
Welcome to:
Case Studies in Quantitative Proteomics

ASMS 2018
San Diego

Schedule

Time	Day 1	Day 2
9:00-9:30	Intro to Protein Quantification (Mike MacCoss)	Case Study 4: Is my mass spectrometer working? System Suitability. How we assess this? (Pino)
9:30-10:15		
10:15-10:45	Coffee Break	Coffee Break
10:45-11:15	Introduction to Skyline (Brendan MacLean)	Intro to DIA (MacCoss)
11:15-12:00	Case Study 1: Label free targeted assay development in a rodent model of heart failure (Brendan MacLean)	Case Study 5: When PRM is not enough -- hybrid assays (Jake Jaffe)
12:00-12:30	Lunch	Lunch
12:30-13:00		
13:00-13:30	What happens when an experiment is designed poorly: Case studies from failed studies (Olga Vitek)	Case Study 6: When do you need a PRM assay -- a case study in Chromatin proteomics (Jake Jaffe)
13:30-14:00		
14:15-14:45	Coffee Break	Coffee Break
14:45-15:15	Case Study 2: Group Comparisons in Skyline (Brendan MacLean)	I have peak areas, what do I do now? Intro to MSStats (Olga)
14:15-15:45		
15:45-16:15	Case Study 3: ABRF iPRG MS1 Quantification in Skyline (Brendan MacLean)	Hands on Analysis of Proteomics Datasets (Meena)
16:15-16:45		

What you should expect

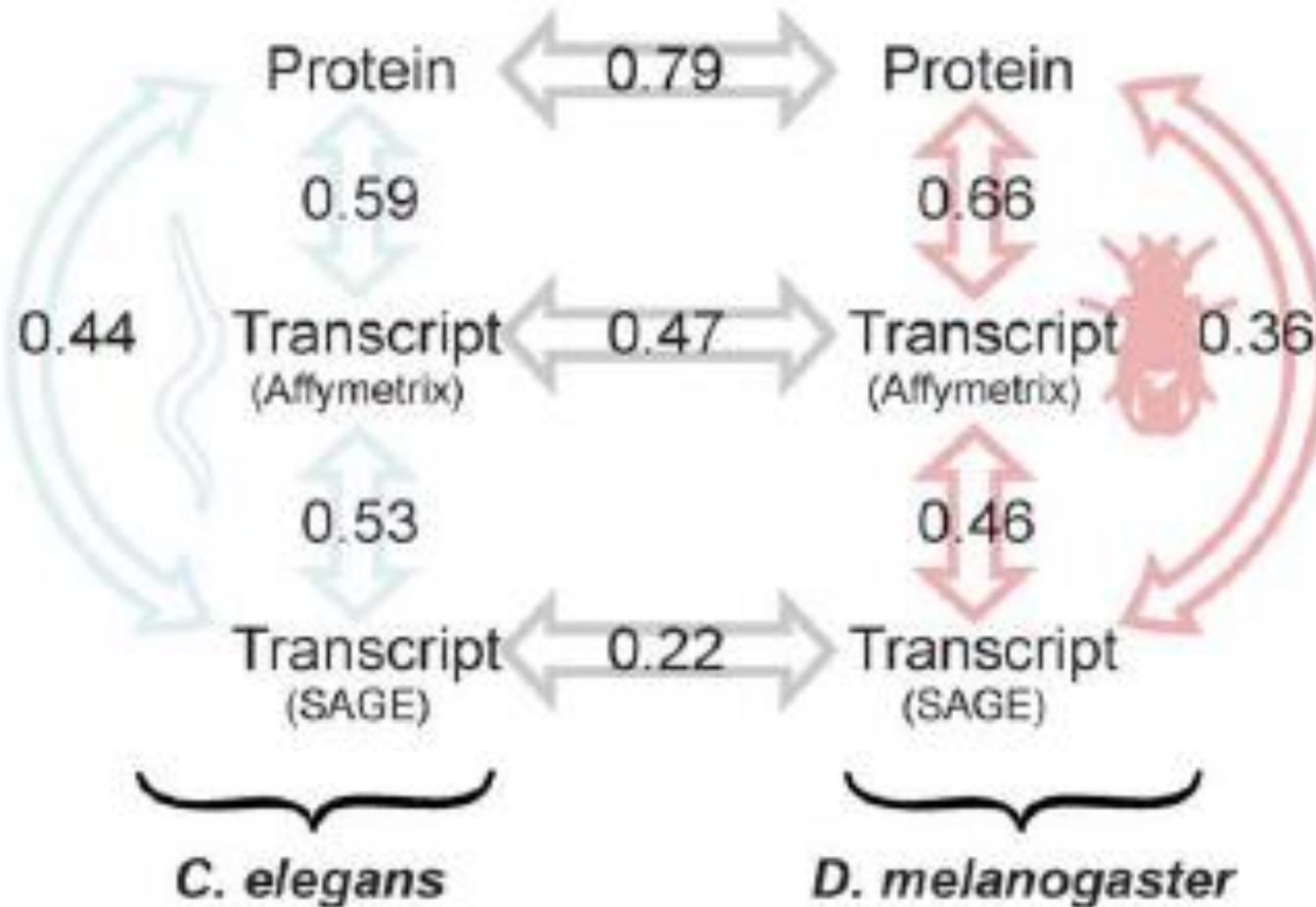
- You should expect to be challenged!
 - There should be material that you know but also there should be material that we present that you don't know.
- You should expect to learn the fundamentals of targeted proteomics.
- You should learn that quantitative proteomics is not possible without extensive quality control.
 - We will teach you ways that have worked for us.
- You should learn how we have handled real-life problems.
- You should have a good introduction to the use of Skyline for targeted proteomics.
- You may hear contradictory things from different instructors.
- The methods presented are not meant to be comprehensive. We will present strategies we used to solve problems in our labs.
- You should expect to have fun!!

What we expect.

- We expect you to challenge us!
 - You won't hurt our feelings – we have thick skin. We want feedback and also to hear that you want more.
- We expect different course participants to come from different backgrounds.
 - We expect to learn from you!
- We expect you to participate. The more you do, the more we will all get out of the course.
 - Don't hesitate to ask a question. Feel free to interrupt us.

Introduction to Quantitative Analysis of Proteins

Selective Pressure is Provided on the Protein and not the Transcript



Why Use Mass Spectrometry?

Problems with clinical immunoassays

Single-plex

Poor standardization

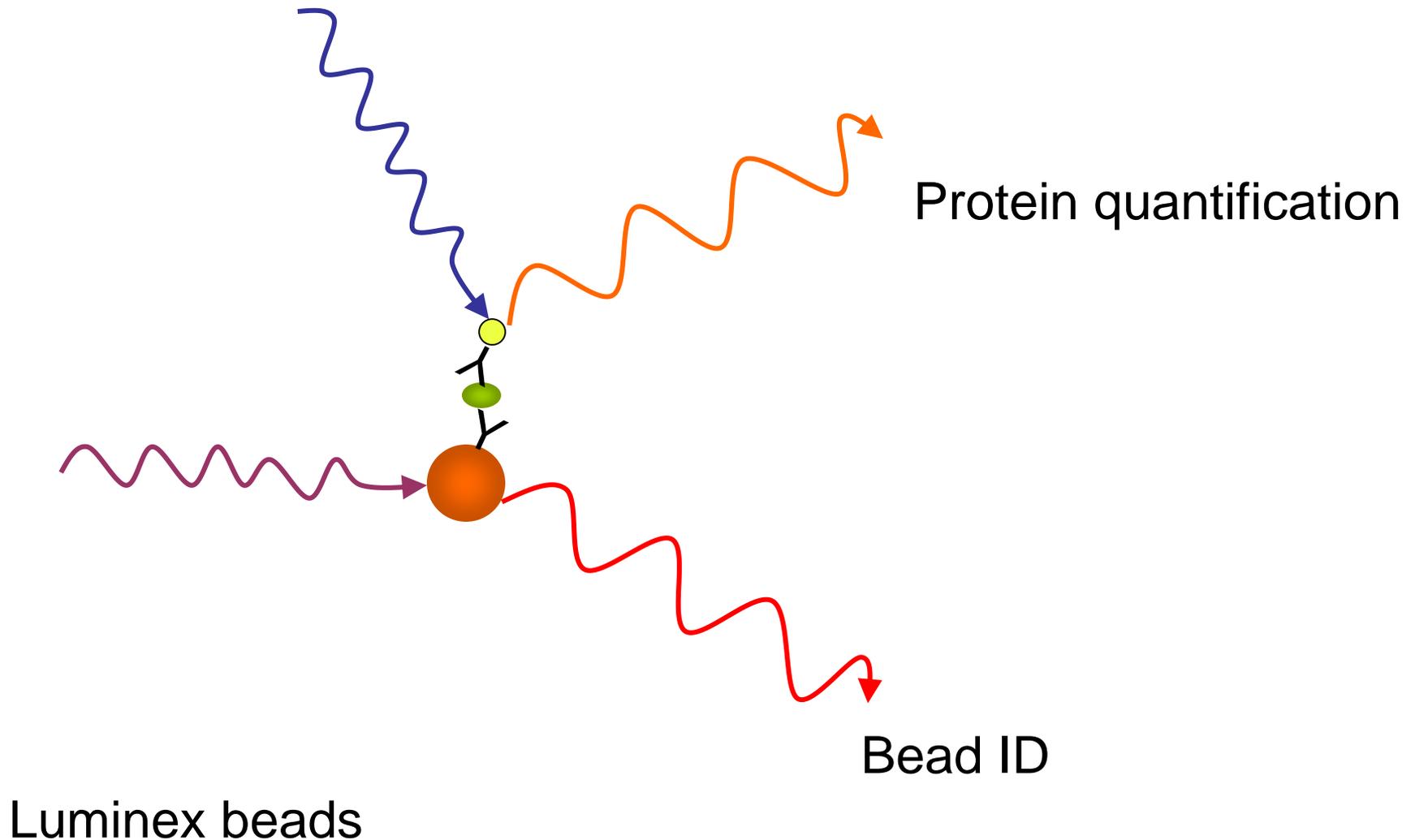
Hook effect

Anti-reagent antibodies

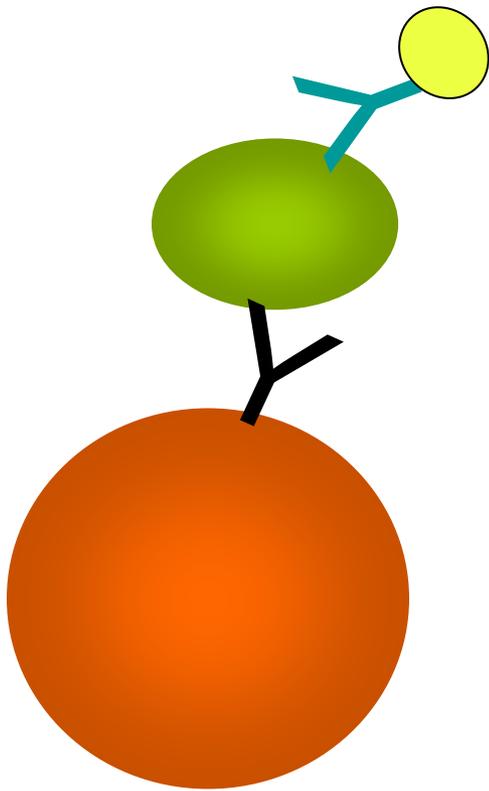
Autoantibodies

Hoofnagle and Wener, J Immunol Methods (2009)

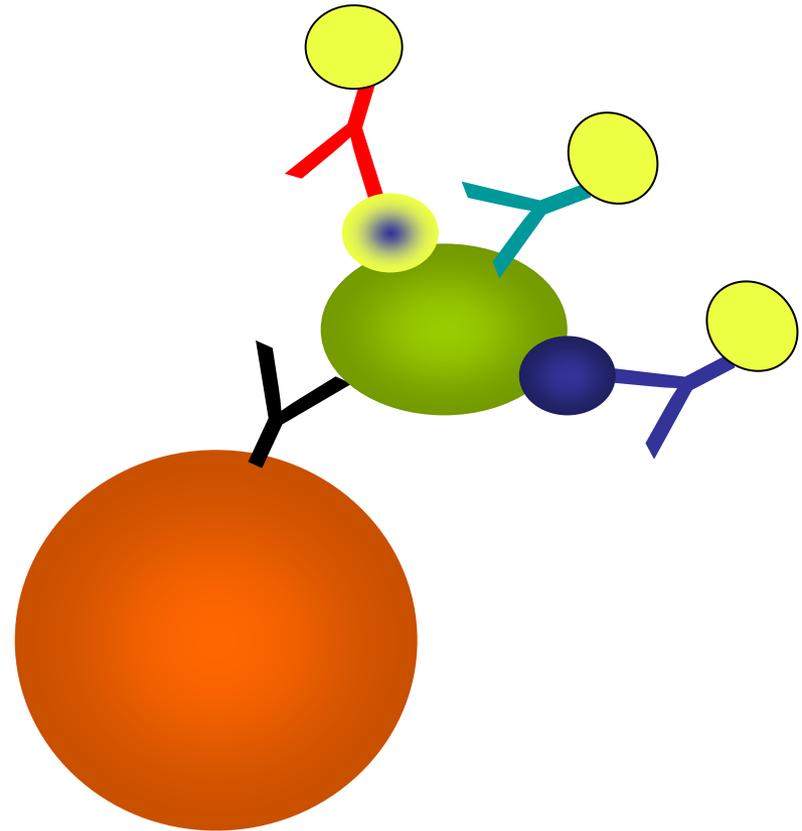
Can We Multiplex Immunoassays?



Can We Multiplex Immunoassays?

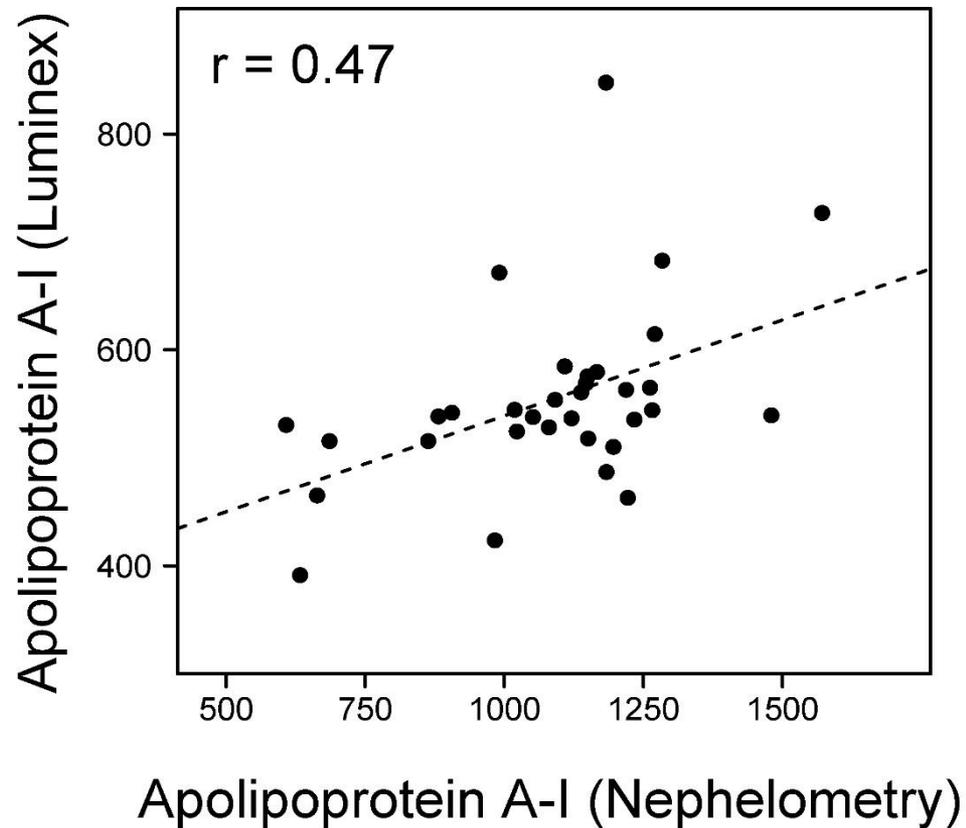


vs.



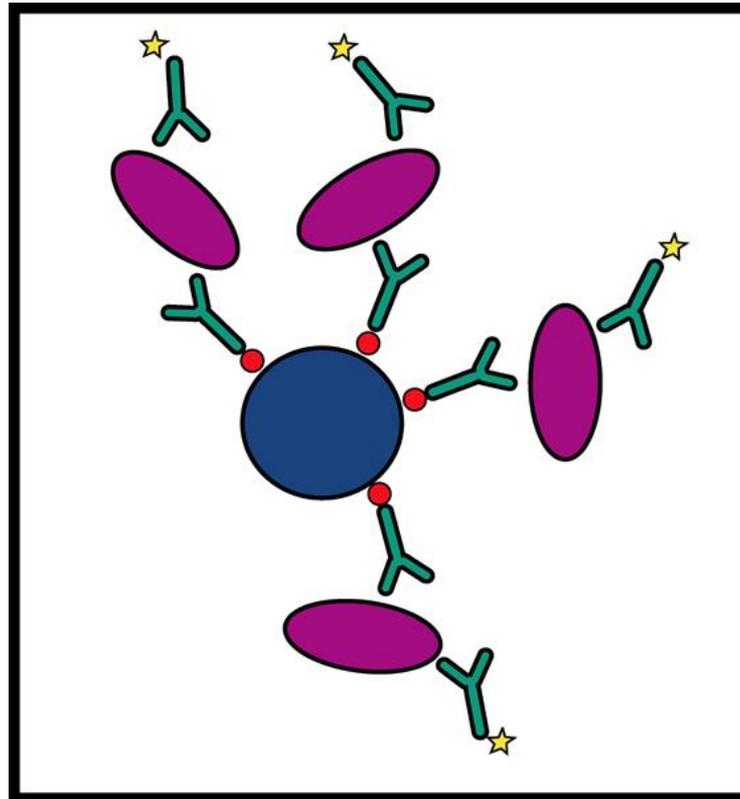
Can We Multiplex Immunoassays?

Not easily.



Traditional Automated Immunoassays

Capture antibody selects sandwiches
detected by enzyme-conjugated secondary
Generally one at a time



Poor Standardization

Immunoassays start with epitope selection

Whole antigen, domain, or peptide

Different assays use different epitopes

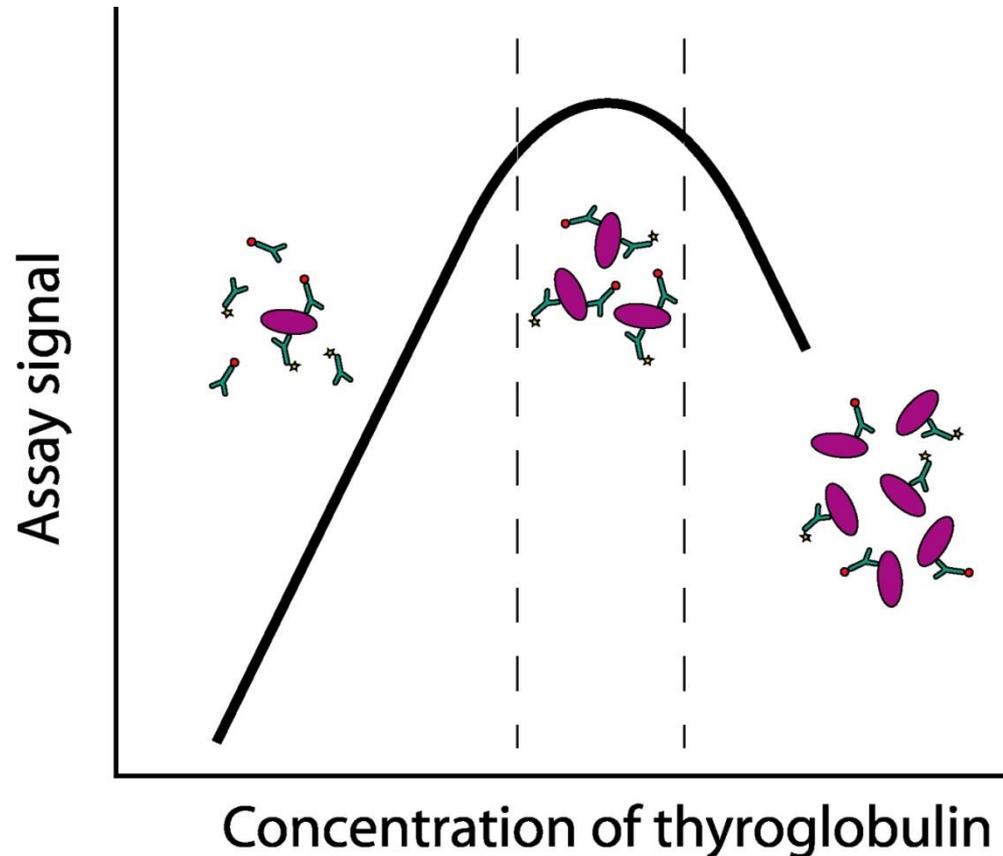
Tumor antigens vary in post-translational modifications

Result: Variable results between assay platforms

Even with an international standard (e.g. BCR-457)

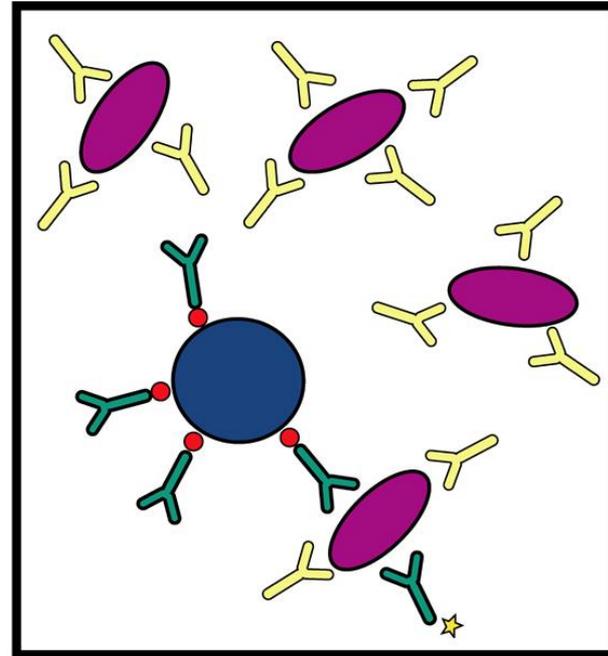
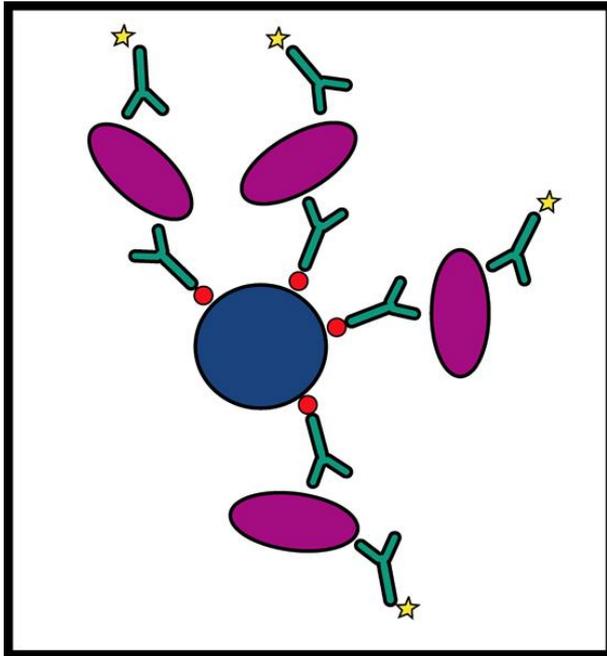
Hook Effect

