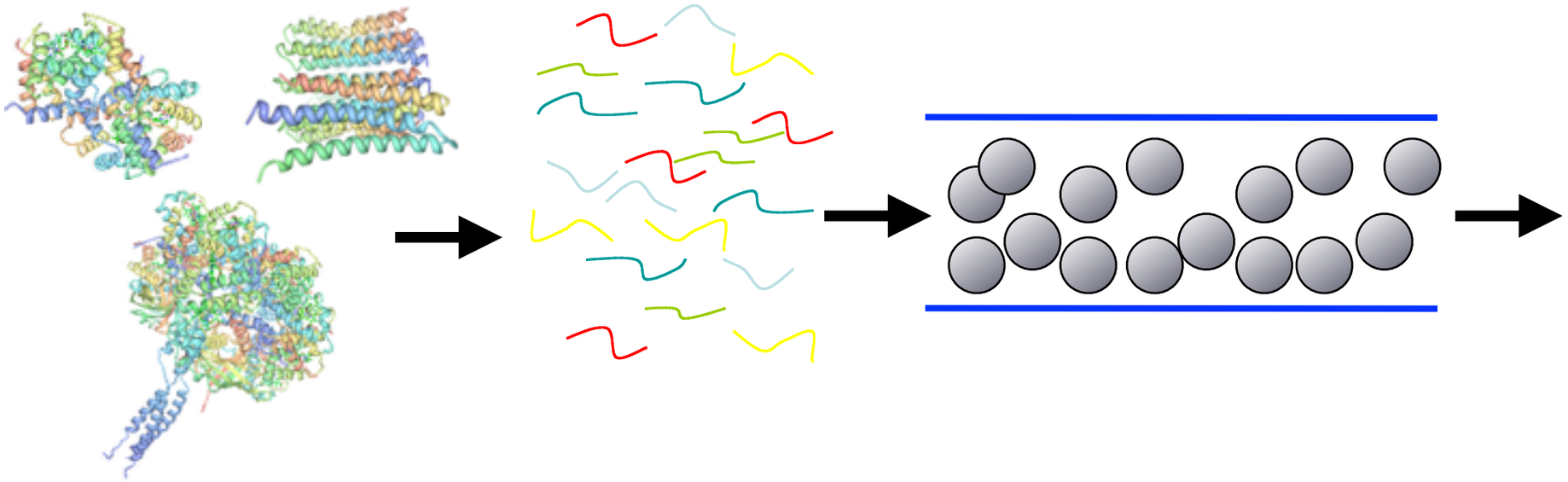
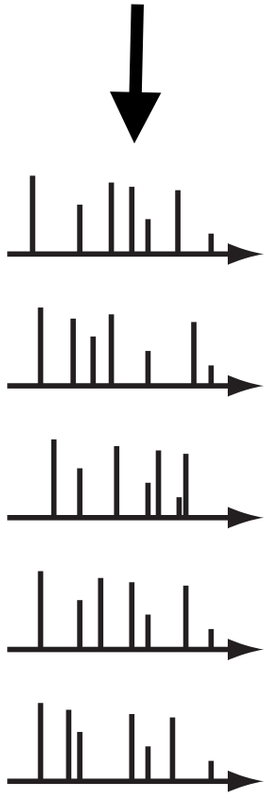
Peptide Identification (Repetition)

