Potential improvements to Skyline to enable DIA analysis of metabolic labeling experiments

Nicholas Shulman[1], Natan Basisty[2], Birgit Schilling[2], Michael J MacCoss[1], Brendan MacLean[1]
[1]University of Washington Genome Sciences, Seattle, WA, United States; [2]Buck Institute for Research on Aging, Novato, CA, United States

Overview

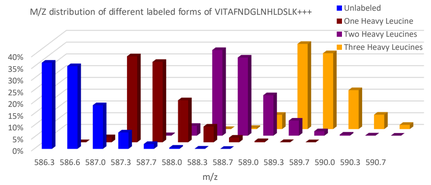
- Skyline is a freely-available, open-source Windows client application for analyzing and viewing many types of mass spectrometry data including DIA/SWATH
- Protein turnover in live organisms can be measured by altering the diet to include a heavy isotope label, and using a mass spectrometer to measure the amount of label incorporated in peptides
- When feeding a heavy leucine labeled to a mouse, there is a lag time of days or months between when the diet is changed, and the amino acid precursor pool in the tissue of interest changes to reflect the new diet
- Peptides containing multiple leucines are of particular interest in metabolic labeling experiments like this because the distribution of unlabeled, partially labeled, and fully labeled forms can be used to infer the precursor pool.
- One of the challenges of using DIA to analyze protein turnover data like this is that peptides with multiple potentially labeled amino acids will have mass distributions that span multiple isolation window. Any particular isolation window will be biased towards either heavy or light forms of the peptide, so care must be taken to remove that bias.
- This poster walks through the analysis of a single MS1 spectrum and two MS2 spectra for a single peptide that had three potentially labeled leucine residues.

Methods

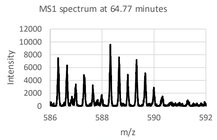
Mice were fed a diet containing $^2H_3$ labeled leucine and were sacrificed after being on that diet for seven days.
Kidney samples were analyzed on a TripleTOF 6600 with a DIA method with 64 isolation windows of varying width covering the range from 400-1200.
The data in this poster shows comes from a sample of the kidney cortex of one mouse.

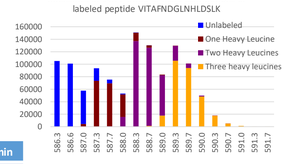
Deconvoluting MS1 spectra

A peptide such as VITAFNDGLNHLDSLK may contain 0, 1, 2, or 3 heavy leucines.
Each of these forms of the peptide has a m/z distribution which can be predicted, and which is independent of which particular leucine residues are heavy.



Using matrix multiplication[1] it is possible to find the scaling factors to apply to each of these distributions that will result in their sum being the closest match of the observed intensities



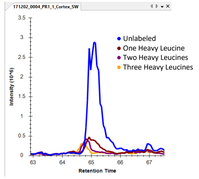


Labeled Form	Scaling Factor at 64.77 min
Unlabeled	$2.87x10^5$
One Heavy Leucine	$2.01x10^5$
Two Heavy Leucines	$3.77x10^5$
Three Heavy Leucines	$2.91x10^5$

Scaling factors which were determined to provide the best fit to the observed intensities in the MS1 scan at 64.77 minutes

Chromatographic trace from deconvoluted MS1 spectra

When the scaling factors for each retention time are plotted, a chromatogram is obtained for each labeled form. This chromatogram shows the characteristic shift where peptides with more deuterium elute earlier.



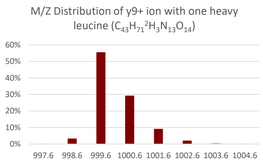
Calculating the expected isotope distribution of the fragment ion depending on the precursor isolation window

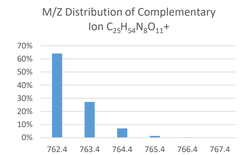
This precursor was isolated in two different MS2 scans:
Scan Type #28: 581.3 – 588.8 m/z
Scan Type #29: 587.8 – 595.8 m/z

Both of these scans only isolate part of the precursor's isotope envelope, so care must be taken to figure out the expected isotope distribution of the fragment ion, given the isolation window of the precursor.

We do this by considering the complementary ion– that is, the ion which is obtained by subtracting the chemical formula of the y9+ ion from the chemical formula of the of the precursor.

Precursor chemical formula: $C_{78}H_{128}N_{21}O_{25}$+++
Y9+ fragment ion formula: $C_{43}H_{74}N_{13}O_{14}$+
Complementary ion formula: $C_{35}H_{54}N_8O_{11}$++



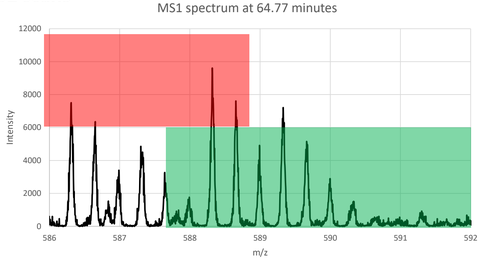


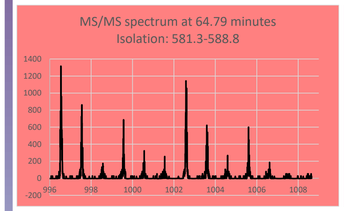
In order for the charge 3 parent ion to fall in the isolation window 581.3-588.8, the sum of the m/z of the y9 ion plus the complementary ions needed to have been less 1766.4 (i.e. 588.8 x 3).