Too much antigen can lead to negative result



Autoantibody Interference

Autoantibodies can sequester antigen away from detectable complexes



Falsely negative results with sandwich assays

Mass Spectrometry

A Potential Solution

Direct detection of analyte by mass spectrometry

Interlaboratory calibration

Improved specificity

No sandwiches formed

No hook effect

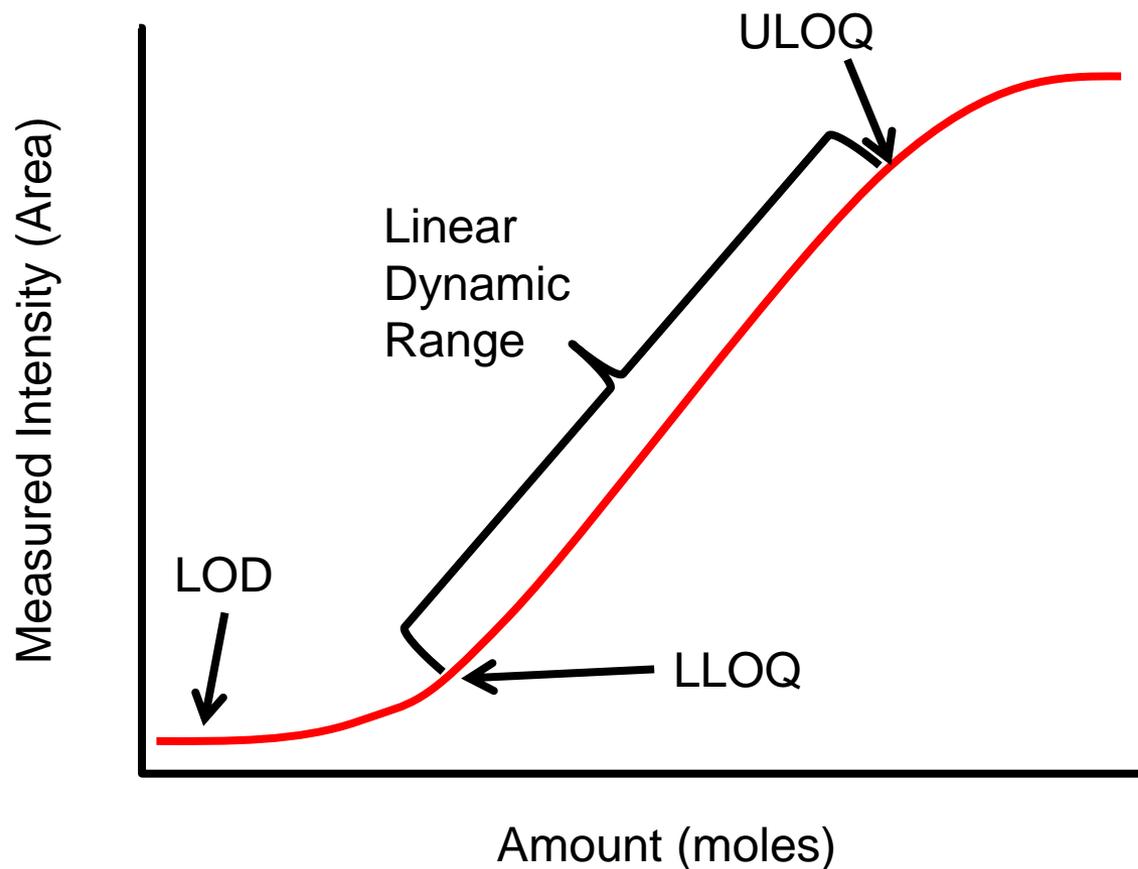
Saturation of signal – plateau of calibration curve

Digestion of all proteins

Destroy autoantibodies

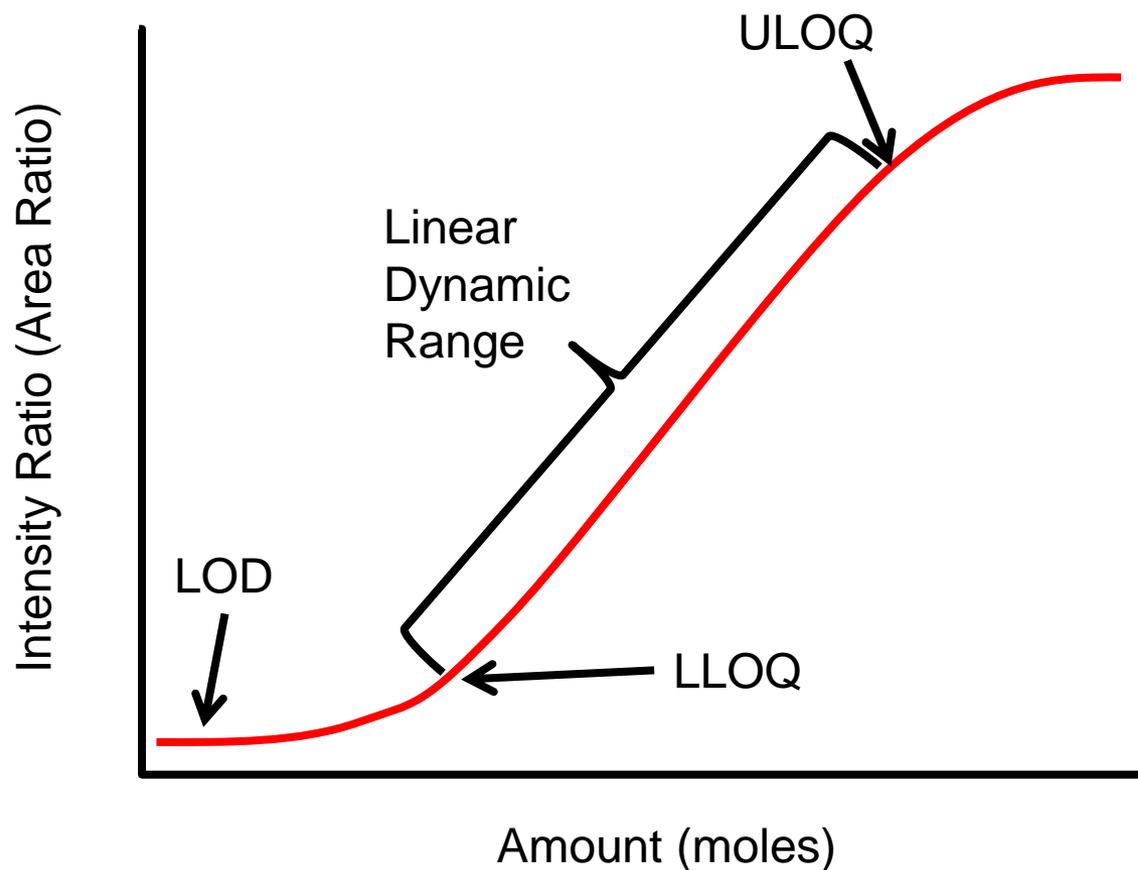
Destroy anti-reagent antibodies

Quantitative Analysis 101



LOD: Limit of Detection
LLOQ: Lower Limit of Quantitation
ULOQ: Upper Limit of Quantitation

Quantitative Analysis 101: Using an Internal Std

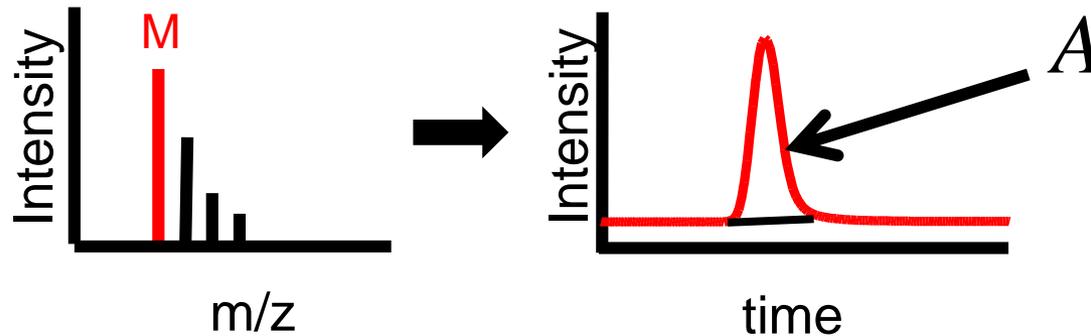


LOD: Limit of Detection
LLOQ: Lower Limit of Quantitation
ULOQ: Upper Limit of Quantitation

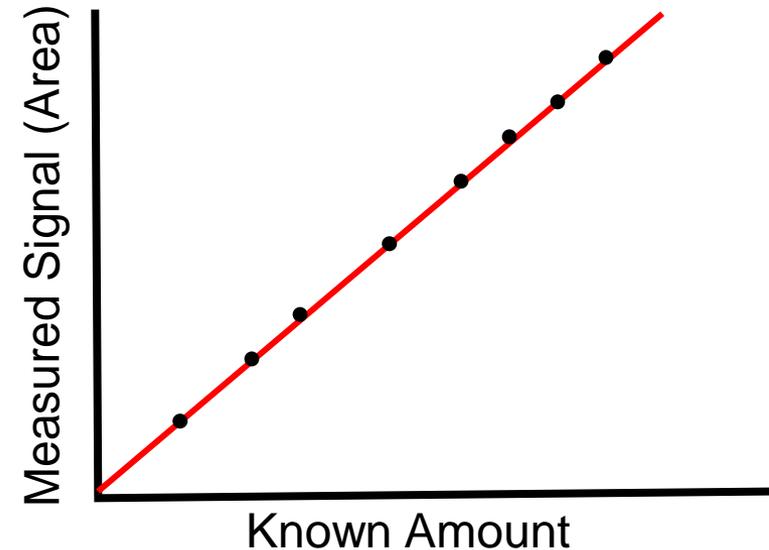
Use of calibration curve to convert signal to quantity

Analyte ($m/z = M$)

$A \propto \text{moles}$



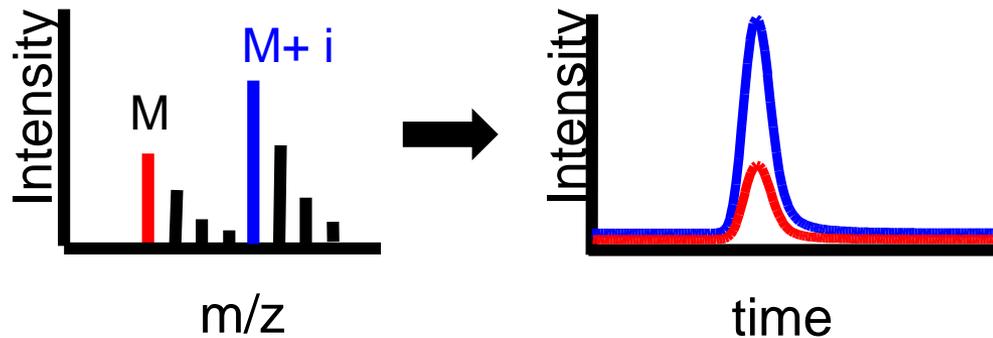
$$A = k \cdot n + A_b \qquad n = \left(\frac{A - A_b}{k} \right)$$



Where: $k_i =$ is the slope of the standard curve
 $A_b \approx 0$ Area from a blank

Use of calibration curve to convert signal to quantity

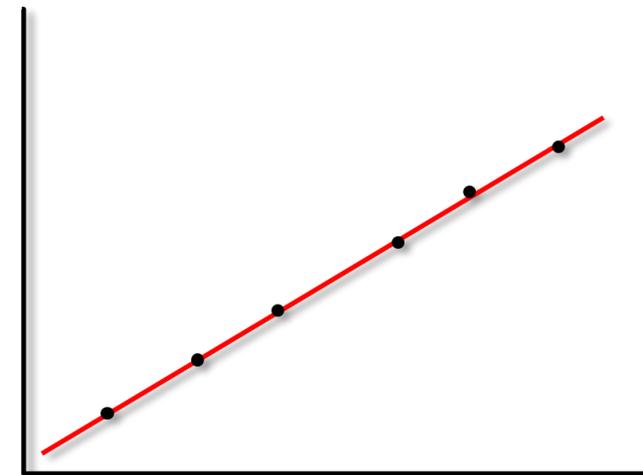
Sample ($m/z = M$) Spiked with
Internal Standard ($m/z = M+i$)



$$R_0 = \frac{A_0}{A_i} \propto \frac{n_0}{n_i}$$

$$R_0 = k_{0/i} \cdot n_0 + R_i \quad n_0 = \left(\frac{R_0 - R_i}{k_{0/i}} \right)$$

Measured Signal (Area Ratio)

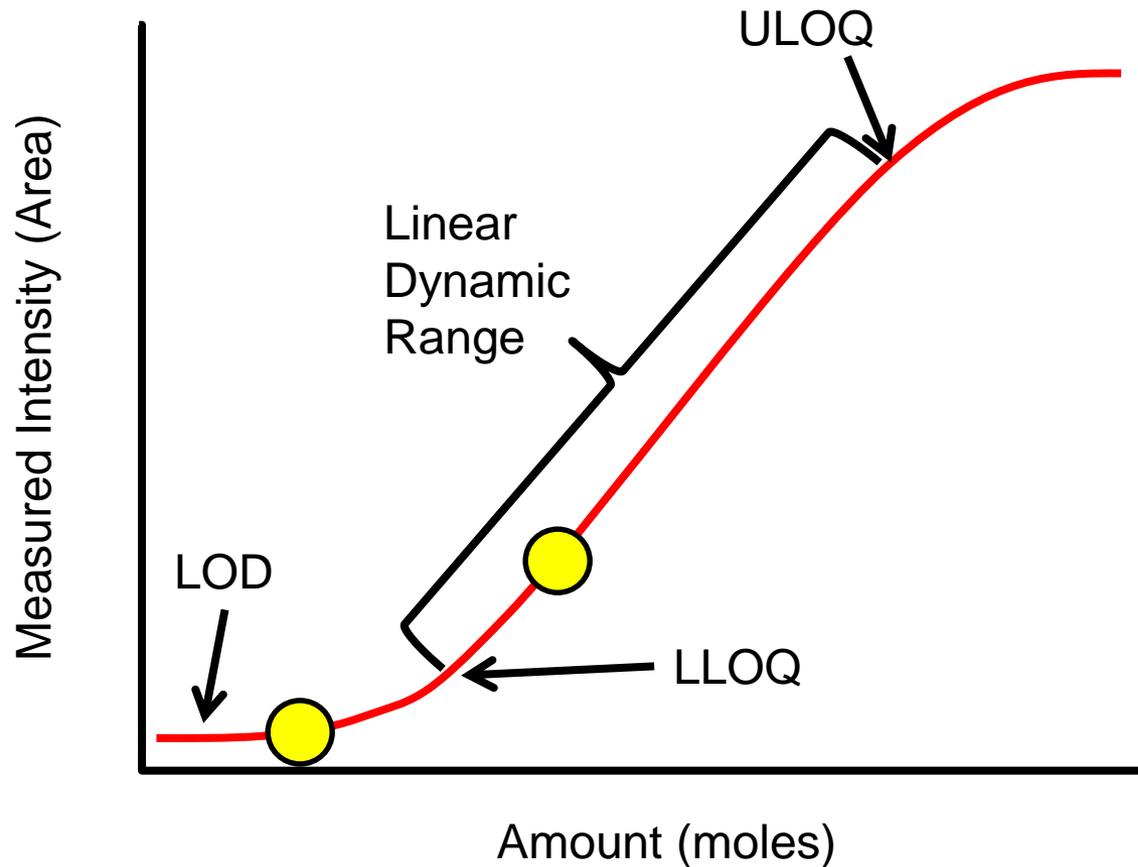


Known Amount

Where: $k_{0/i}$ is the slope of the standard curve

$R_i \approx 0$ Area ratio from a blank ... only internal std

What does it mean if these are the two points?

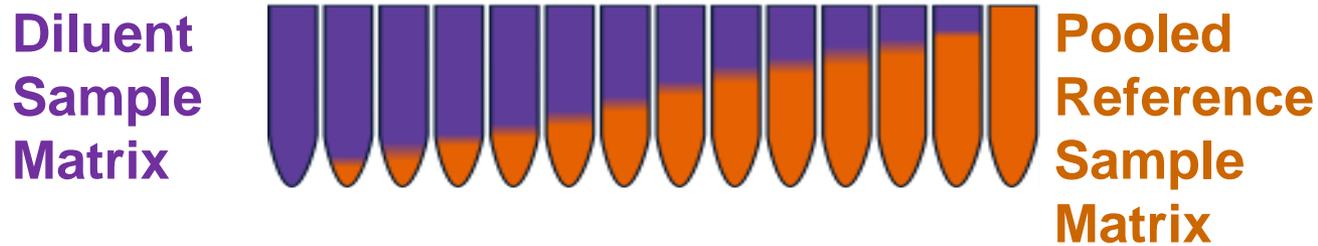


LOD: Limit of Detection
LLOQ: Lower Limit of Quantitation
ULOQ: Upper Limit of Quantitation

Questions:

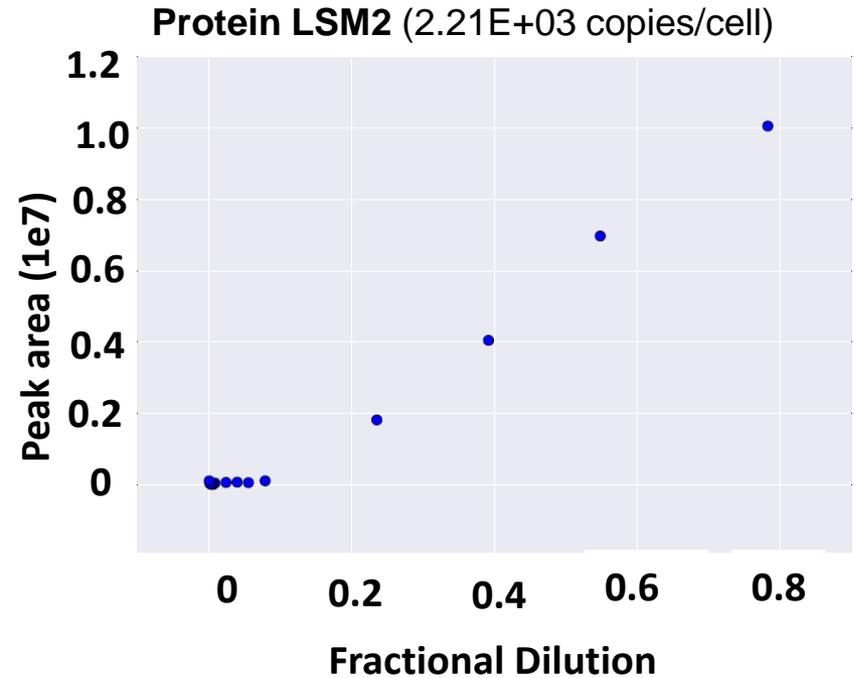
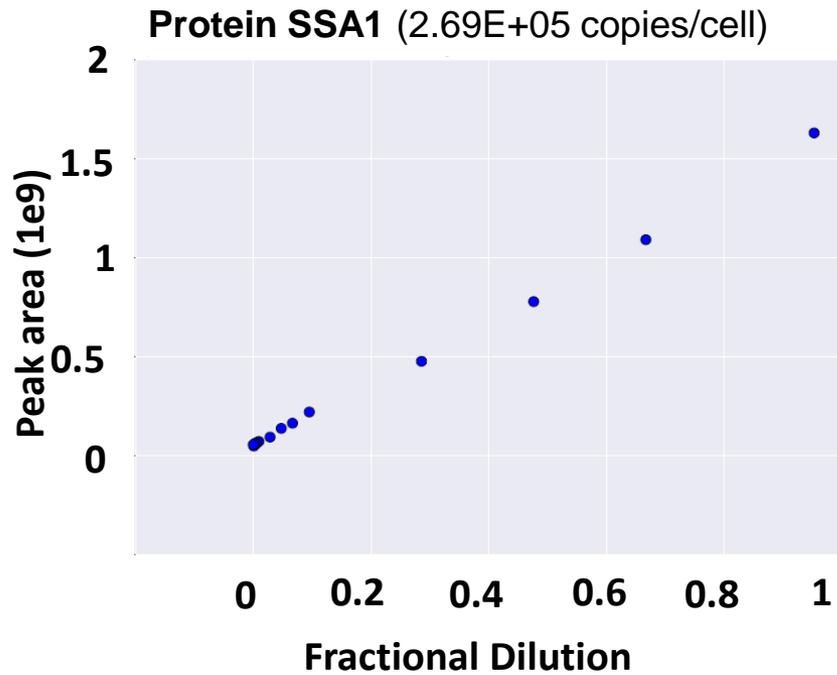
- Design an approach to confirm linearity and LLOQ for peptides in a complex matrix?
 - Assume you have < 50 peptides in your assay.
 - Assume you have >1000 peptides in your assay?
- Design an approach to “calibrate” your method
 - The purpose of the calibration is:
 - To correct for differences between instruments and platforms
 - Minimize between day batch effects
 - Enable another lab to get similar results to your lab

Method to measure both LOQ and Linearity



- Possible Samples to Use as a Diluent Matrix
 - Stable isotope labeled version of the matrix.
 - ^{15}N or SILAC labeled cells
 - A diverged species
 - For human plasma we use chicken plasma.