Shotgun proteomics





GLB4_LUMTE

COX1_LUMTE

ATP6_LUMTE

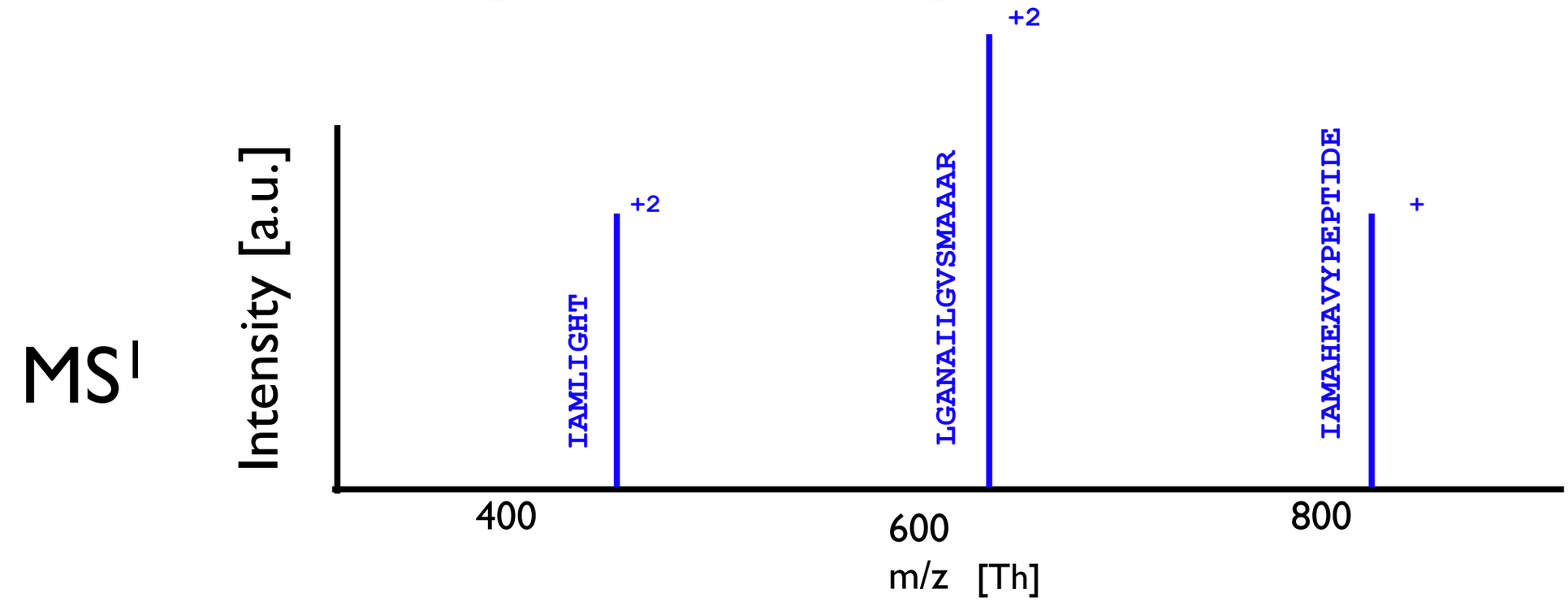
