevisiae nucleus.” *Journal of proteomics* vol 72, (2009):110-120.

3 Meier, Florian et al. “Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer.” *Molecular & cellular proteomics* vol. 17,12 (2018): 2534-2545.