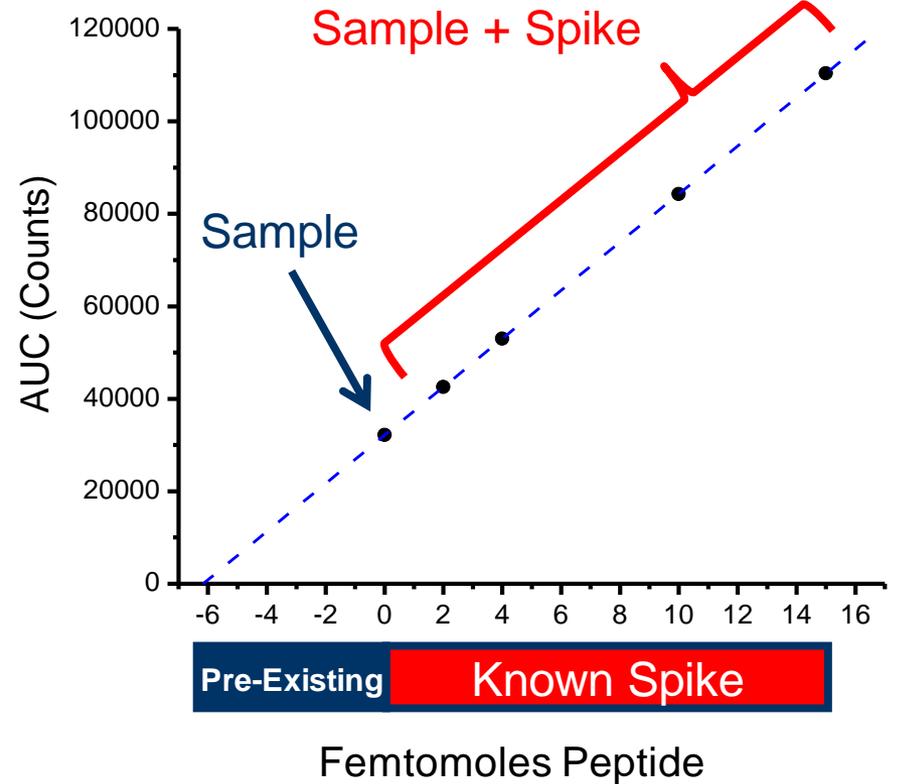
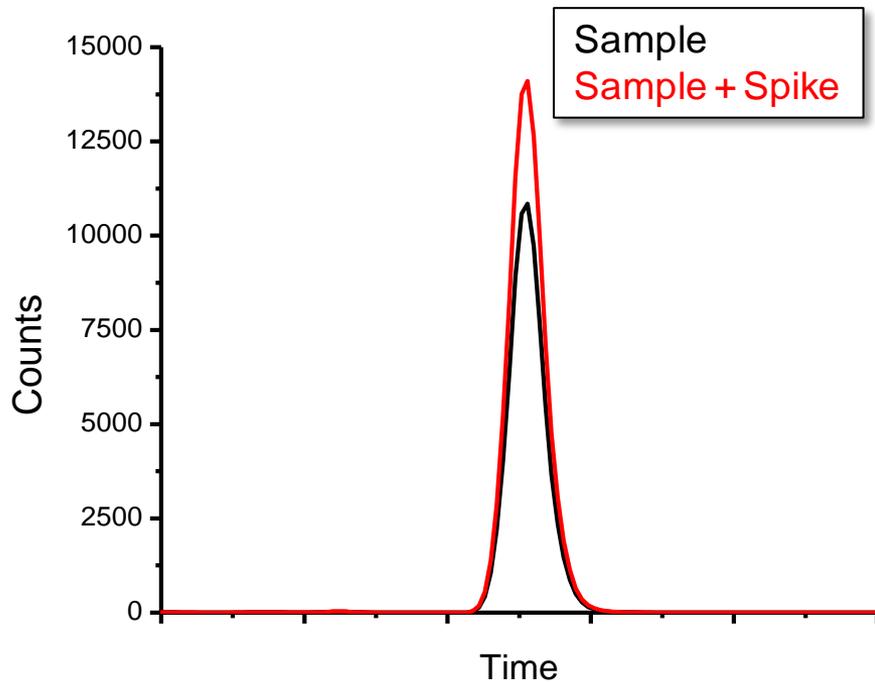
Reference Yeast BY4742 Diluted in ^{15}N Yeast (S288c)



Mixing two samples of extreme conditions

- Assume that you have a healthy and a disease cohort.
- You can take a pool of the healthy and a pool of the disease to demonstrate that the measurement is linear between the two conditions
- Example
 - 100% condition A
 - 80% condition A, 20% condition B
 - 50% condition A, 50% condition B
 - 20% condition A, 80% condition B
 - 100% condition B

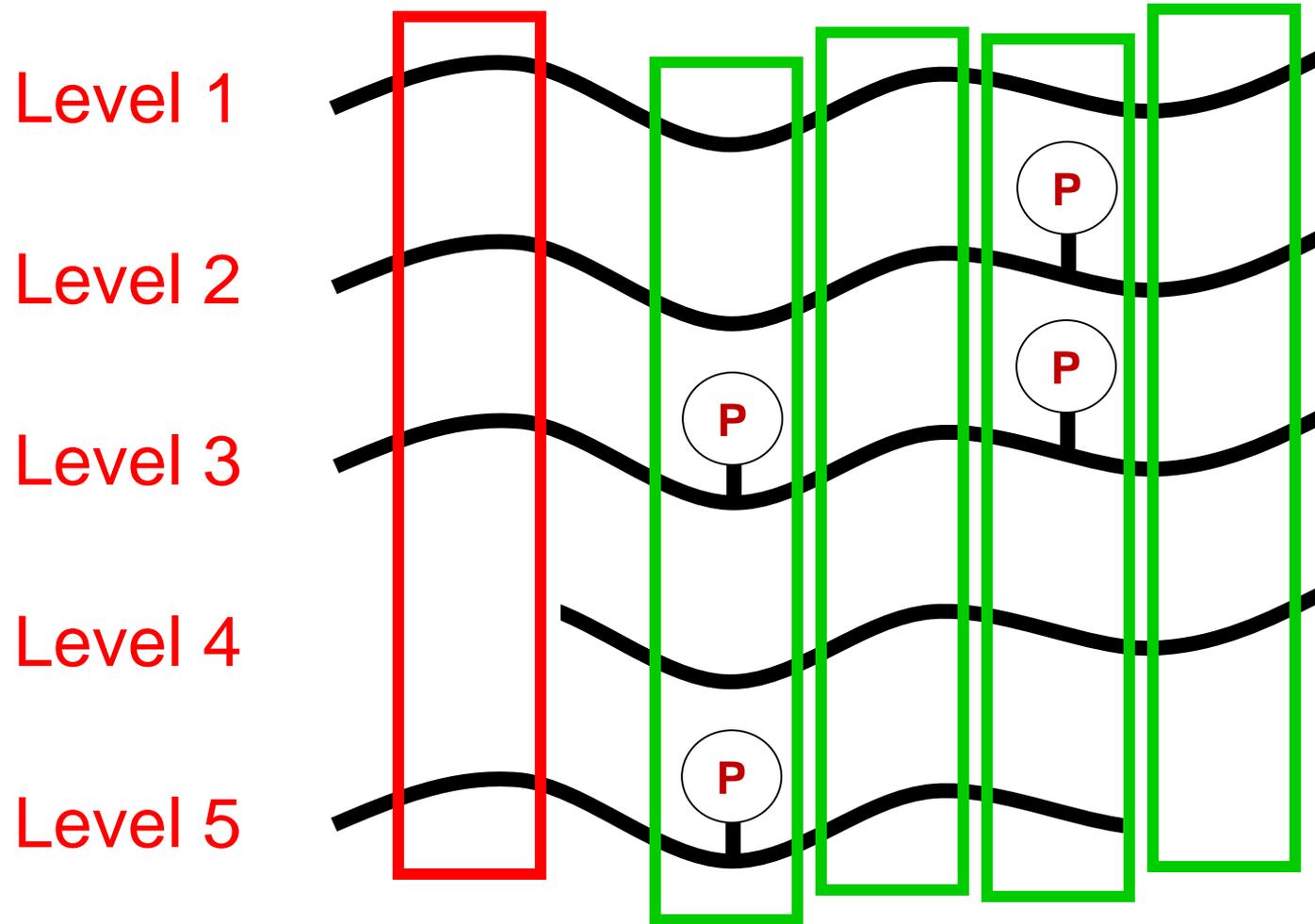
Method of Standard Addition



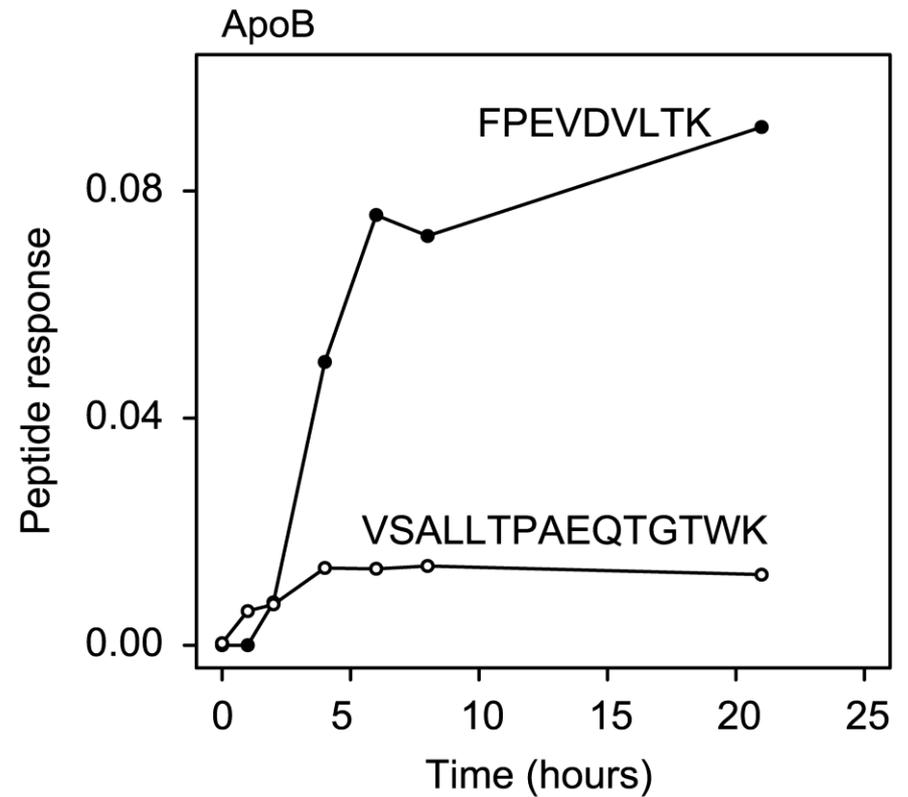
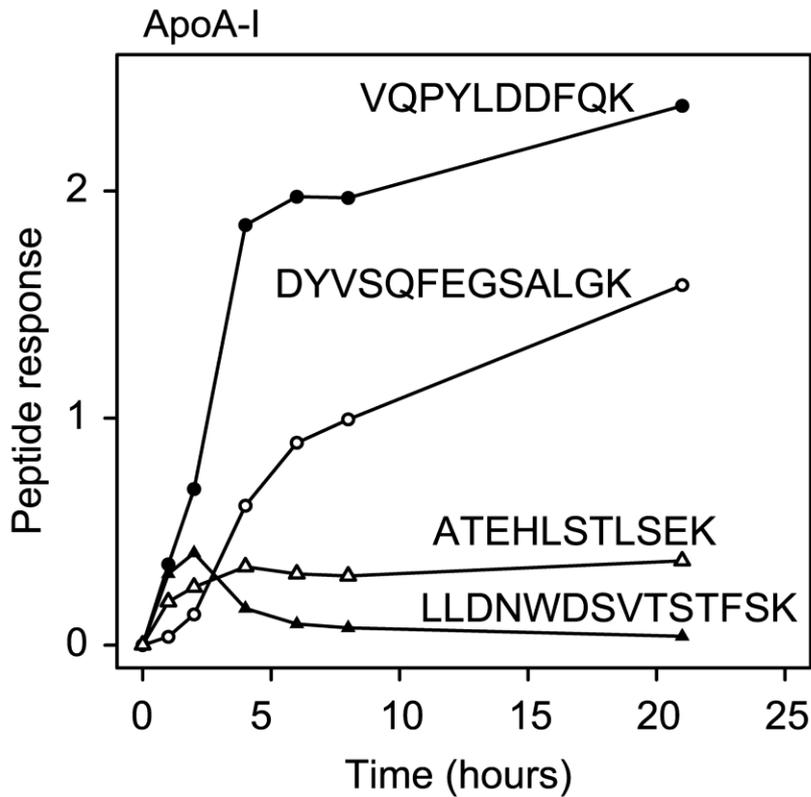
Can't assess Limit of Quantification. Can only assess whether the sample is above the LOQ

OTHER THINGS TO CONSIDER

We are not doing protein quantitation we are doing peptide quantitation!

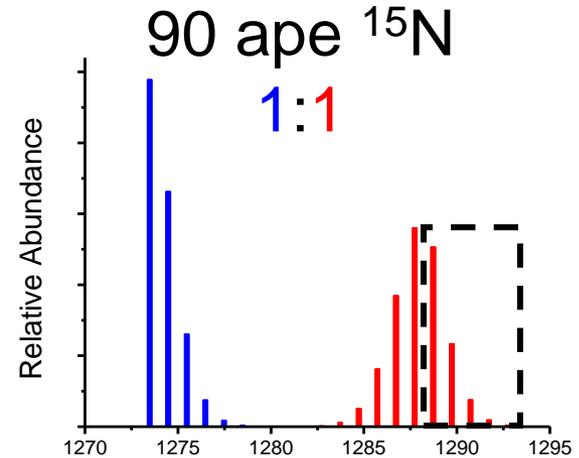
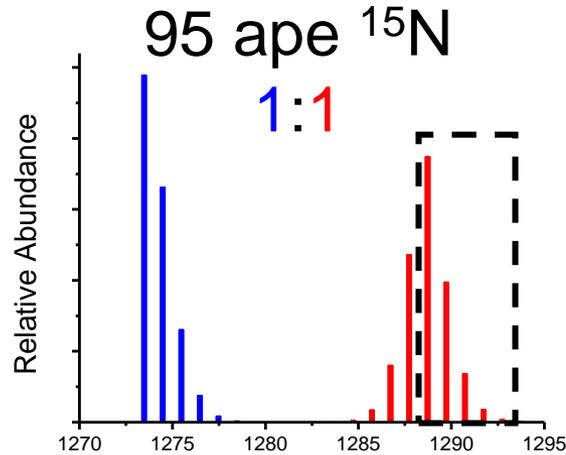
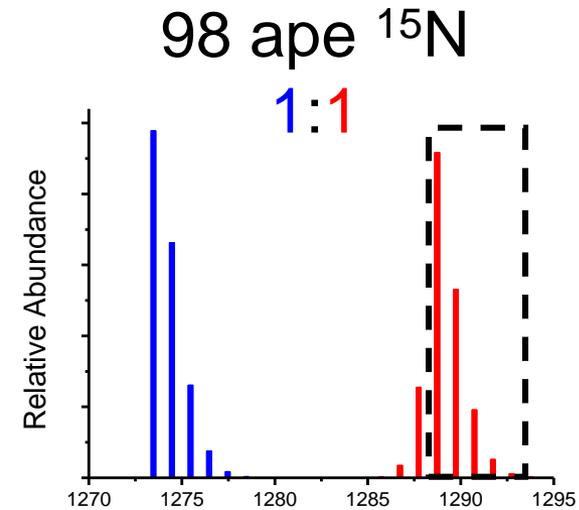
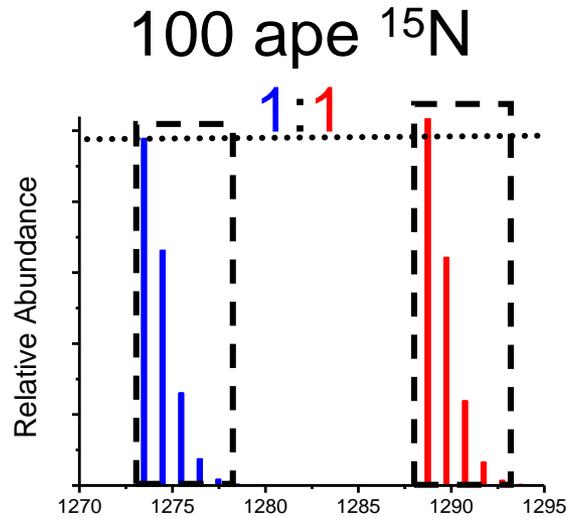


Peptide Intensity Varies Over Course of Digestion



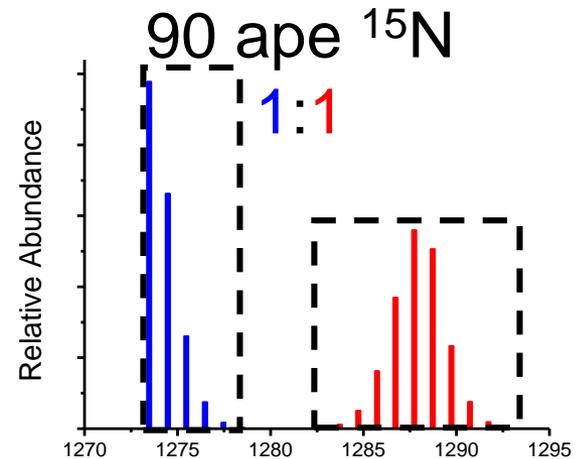
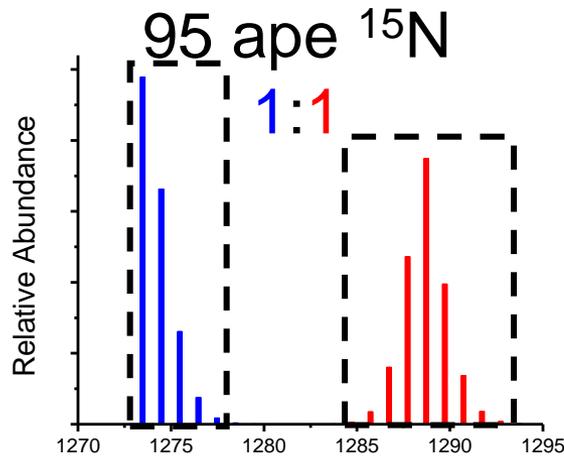
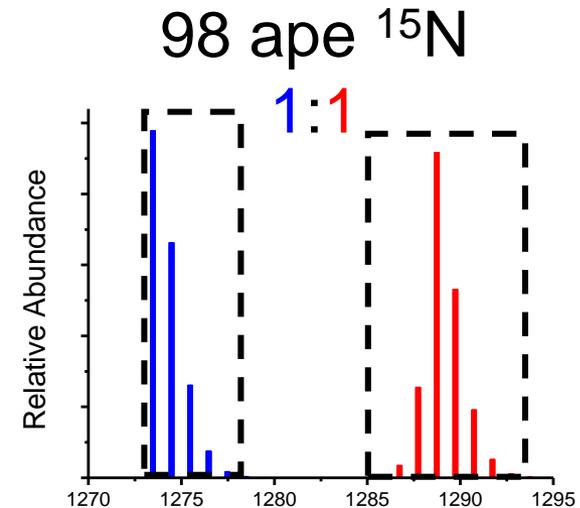
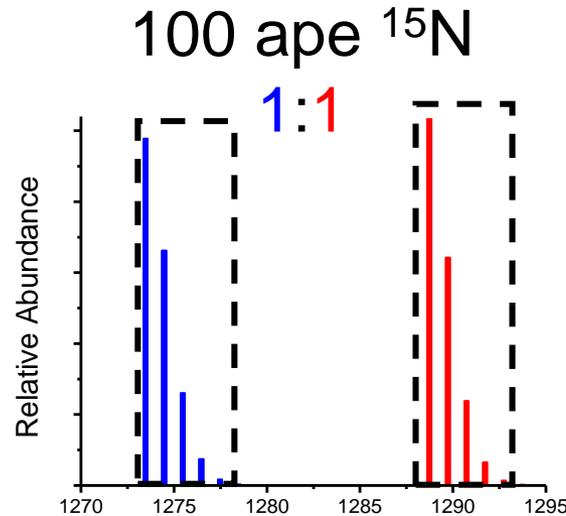
Do you know the enrichment of the standard?

Peptide: DIDIEYHQNK N = 15



Do you know the enrichment of the standard?