IAMAPEPTIDER

MGREATMATCHK

ATRYIDENTIFYMEK

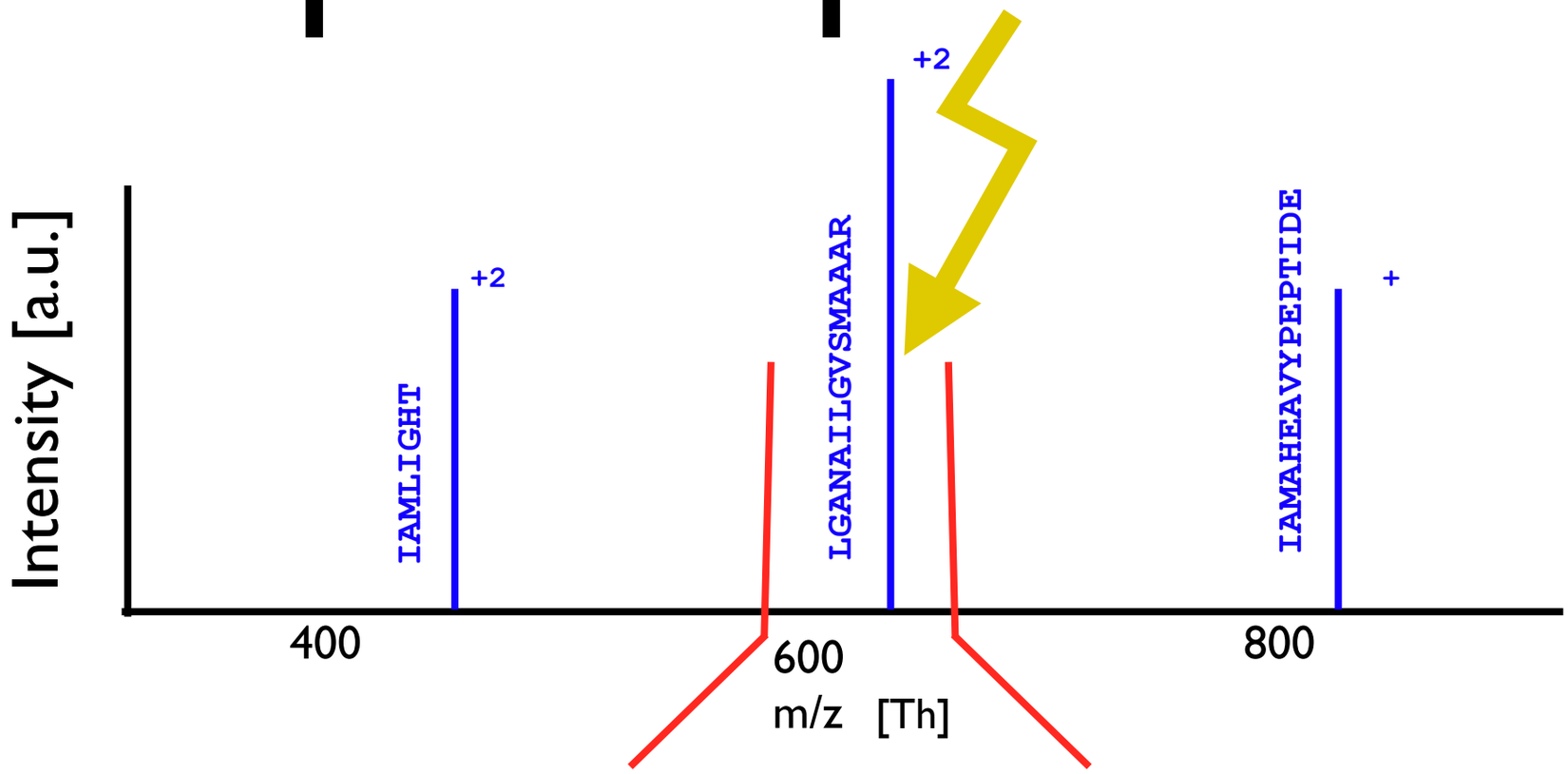