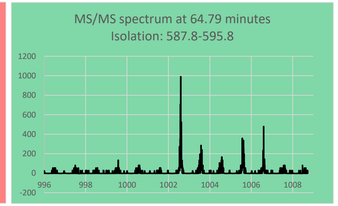
M/Z of fragment ion	Unfiltered abundance	Max m/z of complementary ion	% of complementary ion distribution	Filtered abundance
997.55	0.07%	768.85	100.00%	0.07%
998.56	3.33%	767.84	35.89%	1.19%
999.57	55.55%	766.83	8.63%	4.79%
1000.57	29.31%	765.83	1.35%	0.39%
1001.57	9.20%	764.83	0.00%	0.00%
1002.57	2.12%	763.83	0.00%	0.00%
1003.58	0.38%	762.82	0.00%	0.00%
1004.58	0.04%	761.82	0.00%	0.00%

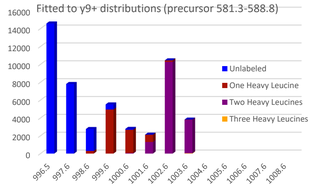
Deconvoluting MS2 scans

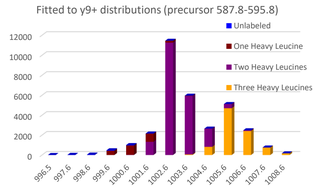
The MS1 spectrum at 64.77 minutes, was followed by two separate MS/MS scans which isolated the m/z range of our peptide of interest. One of these scans isolated the m/z range of 581.3-588.8 and the next scan covered the range 587.8-595.8.
Neither of these ranges cover the entire isotope envelope of precursor, so the expected m/z distribution of the fragment ion had to take that into account.

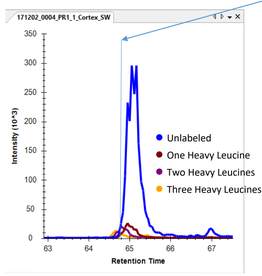












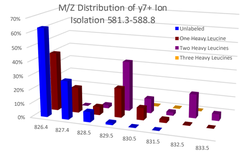
Labeled Form	Scaling factor for MS2 scan 581.3-588.8	Scaling factor for MS2 scan 587.8-595.8
Unlabeled	$25.5x10^3$	$25.8x10^3$
One Heavy Leucine	$8.91x10^3$	$8.93x10^3$
Two Heavy Leucines	$20.8x10^3$	$20.8x10^3$
Three heavy Leucines	$8.93x10^3$	$8.95x10^3$

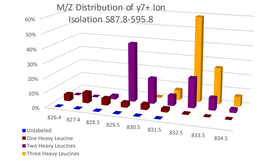
Chromatographic trace from deconvoluted MS2 spectra

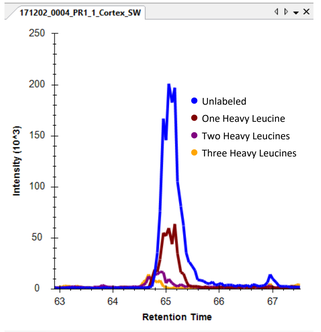
By plotting the fitted scaling factors of each scan, a chromatographic trace can be obtained which closely matches the trace from the MS1 scans. This shows that deconvoluting using the mass distribution of the fragment ion filtered by the precursor isolation window works correctly.

Inferring label amounts when the fragment ion does not have all of the leucines

The y7 ion of VITAFNDGL/NHLDSLK contains only two of the three leucines from the parent ion. If we assume that each leucine is equally likely to be labeled, then the partially labeled forms of the peptide have an m/z distribution which combines the fact that one of the labeled residues might not be in the fragment ion, and also that part of the parent ion was cut off by the isolation window. This might enable us to faithfully recover the exact distribution of unlabeled, partially, and fully labeled peptide, even though the fragment ion is missing one of the potentially labeled residues.







Chromatographic trace obtained from examining the y7 ion.
This technique appears to have overestimated the abundance of the "One Heavy Leucine" form of the peptide, compared to the results obtained from either the MS1 scans or the y9+ ion.
Further investigation is needed to determine whether doing the matrix algebra on more than one MS2 scan at a time is able to remove this artifact.

Conclusions

- Examining the entire isotope envelope of product ions in a DIA experiment can be a good way to determine label ratios in an unbiased way
- The results on this page are for a single peptide which had a very clean and strong signal. MS1 is already sufficient for analyzing high abundance peptides like that. More work is needed to determine whether these results can hold up for lower abundance peptides that require the selectivity of DIA
- If you have any suggestions about how to improve these algorithms to be more useful or to handle noisier data, feel free to contact me at nicksh@uw.edu or on the Skyline support board at https://skyline.ms.

References
[1] Brauman, John I. "Least squares analysis and simplification of multi-isotope mass spectra." *Analytical Chemistry* 38.4 (1966): 607-610.