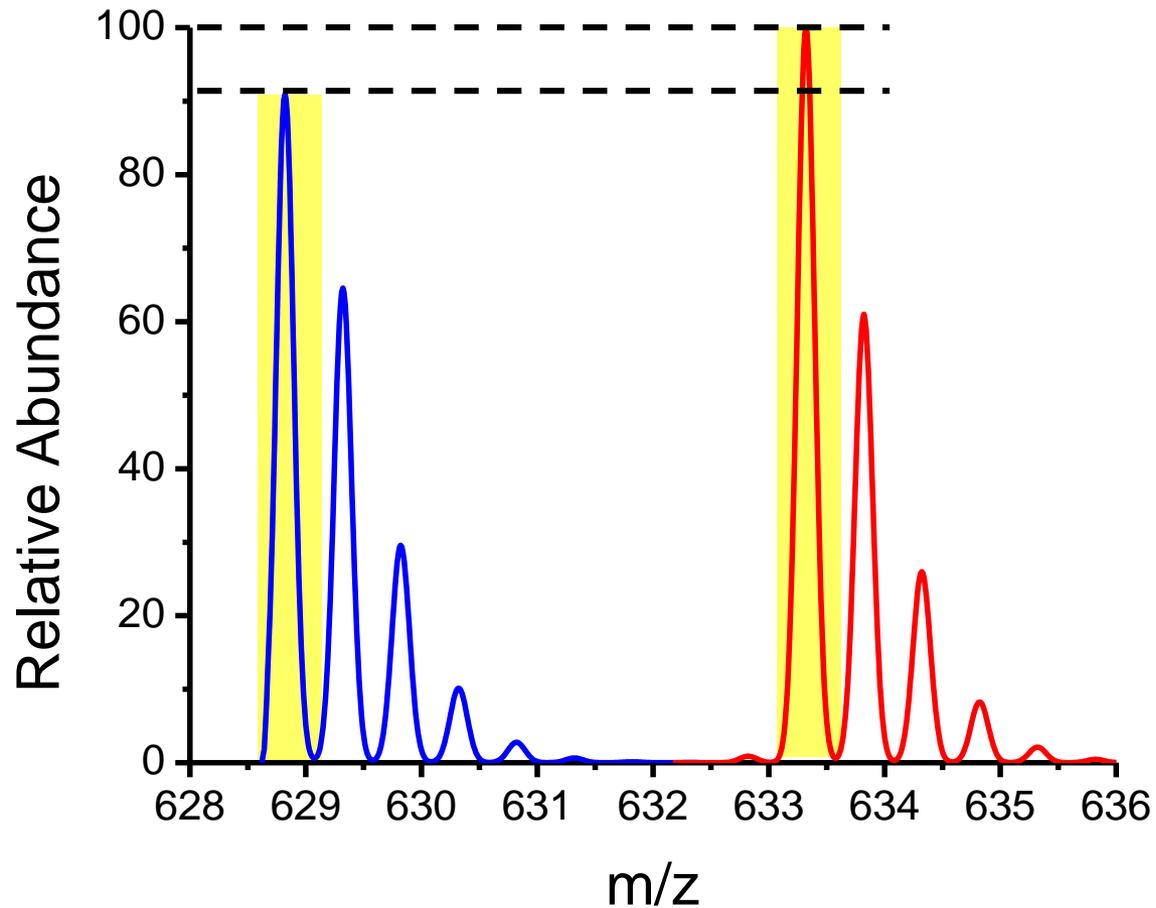
Peptide: DIDIEYHQNK N = 15



Effect of ^{13}C Labeling on Quantitative Accuracy

YAGILDC_{ICAT}FK

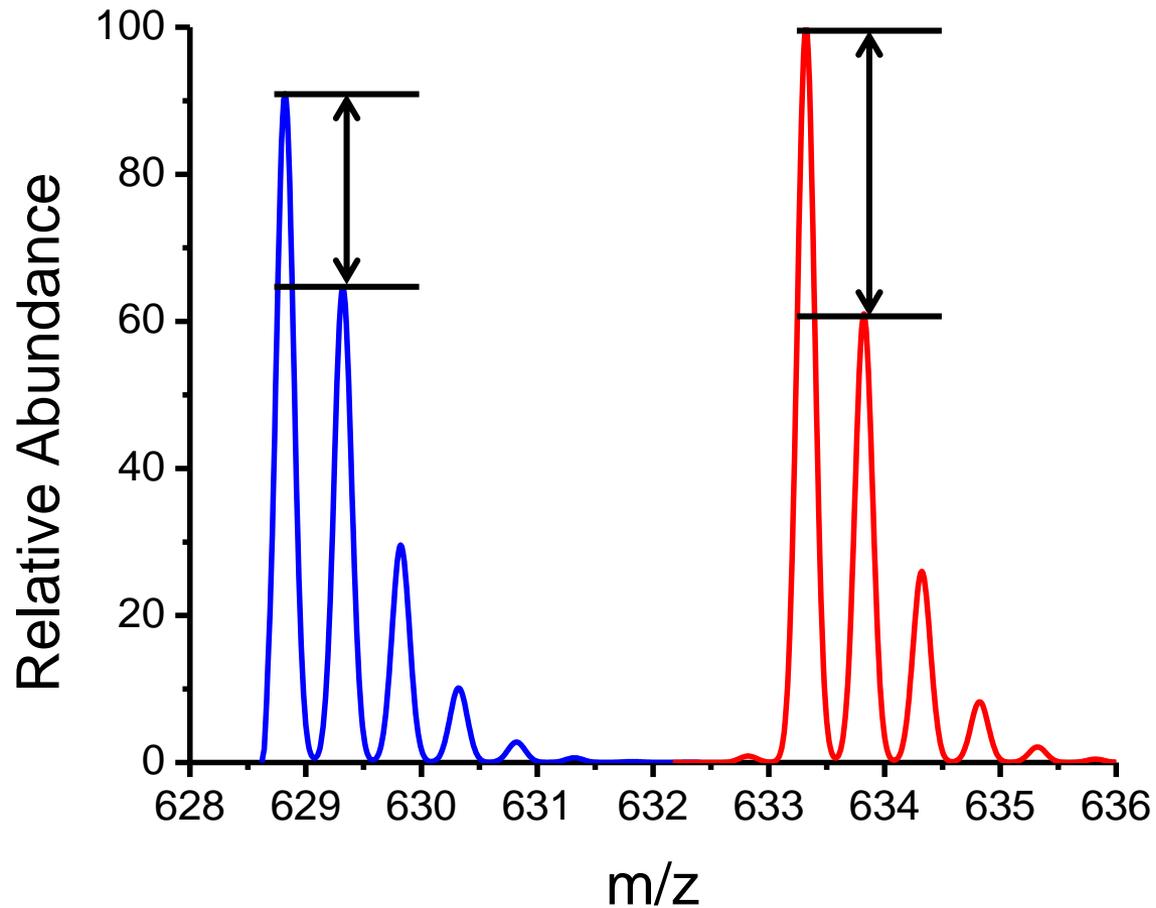
9 x ^{13}C at >99.9 APE



Effect of ^{13}C Labeling on Quantitative Accuracy

YAGILDC_{ICAT}FK

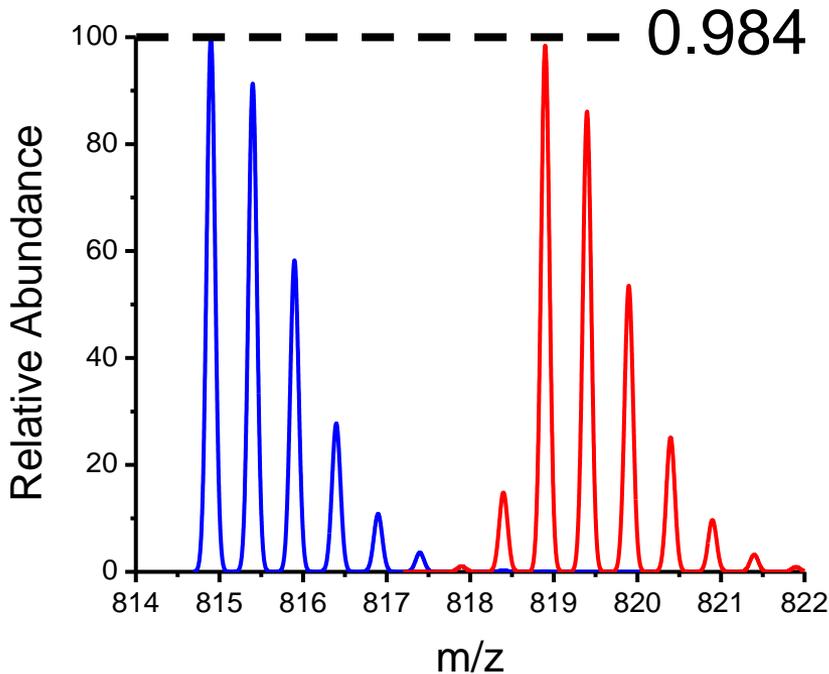
9 x ^{13}C at >99.9 APE



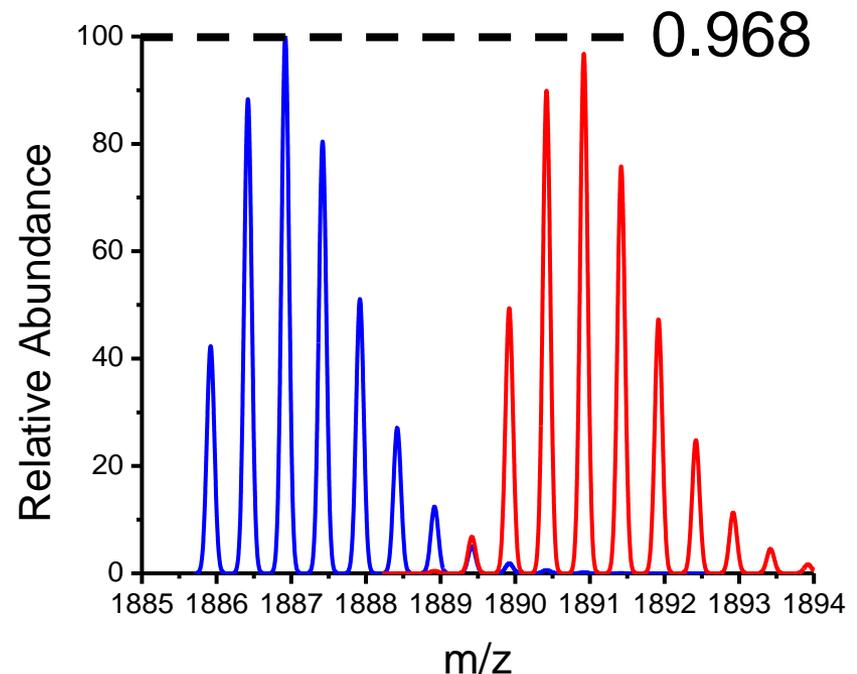
Effect of 98 ape Labeling on $^2\text{H}_8$ -ICAT Accuracy

FGTWQC_{ICAT}LMR

SYIEGTAVSQADFGLGS
GHC_{ICAT}QALDPWVTVFK



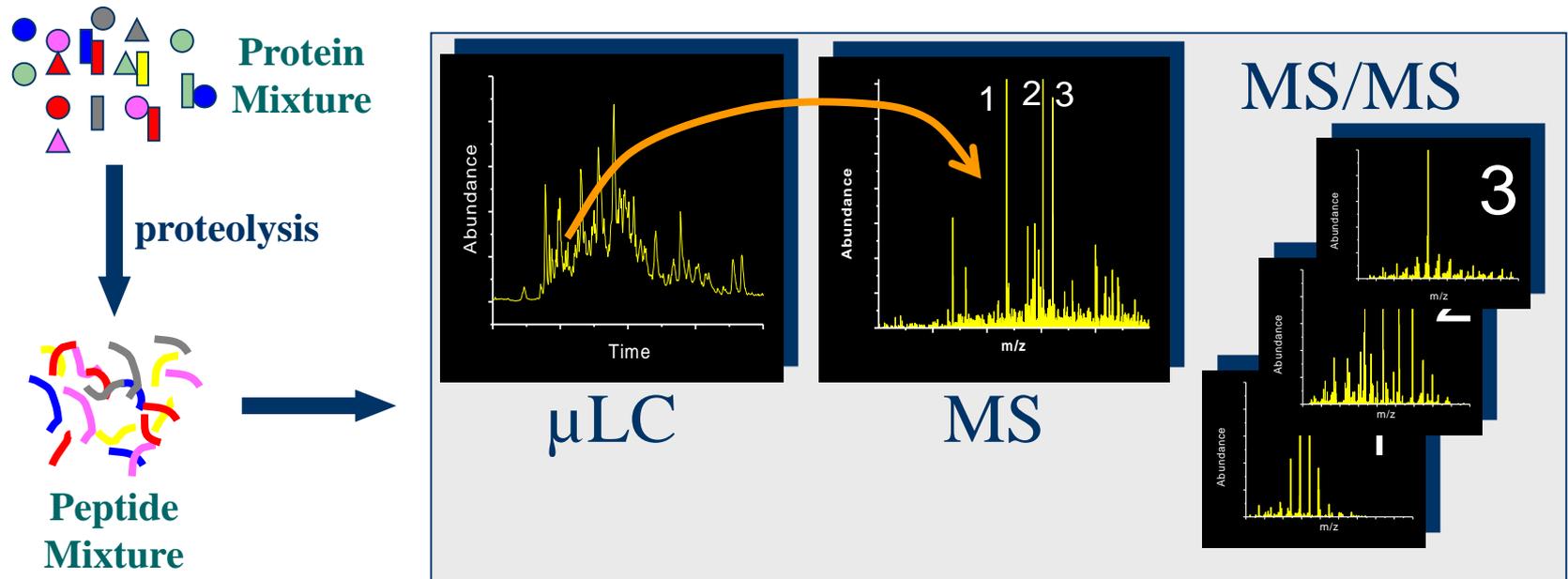
1:1 Mole Ratio



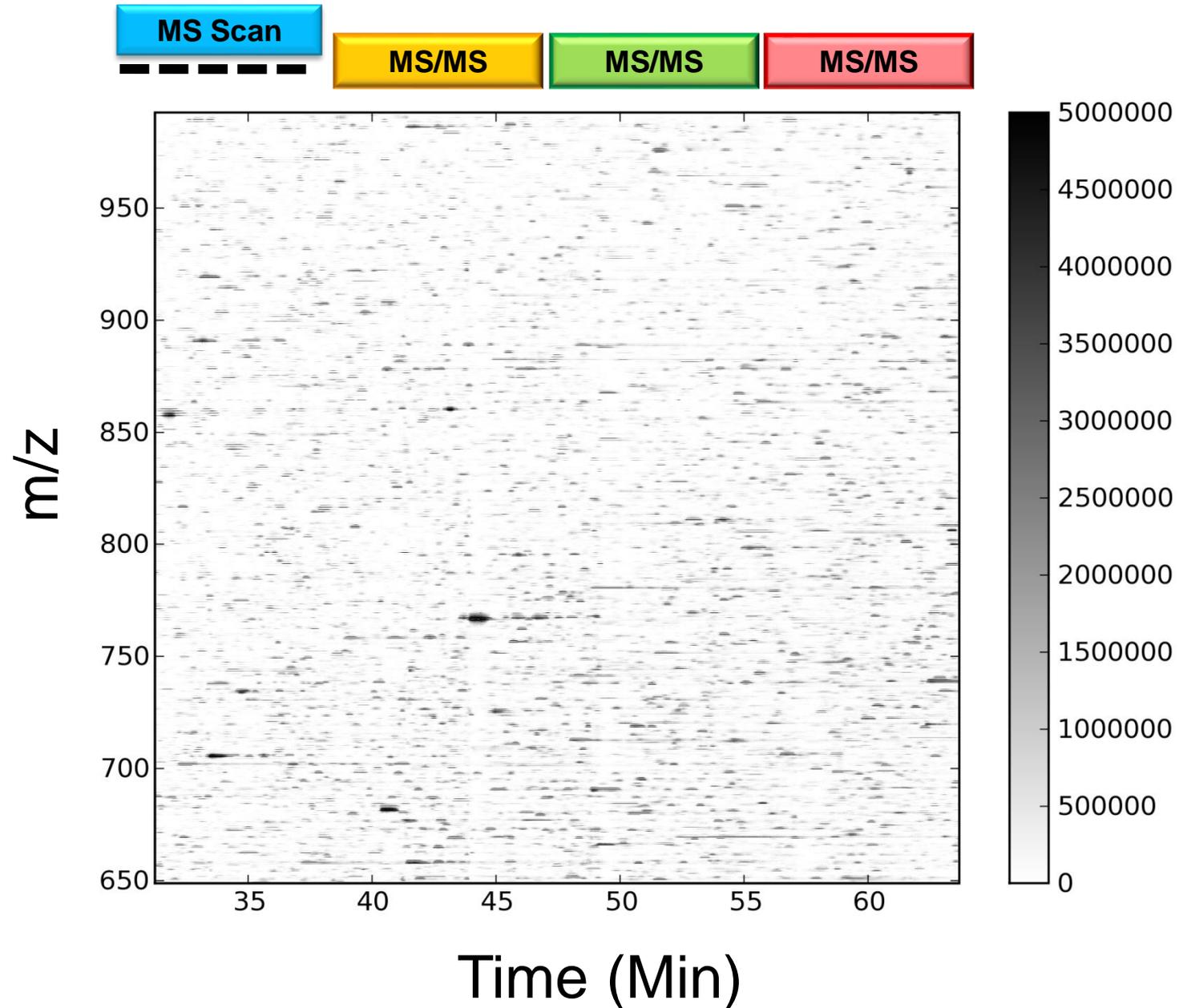
1:1 Mole Ratio

Challenges with Discovery Proteomics and Possible Alternatives

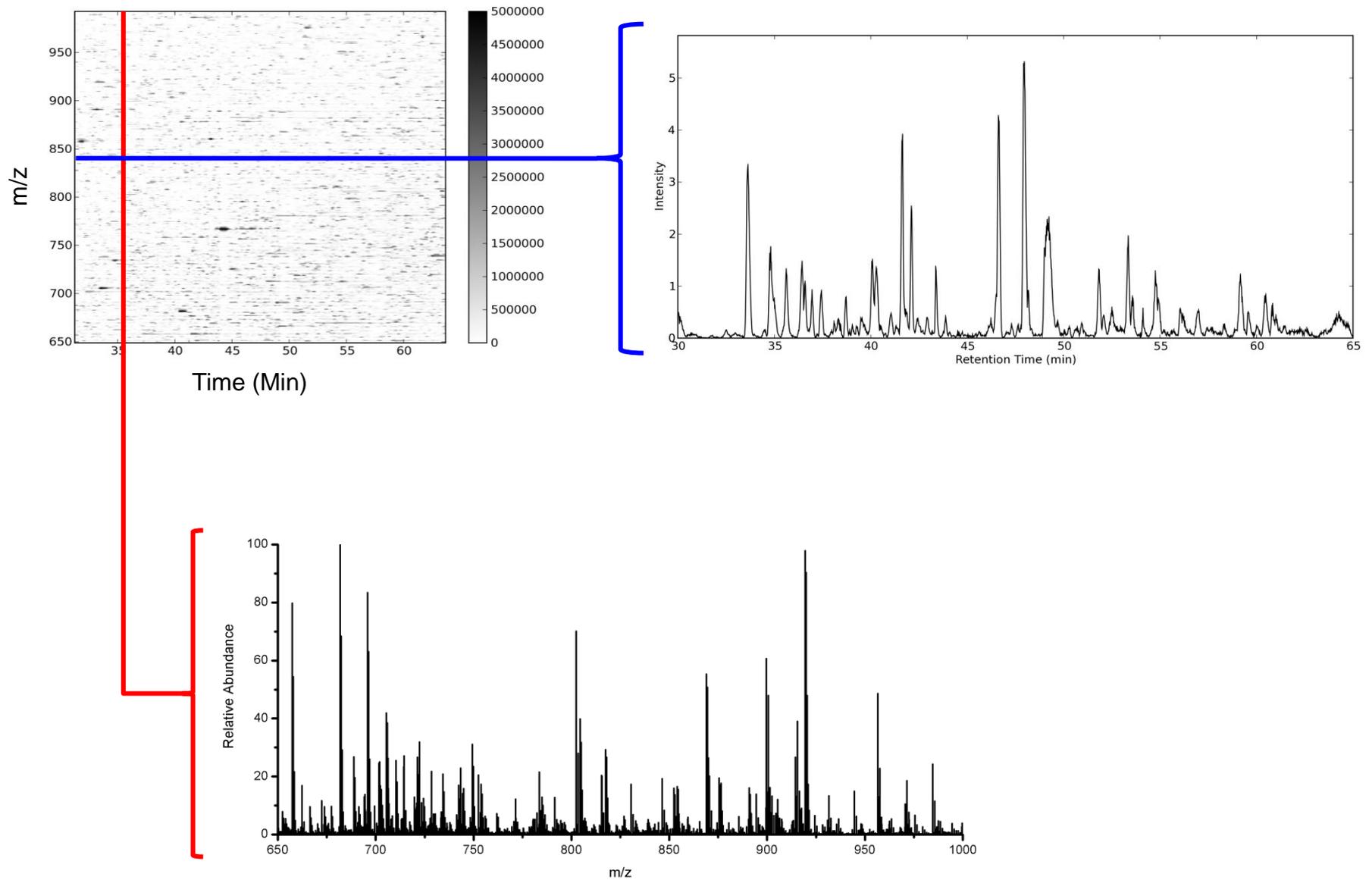
Shotgun Proteomics



These Mixtures are Complicated!

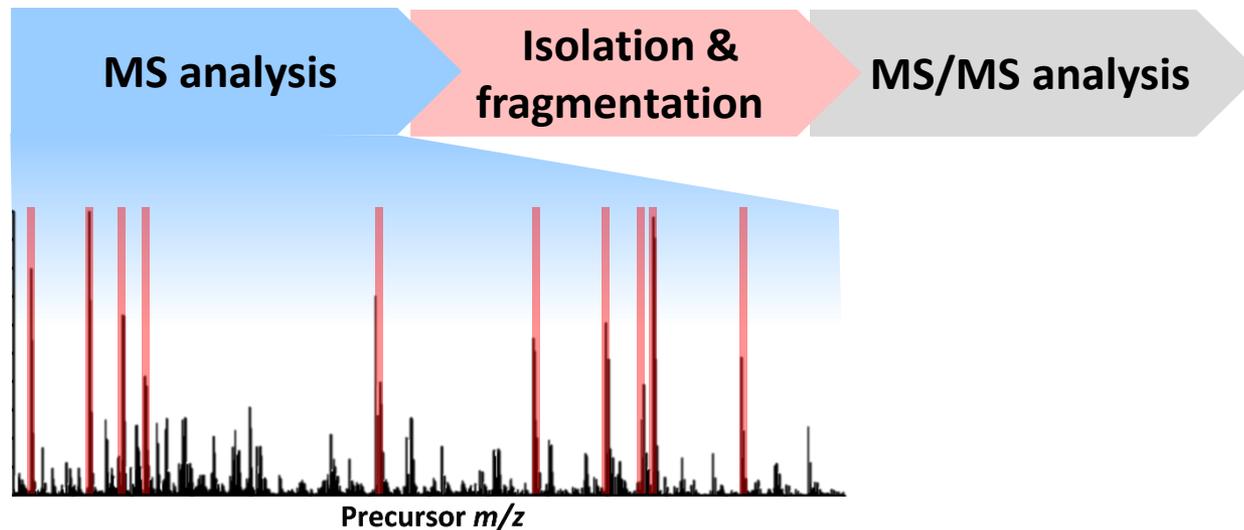


These Mixtures are Complicated!

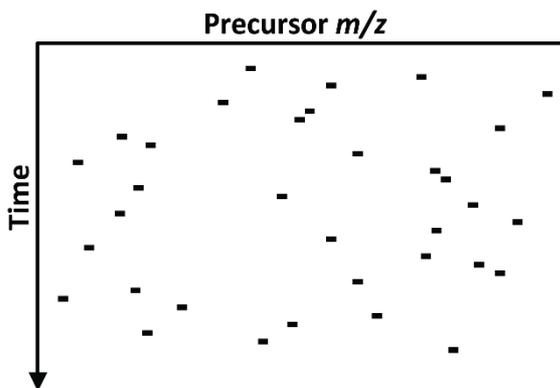


Acquisition methods in shotgun LC-MS/MS

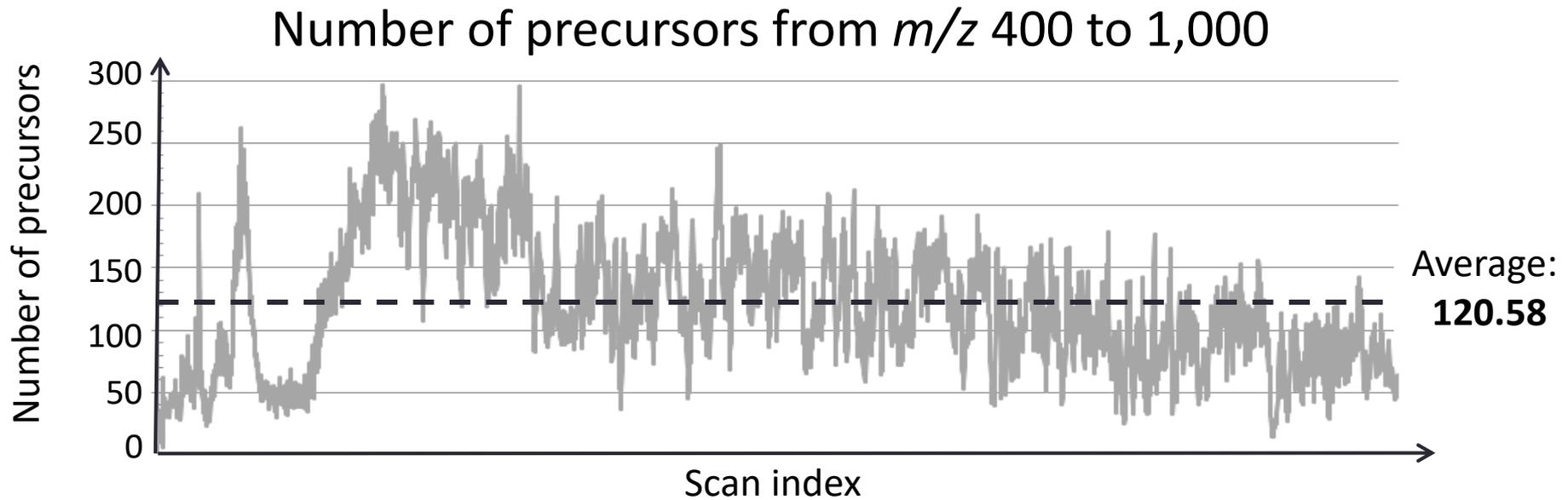
Liquid chromatography



Data Dependent
Acquisition (DDA)



Sampling Every Precursor Using DDA

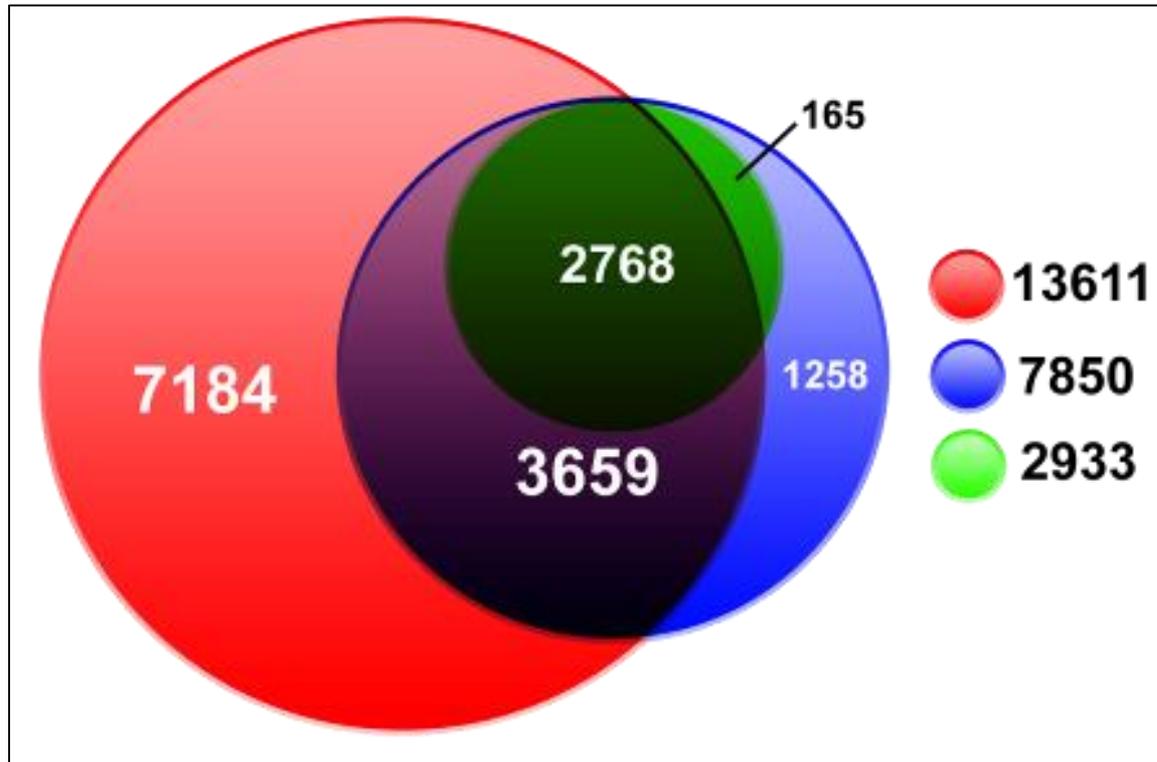


- To get every precursor using DDA, we would have to collect 121 MS/MS spectra on average per chromatographic peak