LMAKEMYDAYR

Peptide spectra

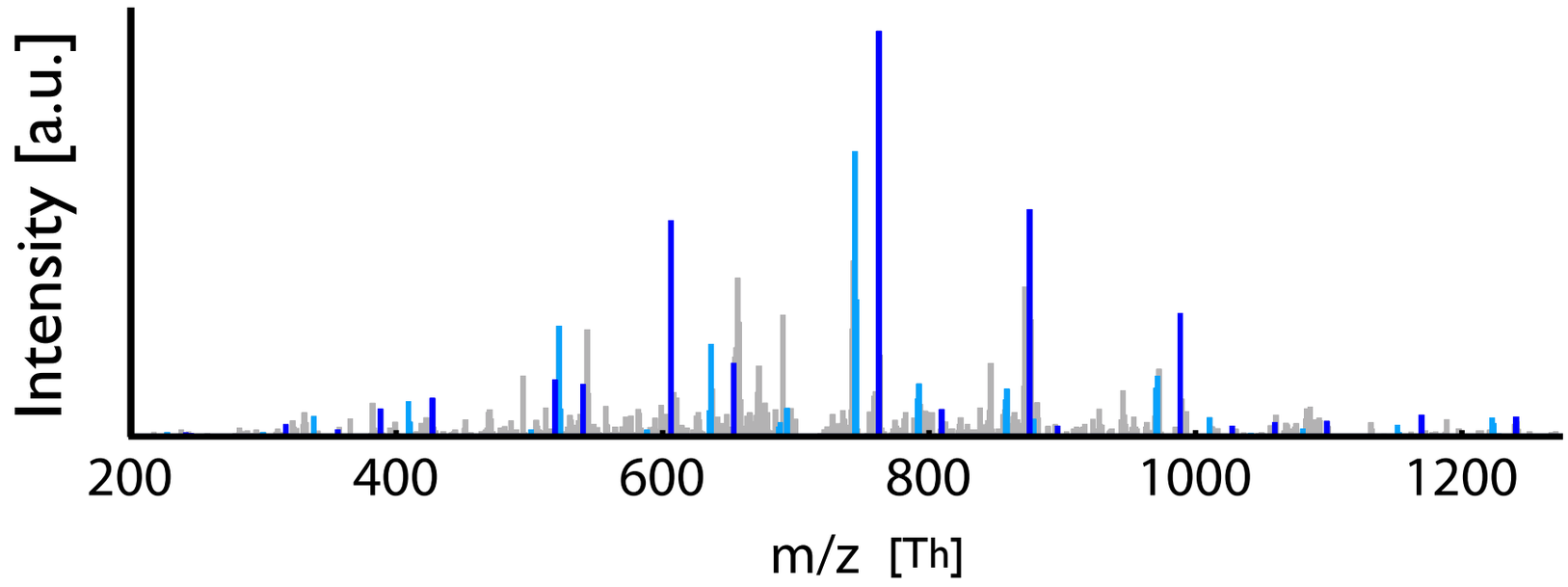


Peptide spectra