Data Dependent Acquisition will Always Undersample



Data Dependent Acquisition will Always Under Sample

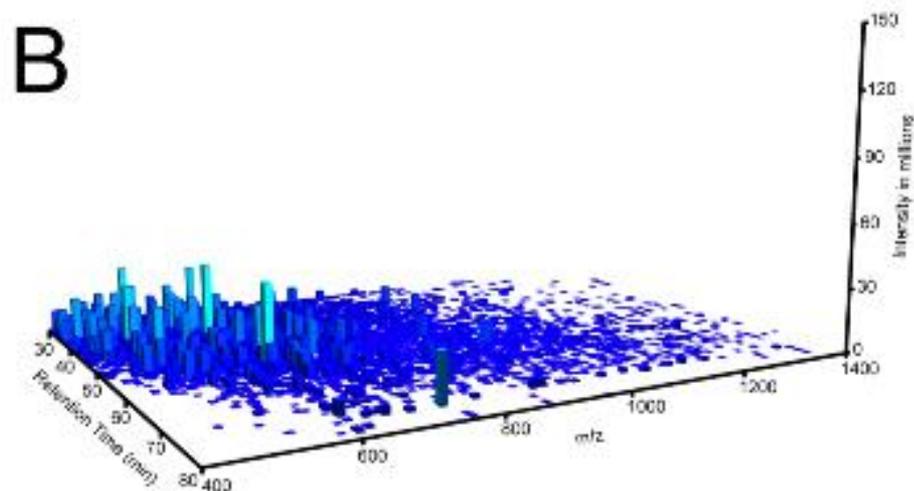
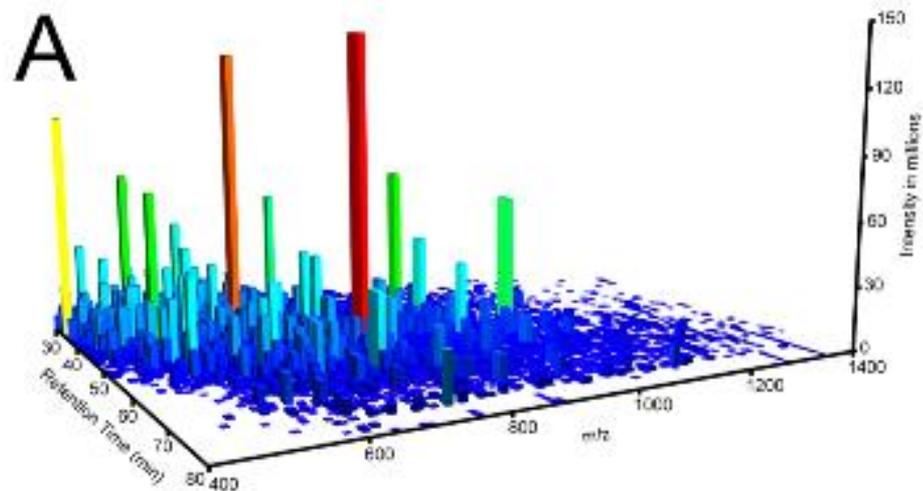


Red: Detected Persistent Peptide Isotope Distributions

Blue: MS/MS Spectra

Green: Peptide Identifications

DDA always misses and is least reproducible on the low abundant signals



Imagine Visiting Kyoto and Visiting the Tallest Buildings First



Copyright Rob Laddish

Analogy Nate Yates

You would miss ...

If you wouldn't learn much about Kyoto by going to the tallest structures why would we use this approach to understand protein mixtures???

So what do we do?

***Build on Other People's
Knowledge***

So what do we do?

kyoto places to visit - Google

https://www.google.com/search?q=map+of+kyoto&hl=en&prmd=imvnsa&source=lnms&tbm=isch&sa=X&ei=VWxZUO6iKazDmQXC94I

Read Later PubMed Home PurchasePATH Login MyFinancial.deskto... Start Page: /MacCoss MSDaPI UW Libraries Proxy B... ADT Pulse(TM) Inter... Other bookmarks

+Michael Search Images Maps Play YouTube News Gmail Drive Calendar More

Google kyoto places to visit maccoss@gmail.com 0 + Share

Search About 2,120,000 results (0.11 seconds) Safe Search moderate

Web
Images
Maps
Videos
News
Shopping
More

Any time
Past 24 hours
Past week
Custom range...

All results
By subject

Any size
Large
Medium
Icon



www.google.com/imgres?num=10&hl=en&biw=1280&bih=675&tbm=isch&tbnid=W1aL1DpDF8rCM:&imgrefurl=http://www.besttraveldestinationinasia.com/the-best-honeymoon-destinations-in-asia-best-places-to-visit-in-asia&...

Windows taskbar: Internet Explorer, File Explorer, Google Chrome, PowerPoint, Word, System tray: (4:10), 5:45 PM 9/19/2012

So what do we do?

kyoto people - Google Search

https://www.google.com/search?q=map+of+kyoto&hl=en&prmd=imvnsa&source=lnms&tbm=isch&sa=X&ei=VWxZUO6iKazDmQXC94I

Read Later PubMed Home PurchasePATH Login MyFinancial.deskto... Start Page: /MacCoss MSDaPI UW Libraries Proxy B... ADT Pulse(TM) Inter... Other bookmarks

+Michael Search Images Maps Play YouTube News Gmail Drive Calendar More

Google kyoto people maccoss@gmail.com + Share

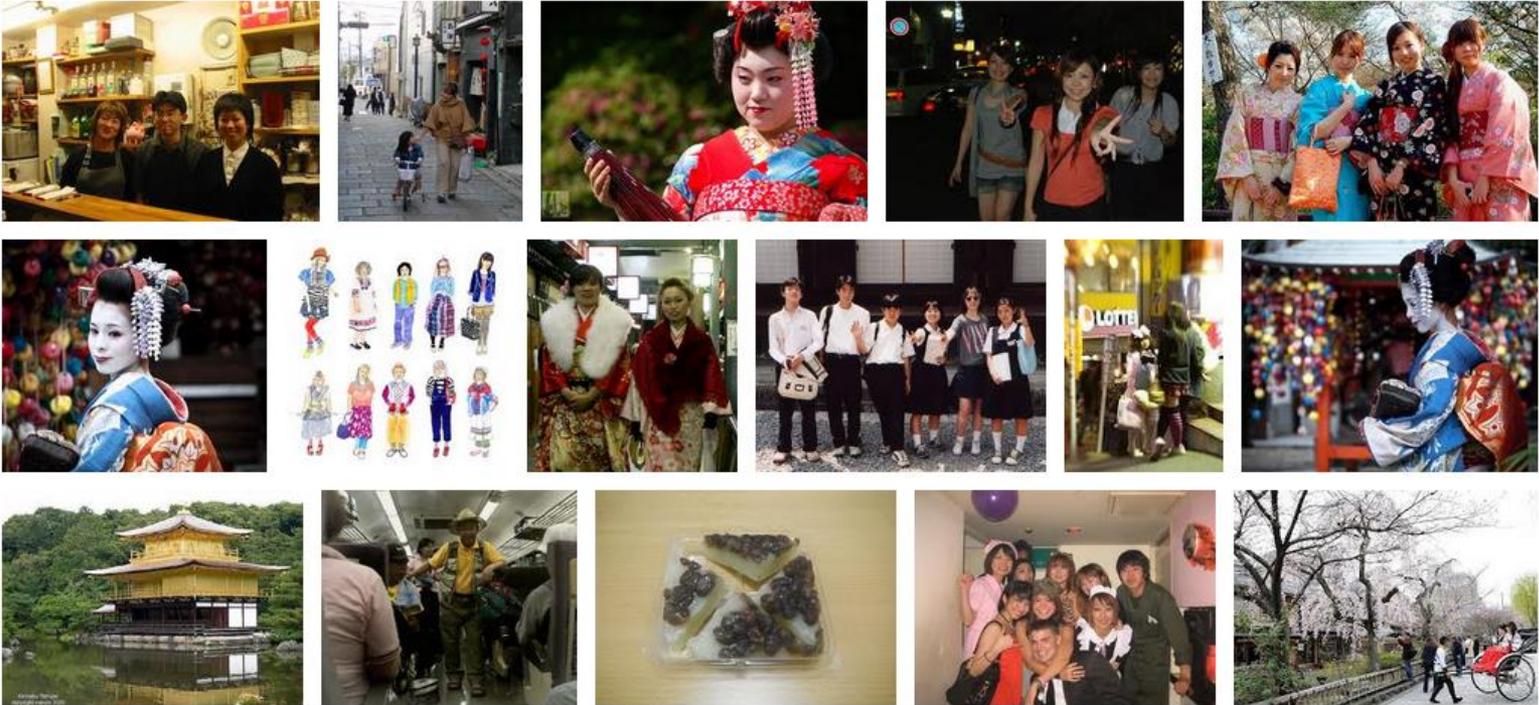
Search About 26,600,000 results (0.34 seconds) Safe Search moderate

Web
Images
Maps
Videos
News
Shopping
More

Any time
Past 24 hours
Past week
Custom range...

All results
By subject

Any size
Large
Medium
Icon
Larger than...



Windows taskbar: Internet Explorer, Firefox, Chrome, PowerPoint, Word, System tray: (4:10), 5:46 PM 9/19/2012

So what do we do?

kyoto food - Google Search

https://www.google.com/search?q=map+of+kyoto&hl=en&prmd=imvnsa&source=lnms&tbm=isch&sa=X&ei=VWxZUO6iKazDmQXC94I

Read Later PubMed Home PurchasePATH Login MyFinancial.deskto... Start Page: /MacCoss MSDaPI UW Libraries Proxy B... ADT Pulse(TM) Inter... Other bookmarks

+Michael Search Images Maps Play YouTube News Gmail Drive Calendar More

Google kyoto food maccoss@gmail.com 0 + Share

Search About 5,670,000 results (0.43 seconds) Safe Search moderate

Web
Images
Maps
Videos
News
Shopping
More

Any time
Past 24 hours
Past week
Custom range...

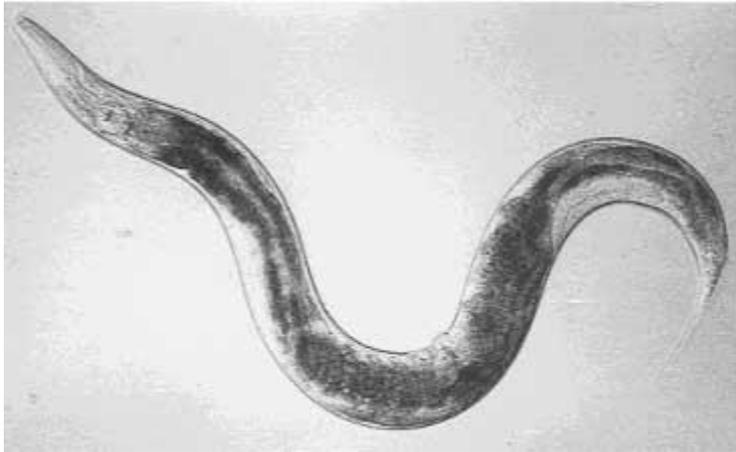
All results
By subject

Any size
Large
Medium
Icon
Larger than...

Page 2

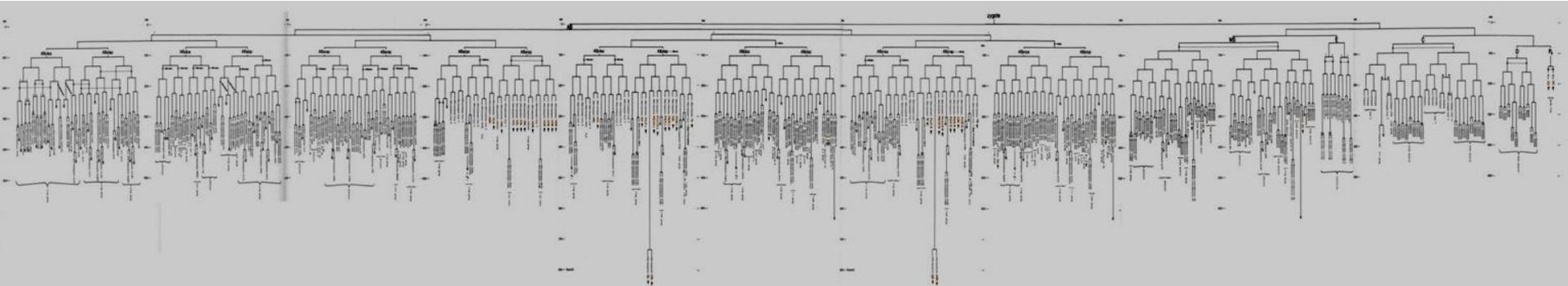


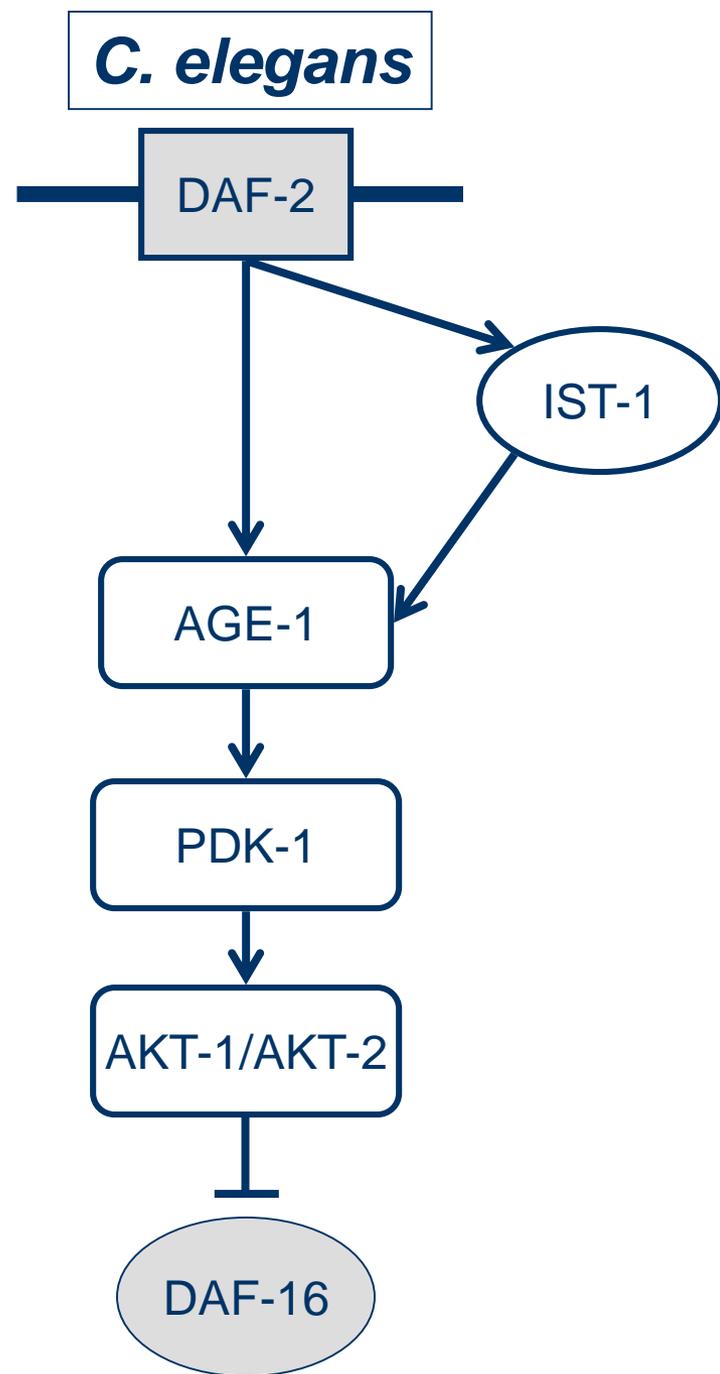
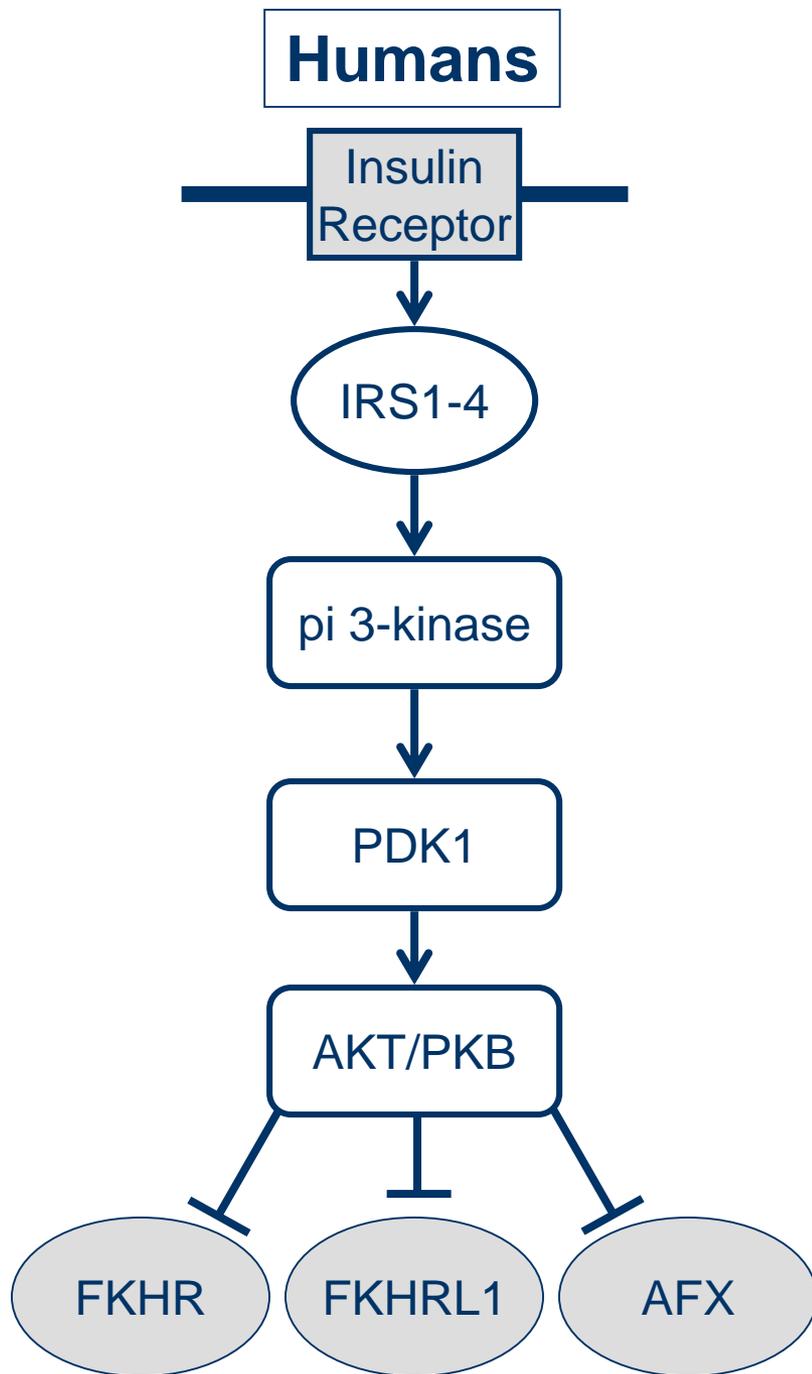
C. elegans as an *In Vivo* Model System for Proteomics



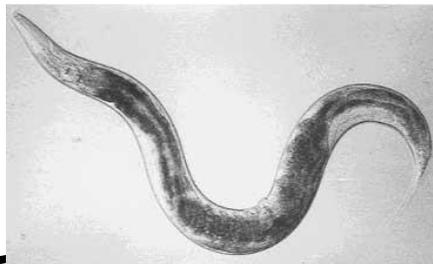
The only multicellular organism with a genome sequenced to completeness

Complete Lineage for all 959 Somatic Cells





Fractionation to Improve the Coverage of Proteins Involved in Insulin / IGF-1 Signaling



0 Proteins Identified that are Known to be Involved in the Insulin / IGF-1 Signaling Pathway

Rev
Pr

s
tion

10 Fractions

8 Fractions

10 Fractions

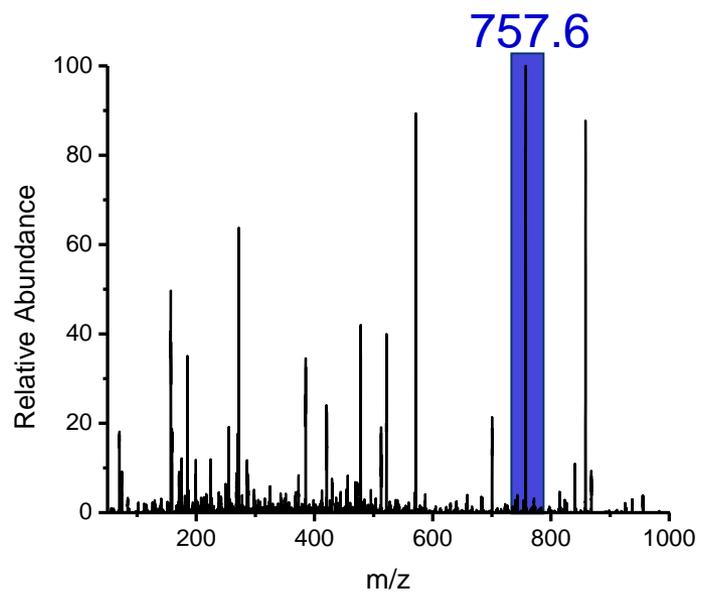
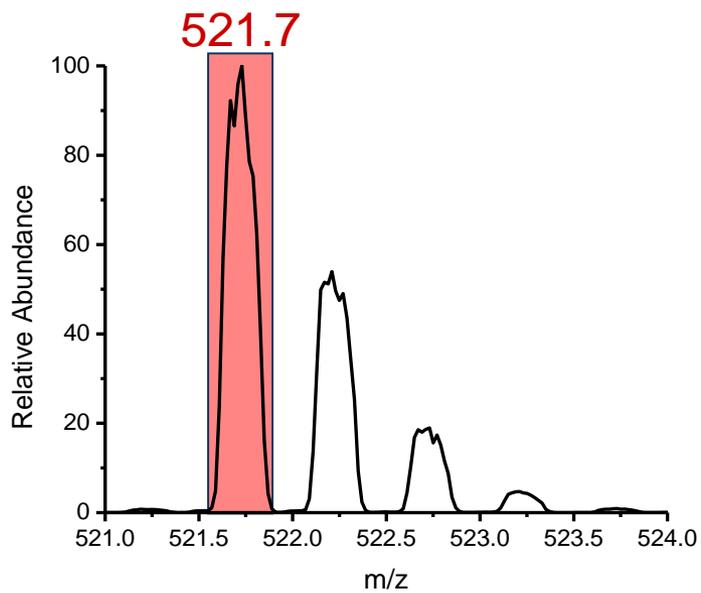
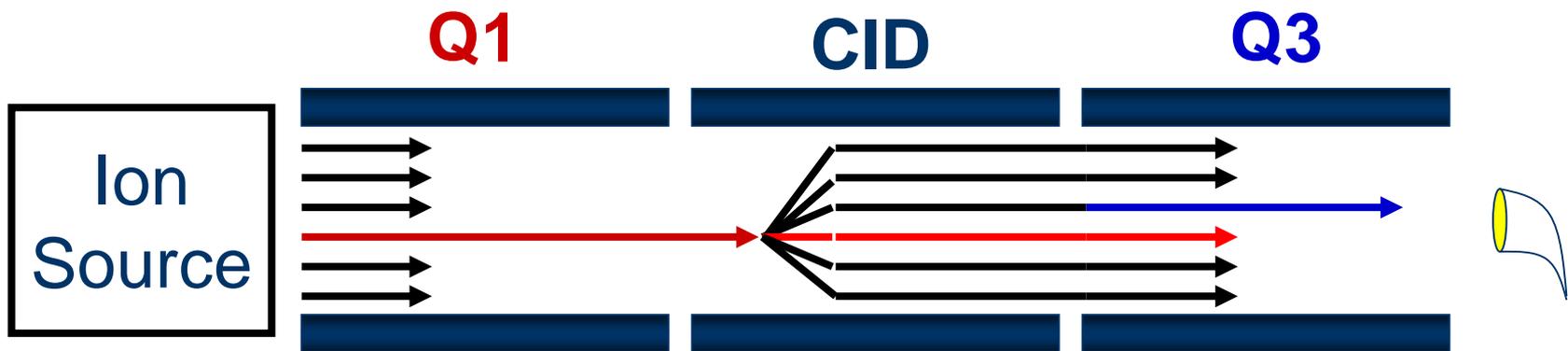


10 MudPITs

8 MudPITs

10 MudPITs

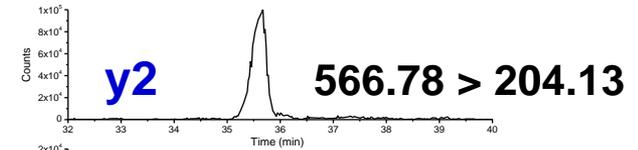
>6,000,000 MS/MS Spectra, 67,047 Unique Peptide Identifications, and 6,779 Unique Protein Identifications



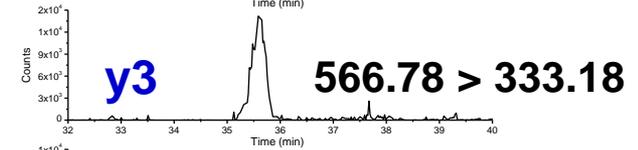
Precursor Ion > Product Ion

SRM Chromatograms

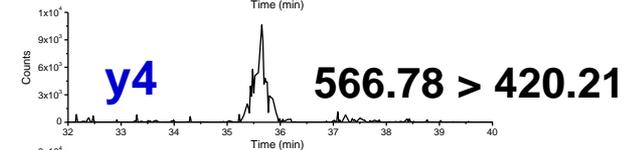
TGASEAVPSEGK > GK



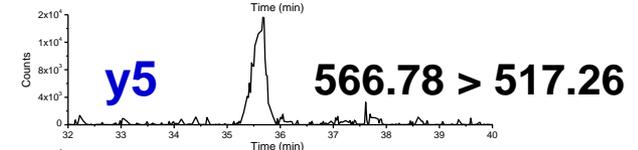
TGSAEAVPSEGK > EGK



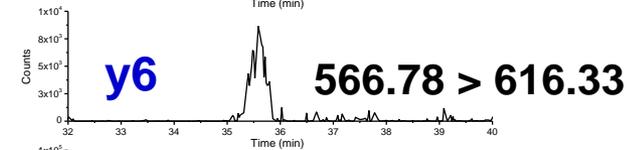
TGSAEAVPSEGK > SEGK



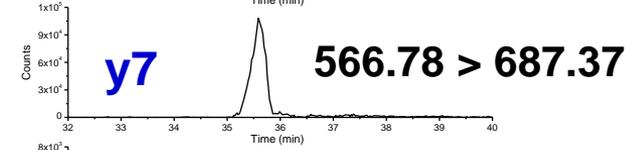
TGSAEAVPSEGK > PSEGK



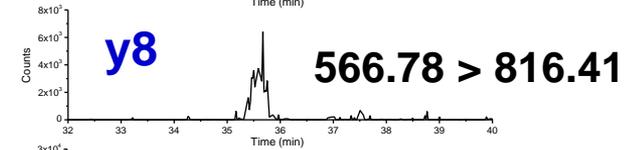
TGSAEAVPSEGK > VPSEGK



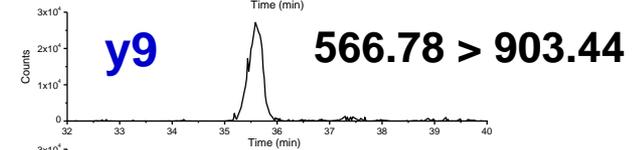
TGSAEAVPSEGK > AVPSEGK



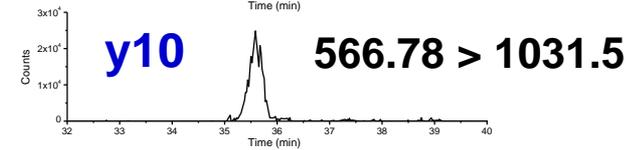
TGSAEAVPSEGK > EAVPSEGK



TGSAEAVPSEGK > AEAVPSEGK

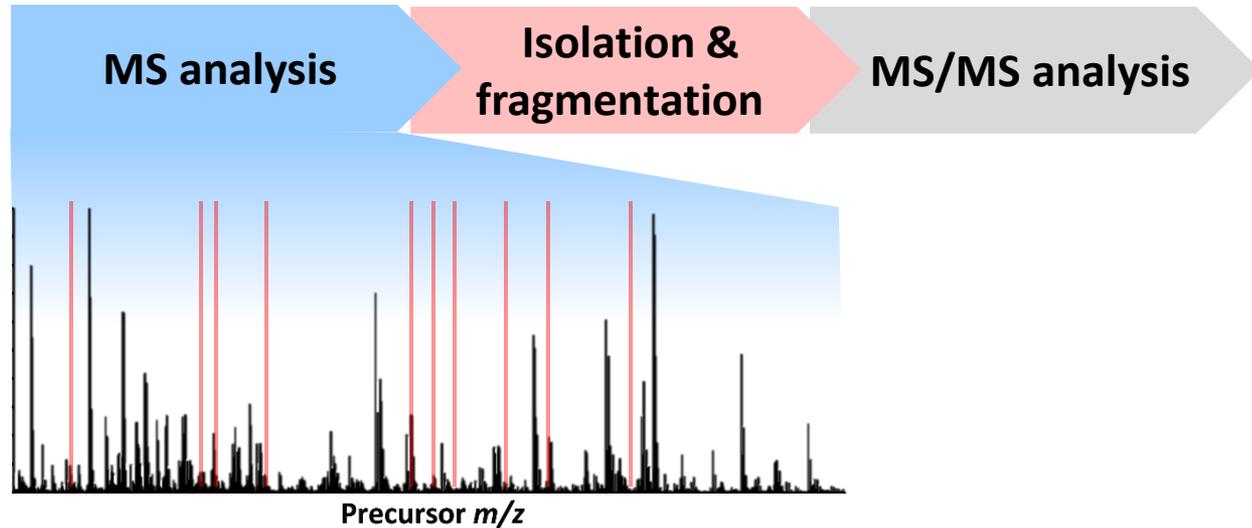


TGSAEAVPSEGK > SAEAVPSEGK



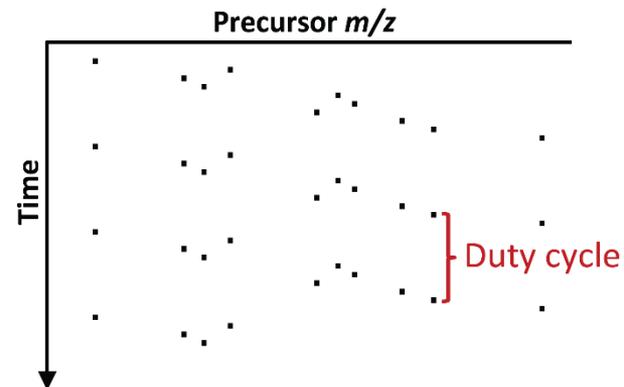
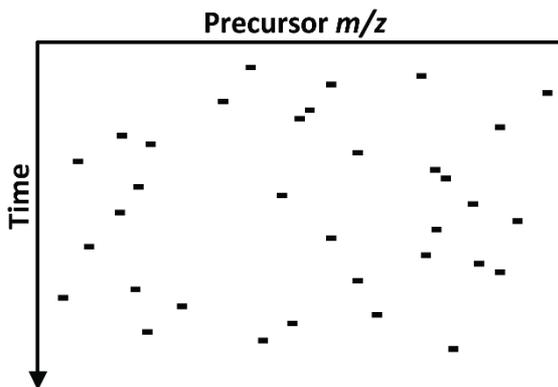
Acquisition methods in shotgun LC-MS/MS

Liquid chromatography

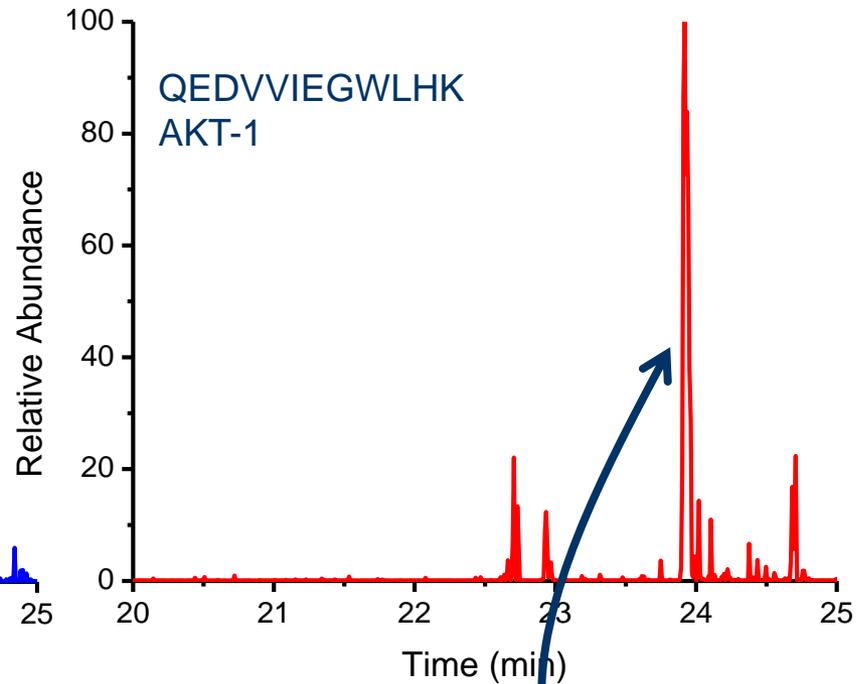
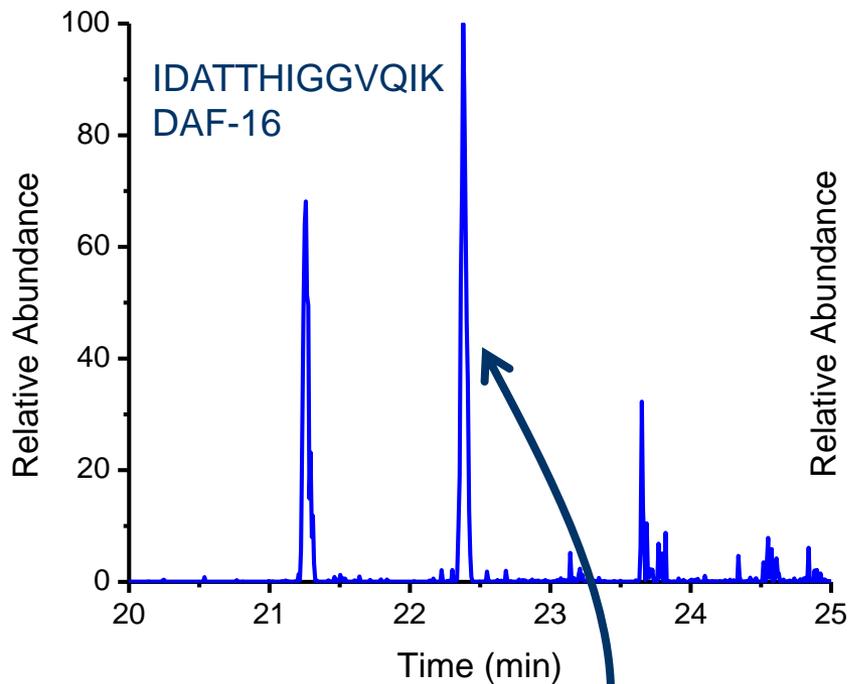


Data Dependent Acquisition (DDA)

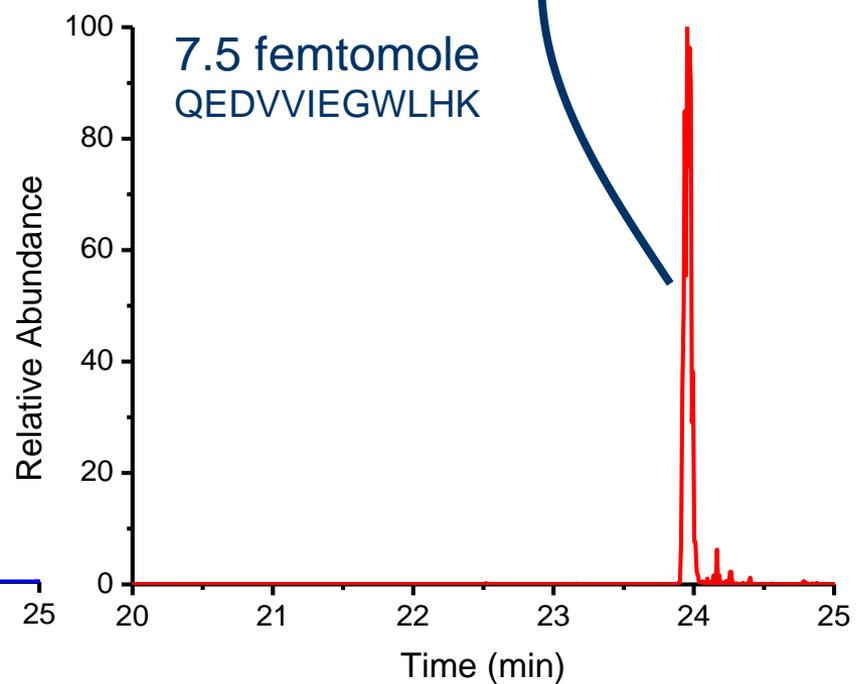
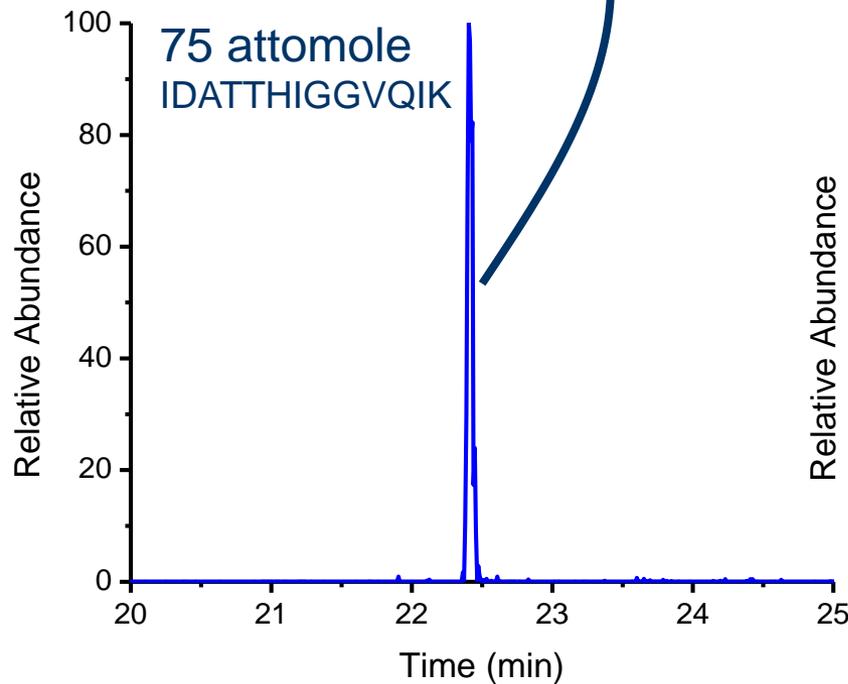
Selected/Parallel Reaction Monitoring (S/PRM)