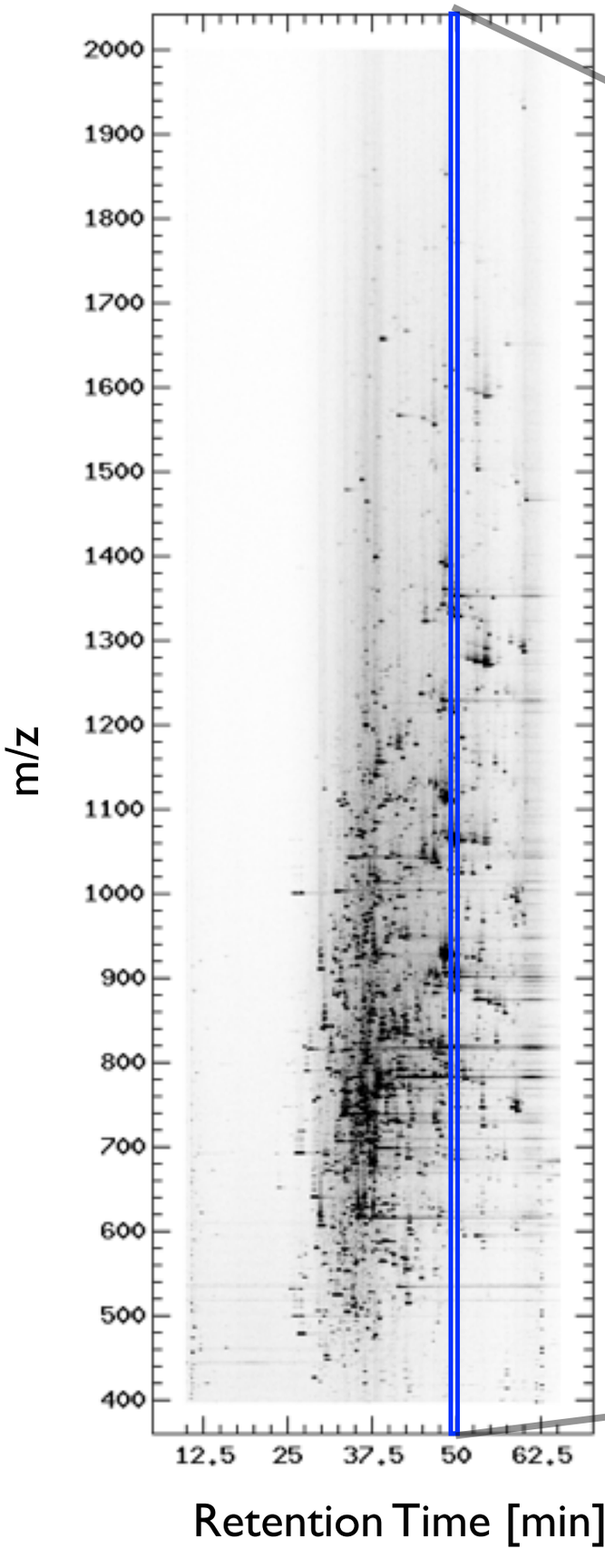
MS¹

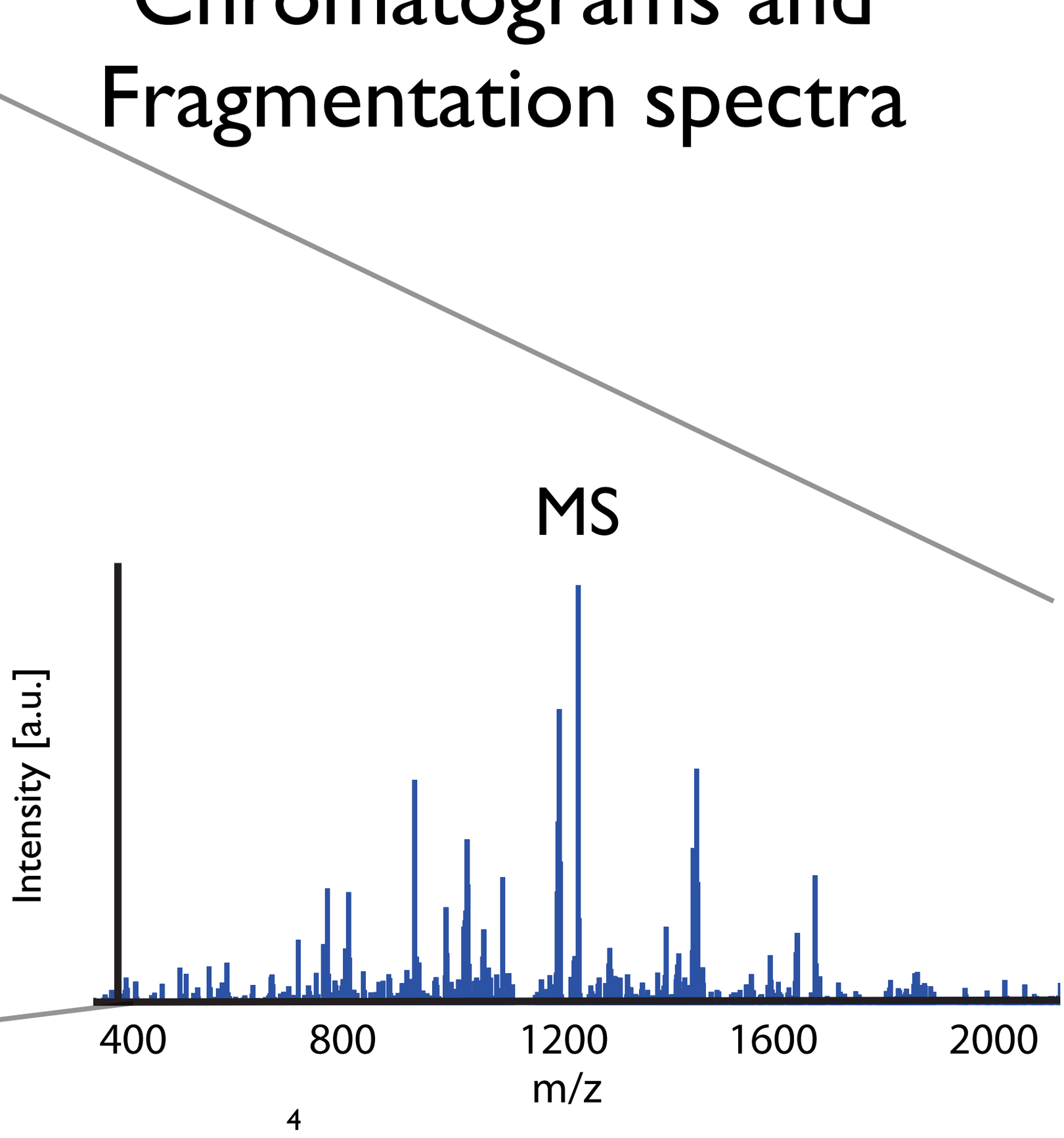