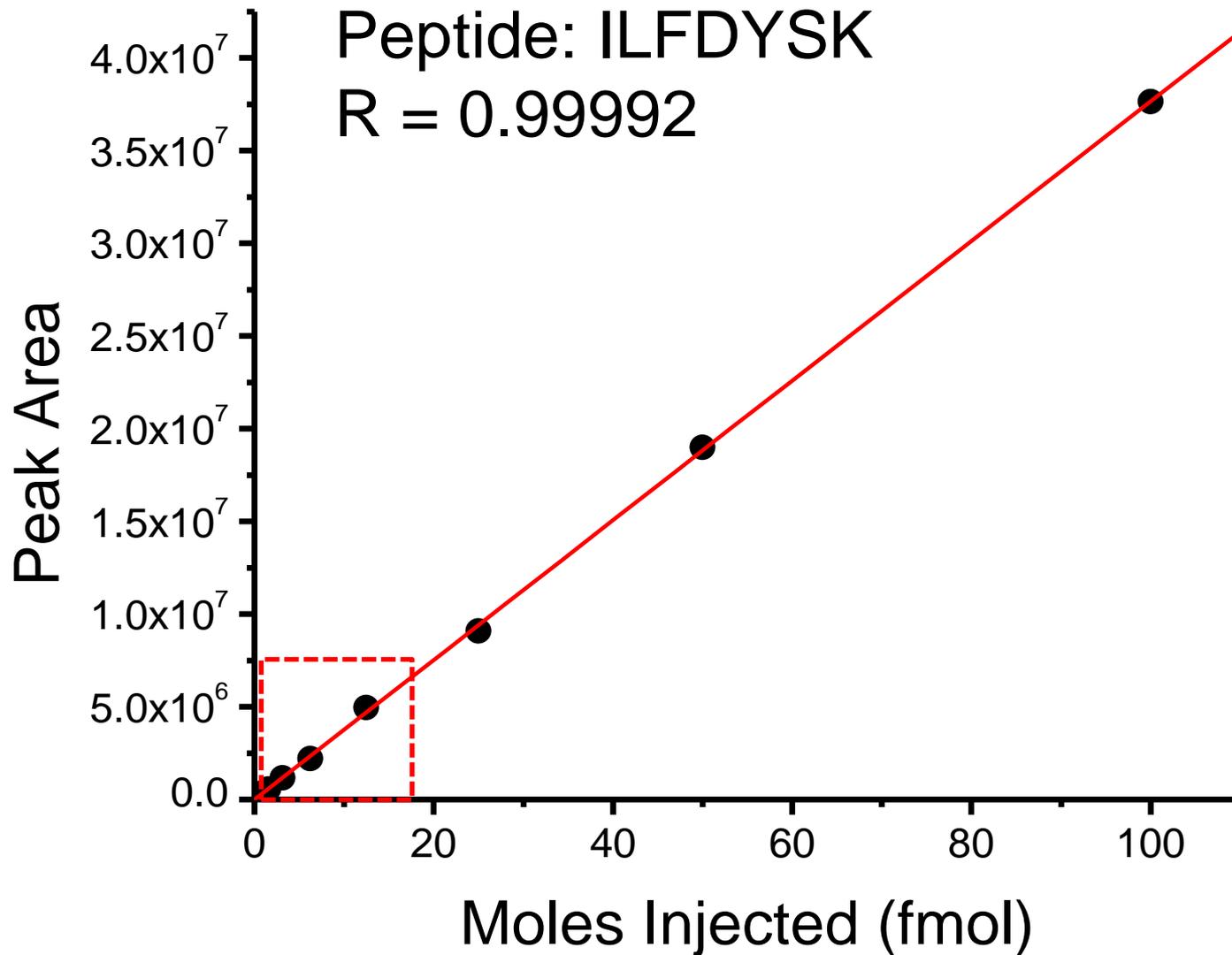
Worm Homogenate 25 μ g



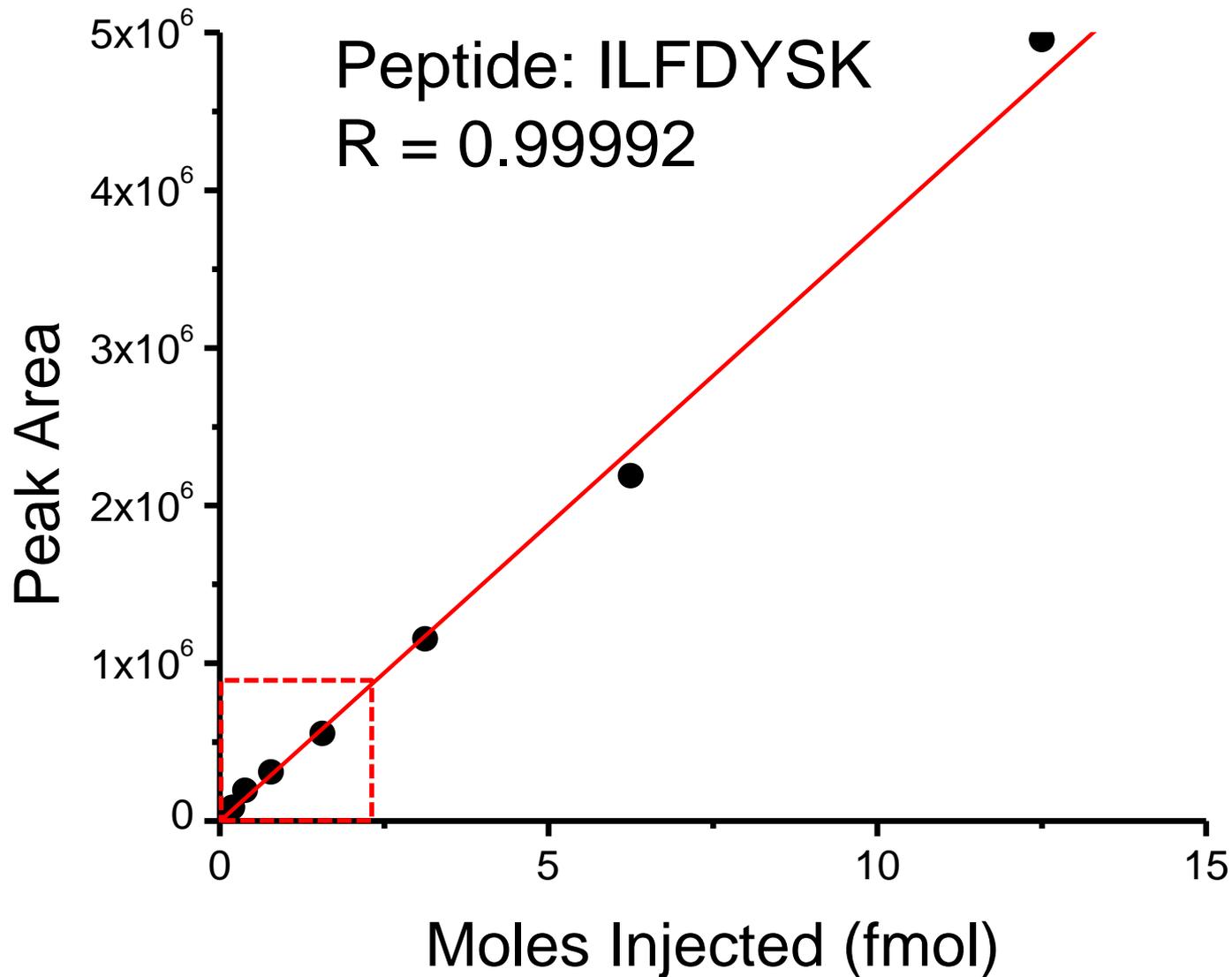
Synthetic Peptide Standards



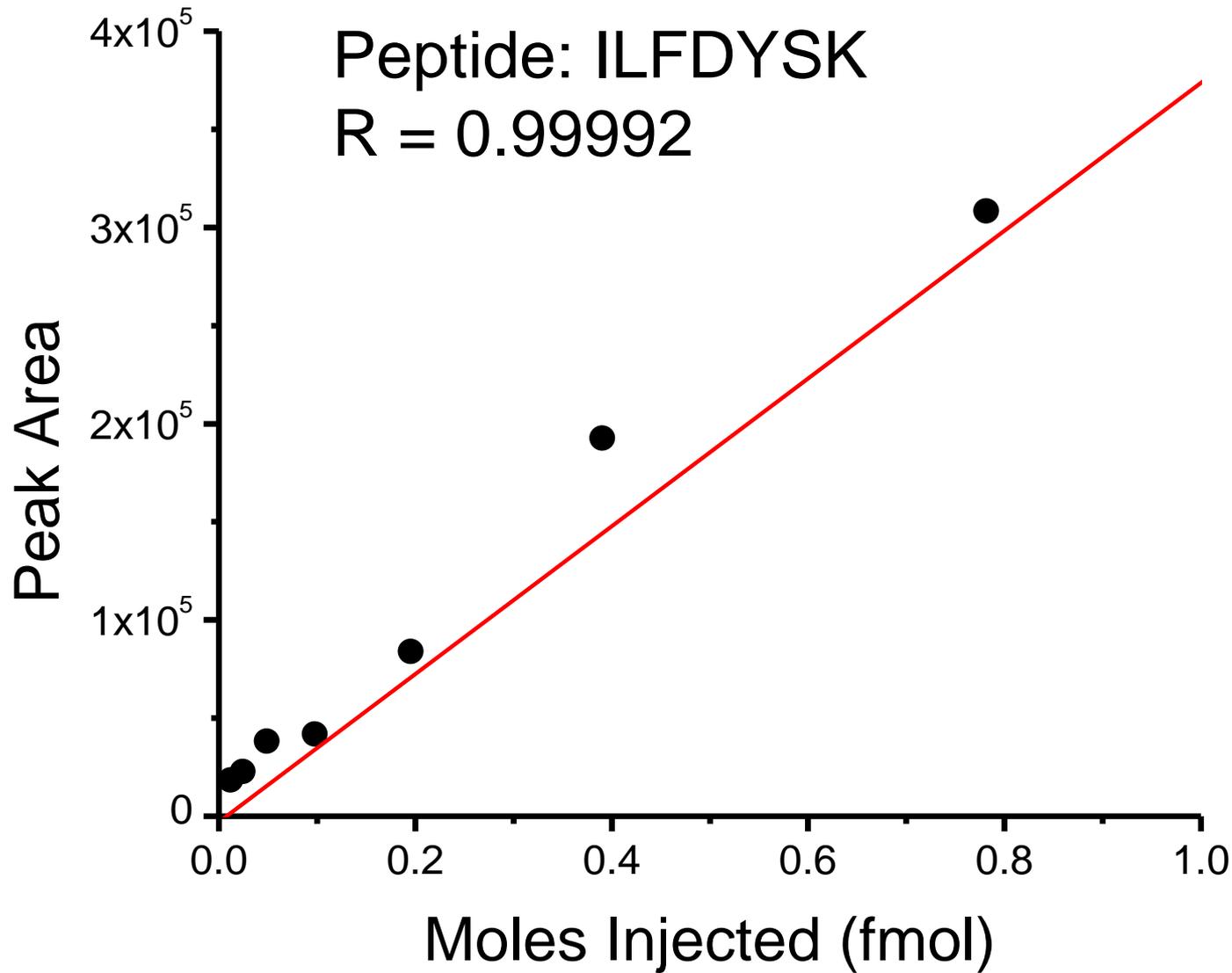
Dilution of the Yeast Protein PGI₁ in Undepleted Human Plasma



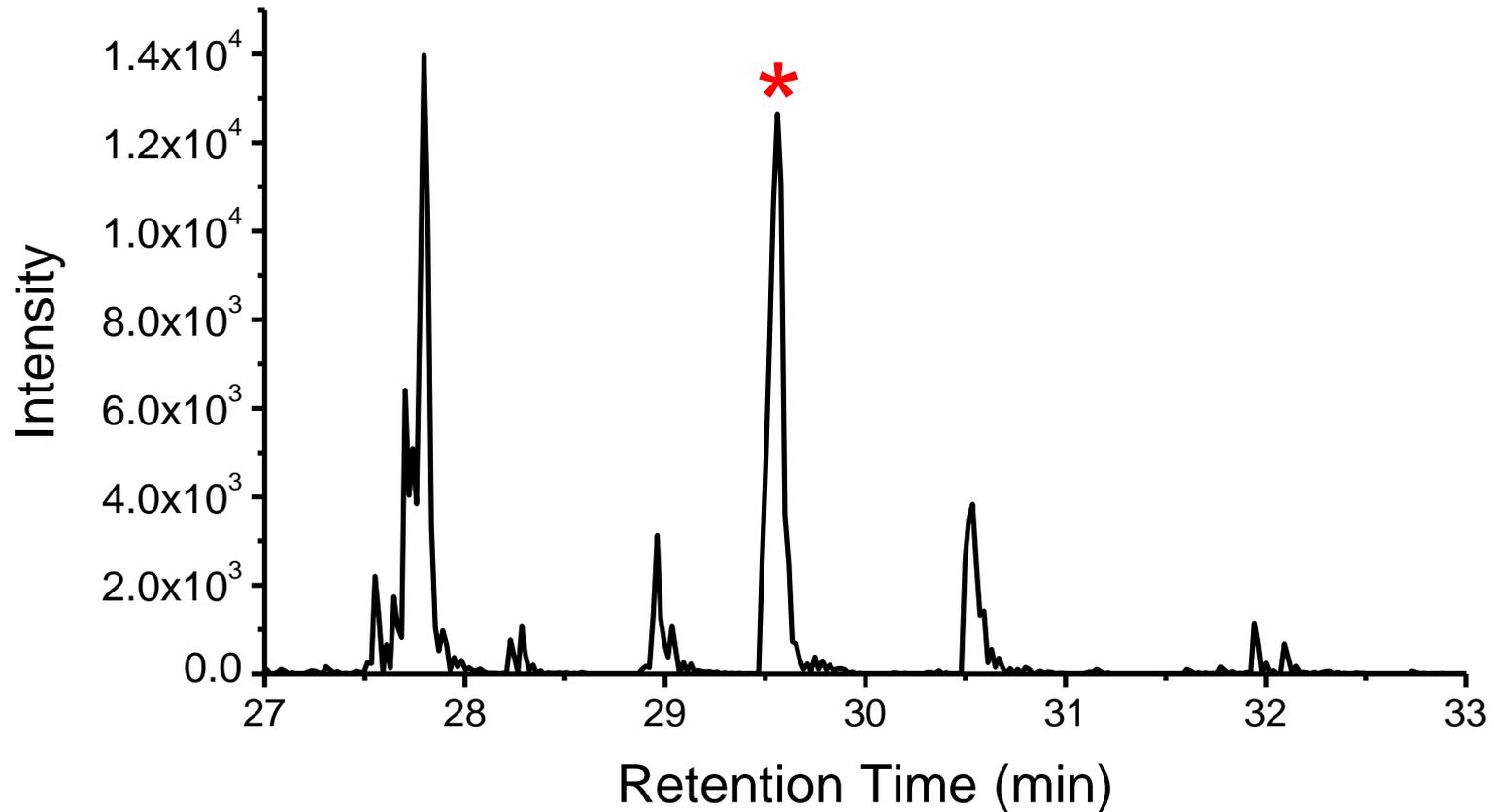
Dilution of the Yeast Protein PGI₁ in Undepleted Human Plasma



Dilution of the Yeast Protein PGI₁ in Undepleted Human Plasma



12 Attomoles of ILFDYSK on Column in 10 μ g of Unfractionated Human Plasma



SRM of Ceruplasmin - ETFTYEWTVPK

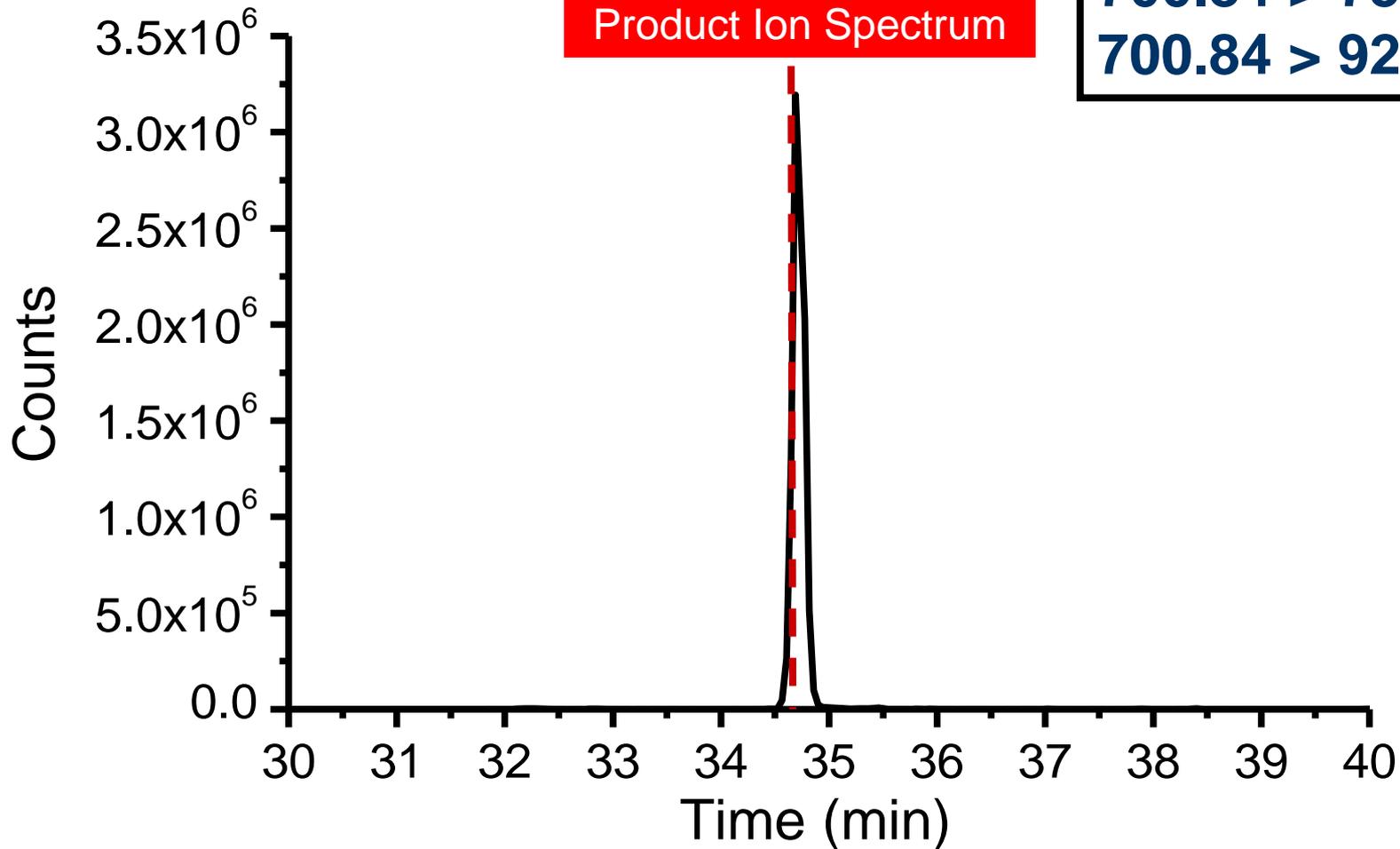
SRM Transitions

700.84 > 244.17

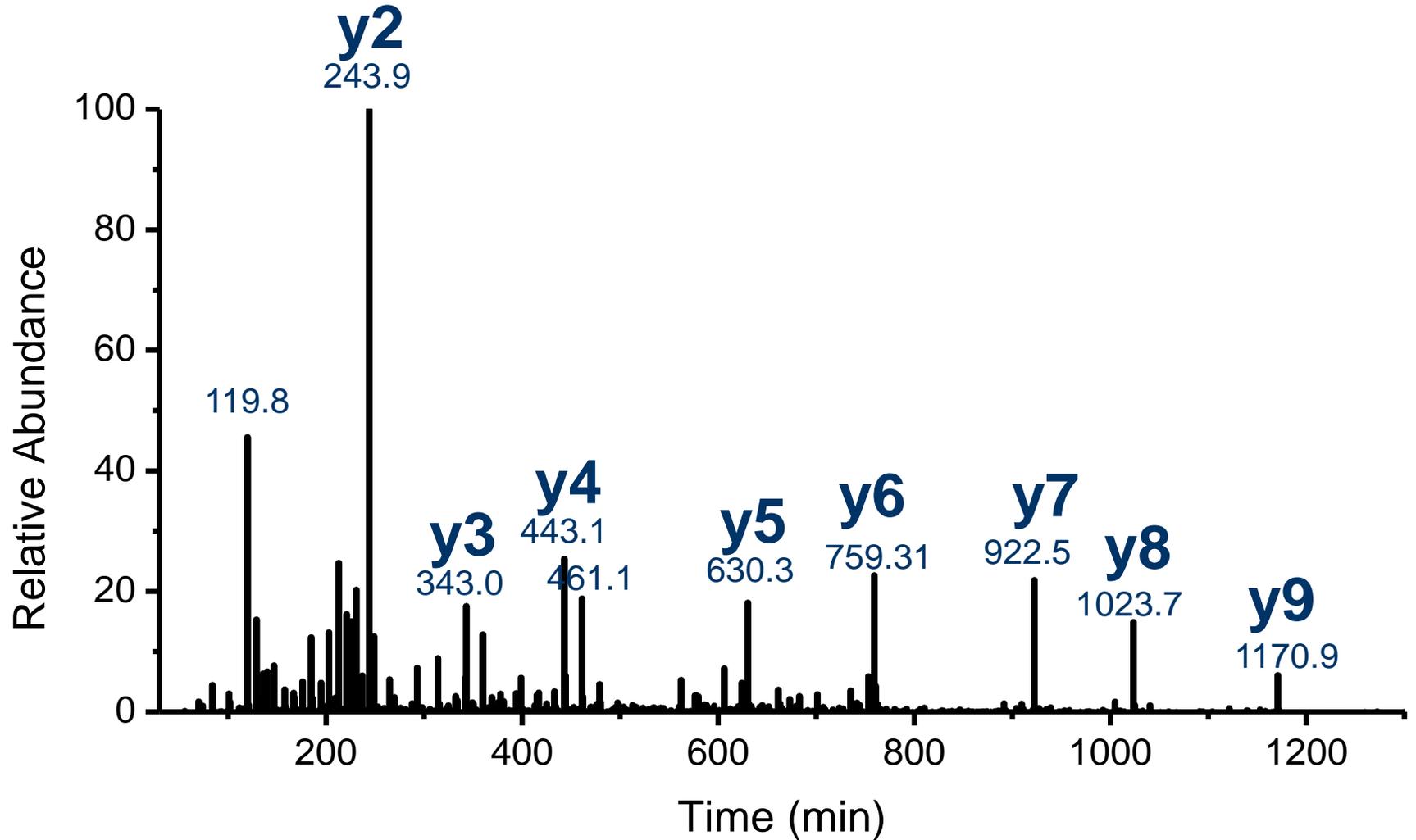
700.84 > 759.4

700.84 > 922.47

SRM Triggered
Product Ion Spectrum

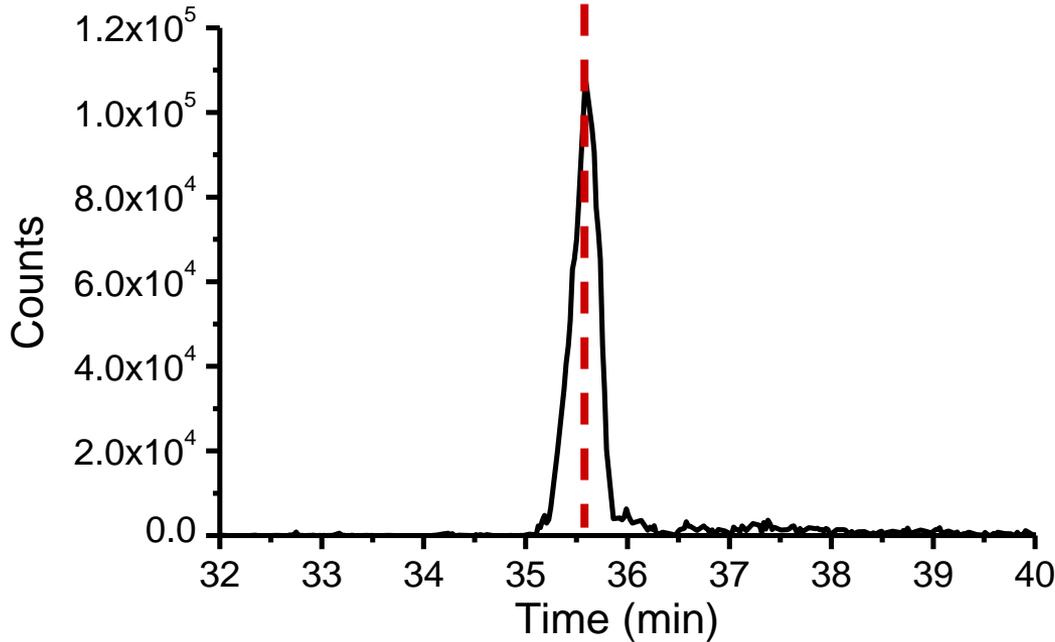


Ceruplasmin – ETFTYEWTVPK

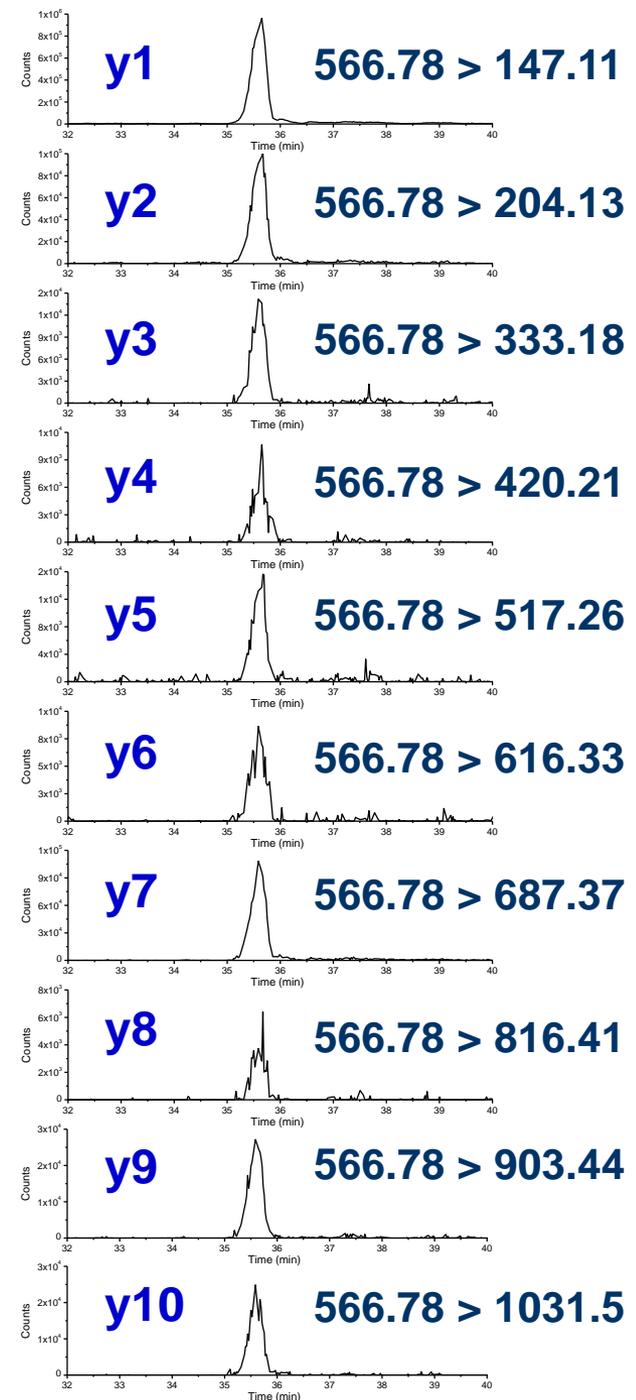


Chromogranin A TGASEAVPSEGK

SRM Triggered
Product Ion Spectrum

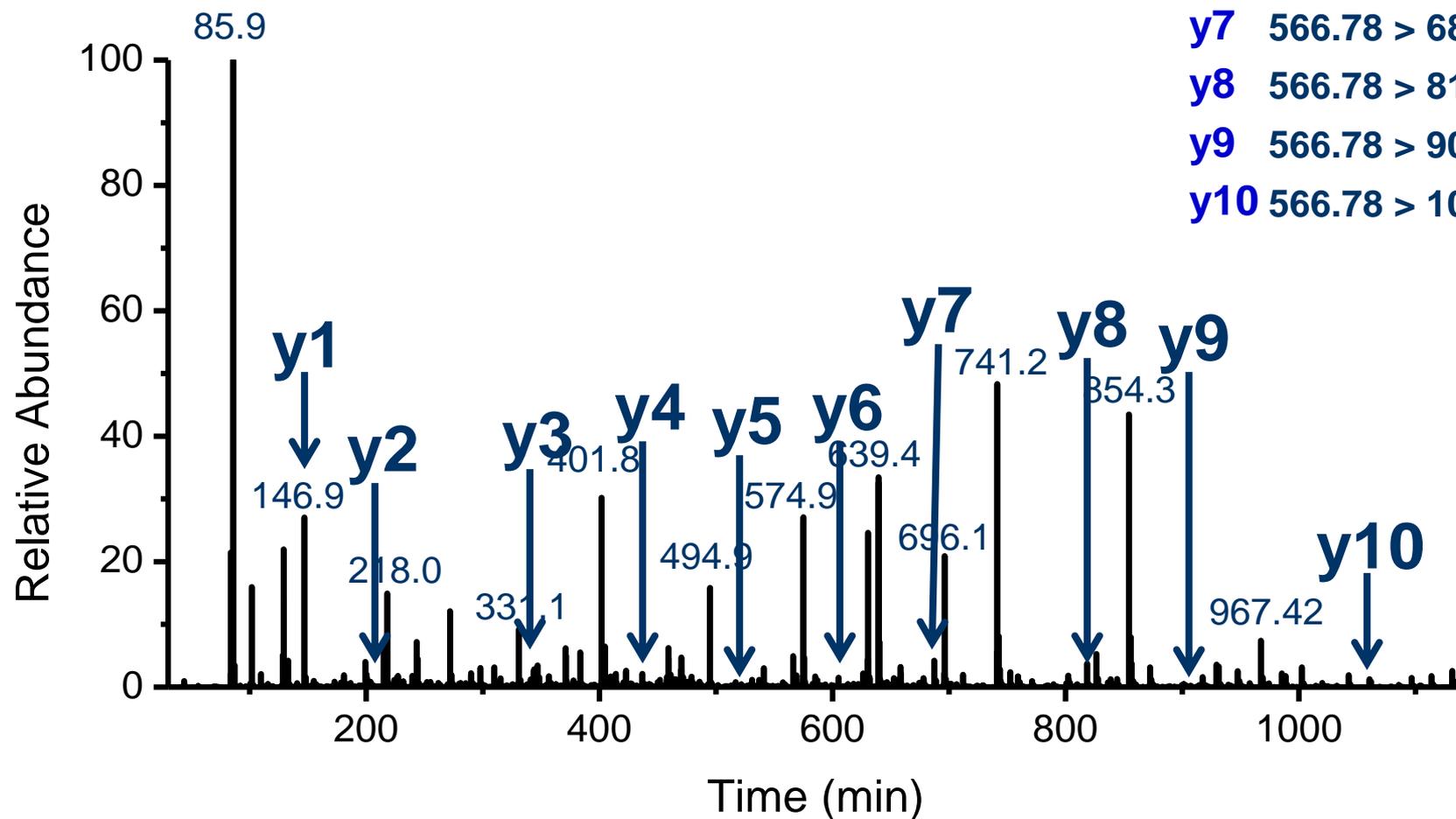


Chromogranin A is ~10 ng/mL in plasma
Anderson et al. J. Physiology 2005



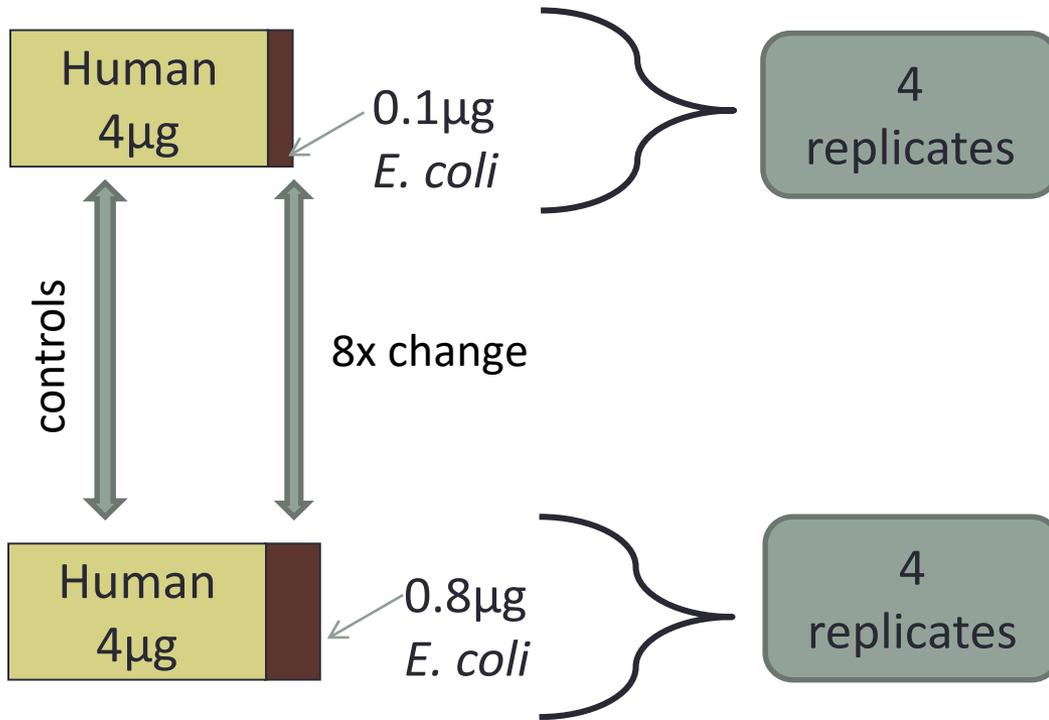
Chromogranin A TGASEAVPSEGK

- y1** 566.78 > 147.11
- y2** 566.78 > 204.13
- y3** 566.78 > 333.18
- y4** 566.78 > 420.21
- y5** 566.78 > 517.26
- y6** 566.78 > 616.33
- y7** 566.78 > 687.37
- y8** 566.78 > 816.41
- y9** 566.78 > 903.44
- y10** 566.78 > 1031.5

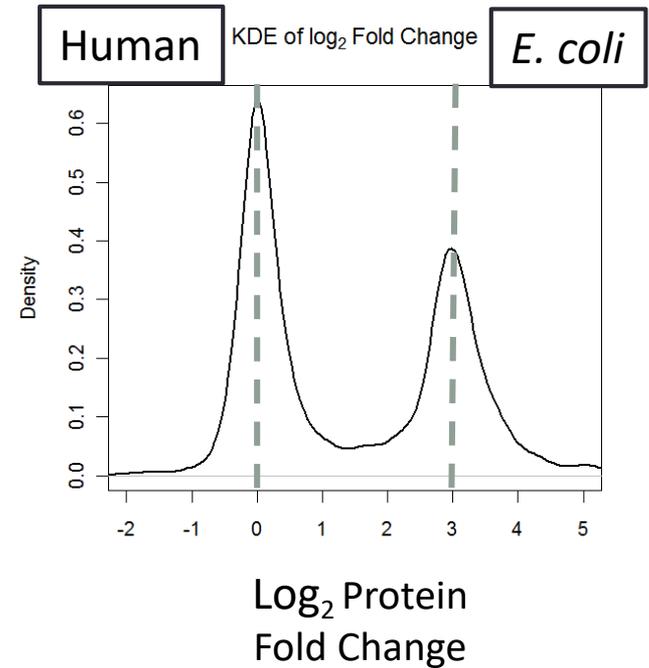


Should we be putting in all this effort toward measuring reliable peptide peak areas?

Spike-in Validation Experiment



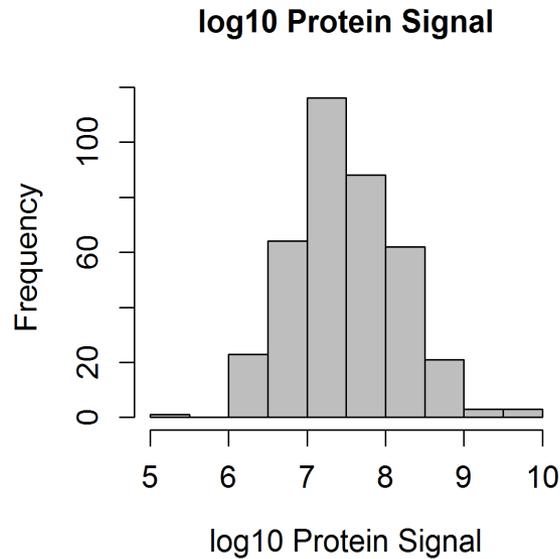
Relative Quantitation



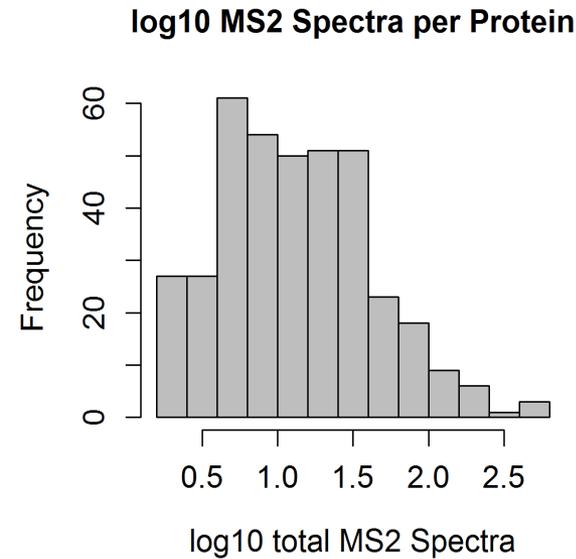
Protein ID Summary

#Proteins Detected	# <i>E. coli</i> Proteins	#Human Proteins	# <i>E. coli</i> peptides	#Human peptides	Spectra / PSMs at q-value < 0.01
2073	507	1566	1633	2346	172241; 29163 (16.9%)

MS₁ Filtering compared with Spectral Counting

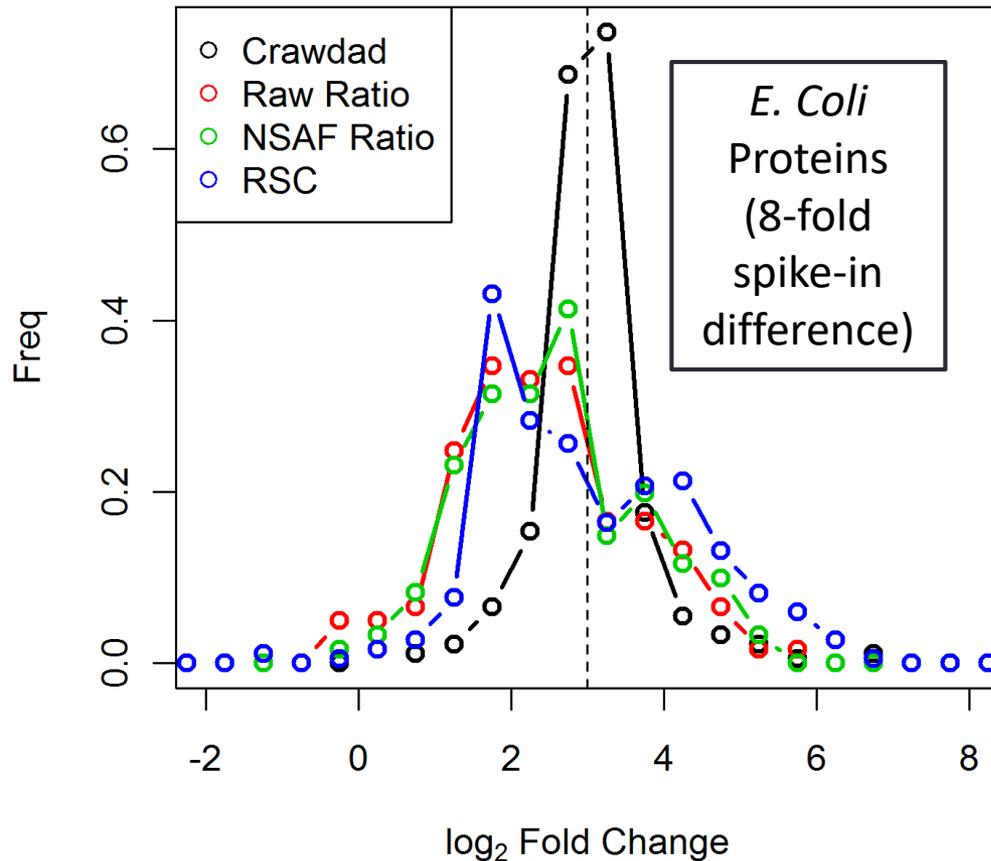


4 orders of magnitude
of dynamic range



2.5 orders of magnitude
of dynamic range

MS₁ Filtering and Spectral Counting Fold-Change Distributions



Limited to proteins where SPC for both
1x, 8x sets are > 0

Crawdad – AUC quantitation

Raw Ratio: n_1 / n_2

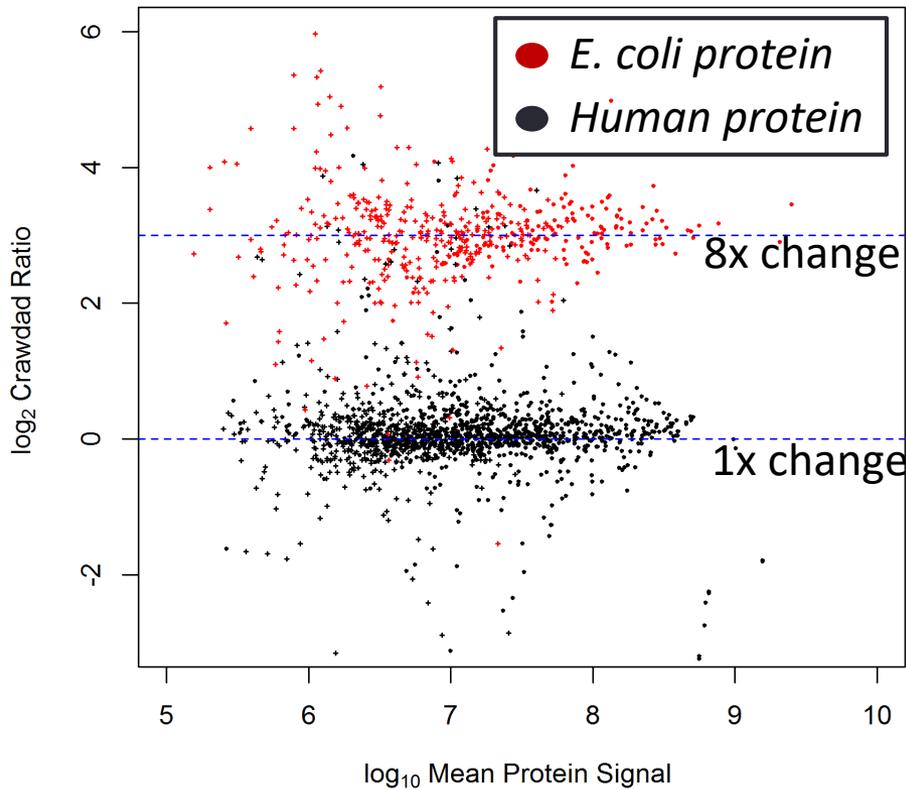
NSAF Ratio: ratio of SPC normalized by
protein length

$$s_c = \log_2 \left(\frac{n_2 + f}{n_1 + f} \right) + \log_2 \left(\frac{t_1 - n_1 + f}{t_2 - n_2 + f} \right)$$

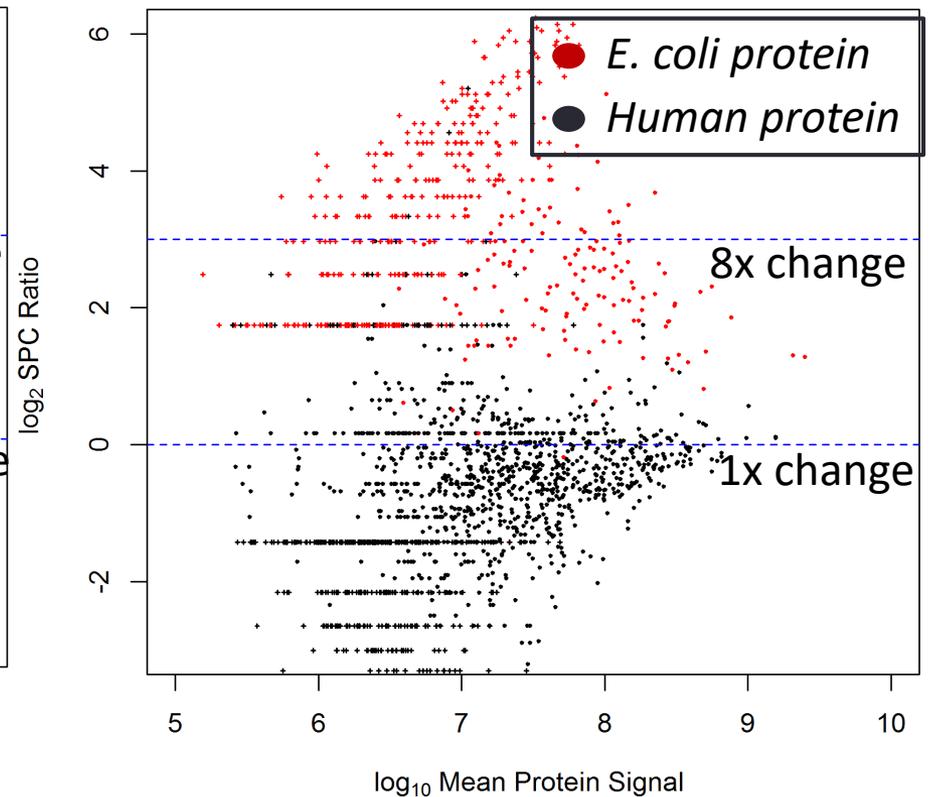
$f = 0.5$

Measured Protein Fold-Change Ratios: MS₁ Filtering vs. Spectral Counting

Crowdad Ratio vs. Protein Signal



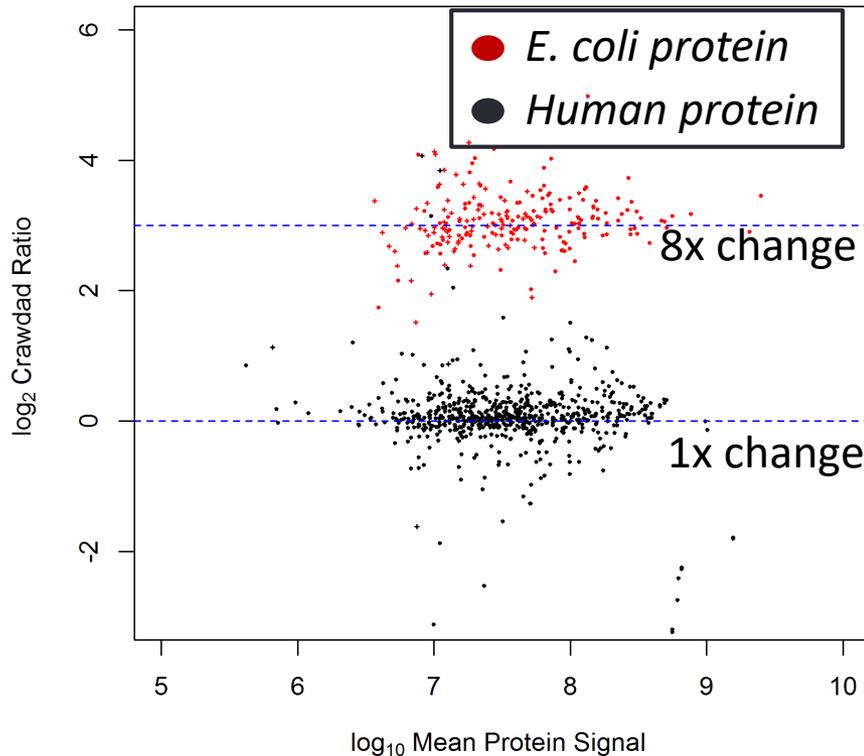
SPC Ratio vs. Protein Signal



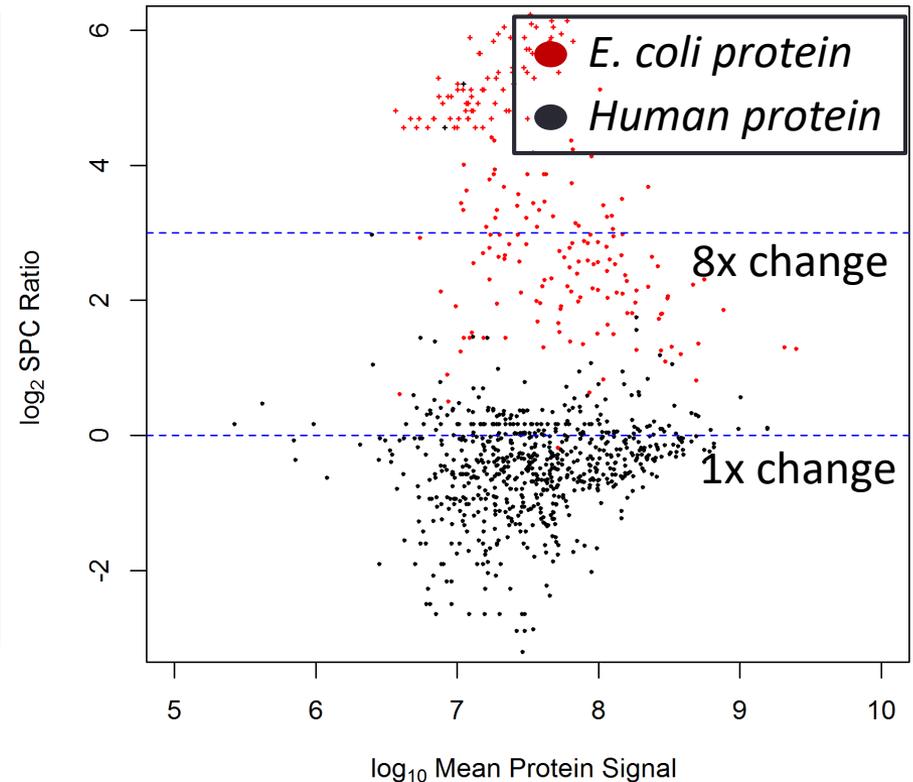
Measured Protein Fold-Change Ratios: MS₁ Filtering vs. Spectral Counting

Limited to Proteins with ≥ 10 MS/MS Spectra

Crawdad Ratio vs. Protein Signal



SPC Ratio vs. Protein Signal



Summary

- Immunoassays are an imperfect solution to the protein quantitation problem.
- Add the internal standard as early as possible to account for sample preparation errors
- Systematic errors do occur and need to be corrected.
- High sequence coverage improves the comprehensiveness of shotgun proteomics data
- Know the isotope enrichment of the internal standard.
 - Can be measured by a number of different techniques
- ^{13}C labeled internal standards aren't perfect and require special considerations to get accurate isotopomer ratios
- Think about how you are going to calculate the significance before you do the study. Get help early on if you are not sure.
- Is your sample size and experimental design going to answer your questions?
- Don't tolerate experiments with an N of 1!

Schedule

Time	Day 1	Day 2
9:00-9:30	Intro to Protein Quantification (Mike MacCoss)	Intro to Data Independent Acquisition (Mike MacCoss)
9:30-10:00		Principles of Study Design: Possibly examples of failed studies? (Olga Vitek)
10:00-10:30		
10:30-11:00	Introduction to Skyline (Brendan MacLean)	Case Study 3b: Using Sentinel Profiles for Drug Characterization: Clustering and connectivity (Jake Jaffe)
11:00-11:30	Case Study 1a: Targeted Method Refinement (Brendan MacLean)	
11:30-12:00		
12:00-12:30		Lunch
12:30-13:00	Case Study 1b: Group Comparisons in Skyline (Brendan MacLean)	Intro to MSstats (Olga Vitek)
13:00-13:30		
13:30-14:00		Case Study 2: MS1 Quantification in Skyline (Brendan MacLean)
14:00-14:30		
14:30-15:00	Case Study 3a: Intro to Sentinel Assay Concept, Development of P100 Discovery Data Set (Jake Jaffe)	System Suitability and Quality Control. How we assess this, AutoQC/Panorama/MSStatsQC (Lindsay/Eralp)
15:00-15:30		
15:30-16:00		