MS²

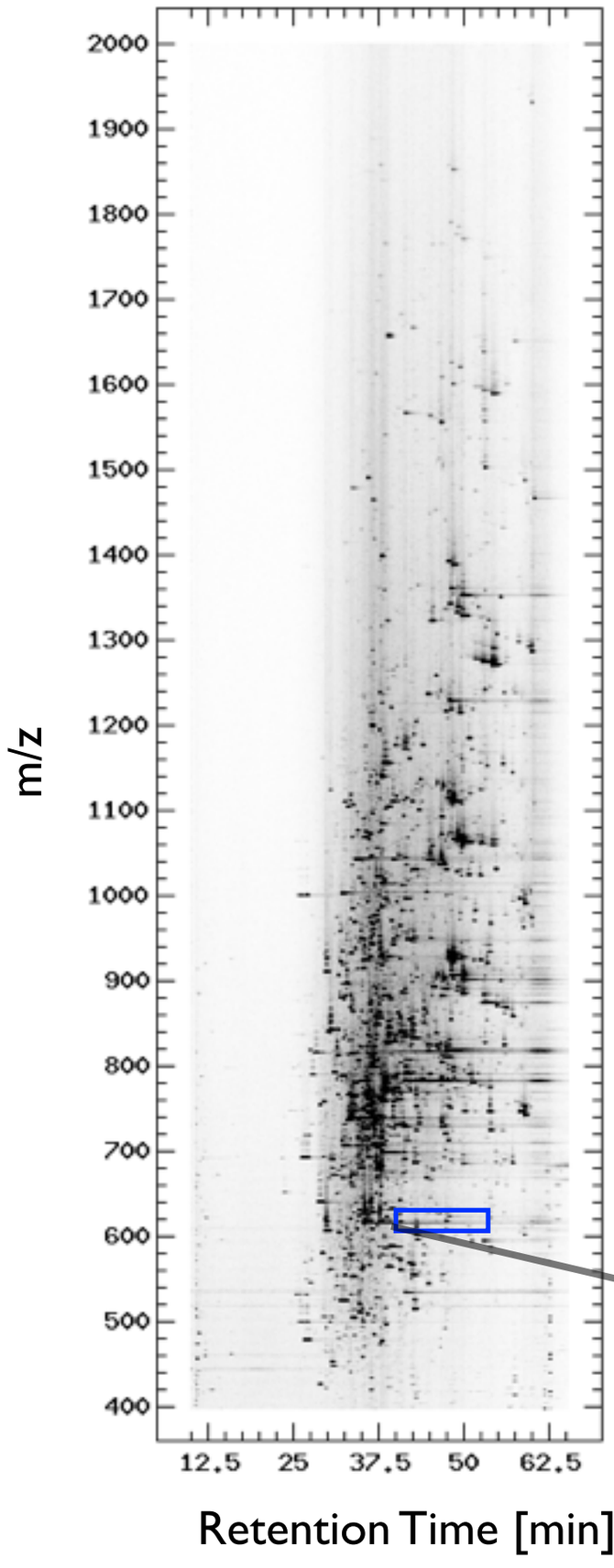


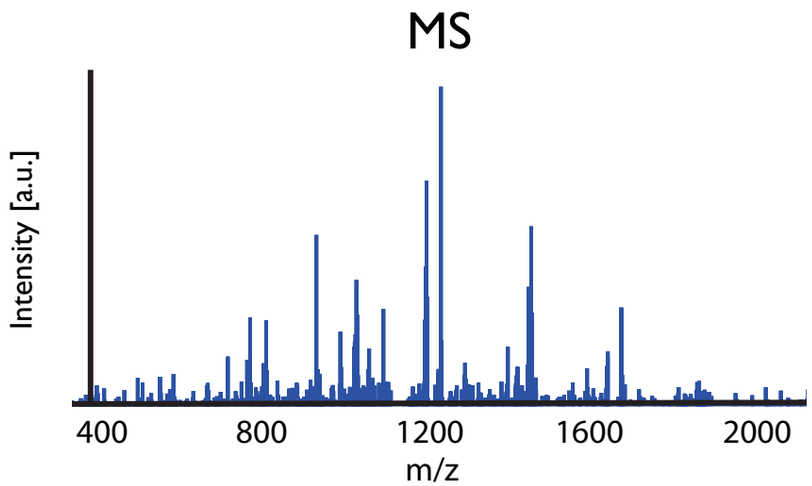
Chromatograms and Fragmentation spectra

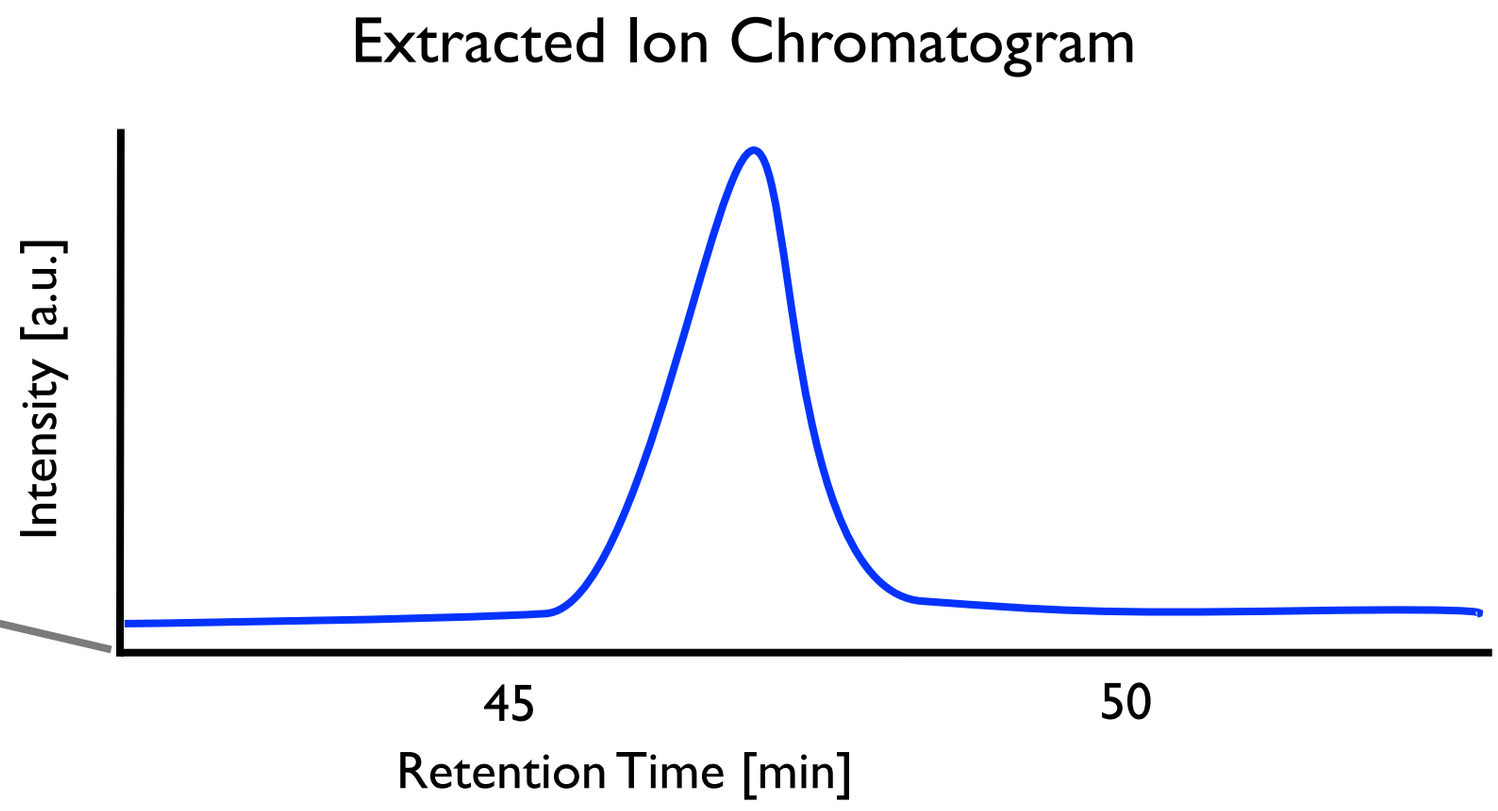




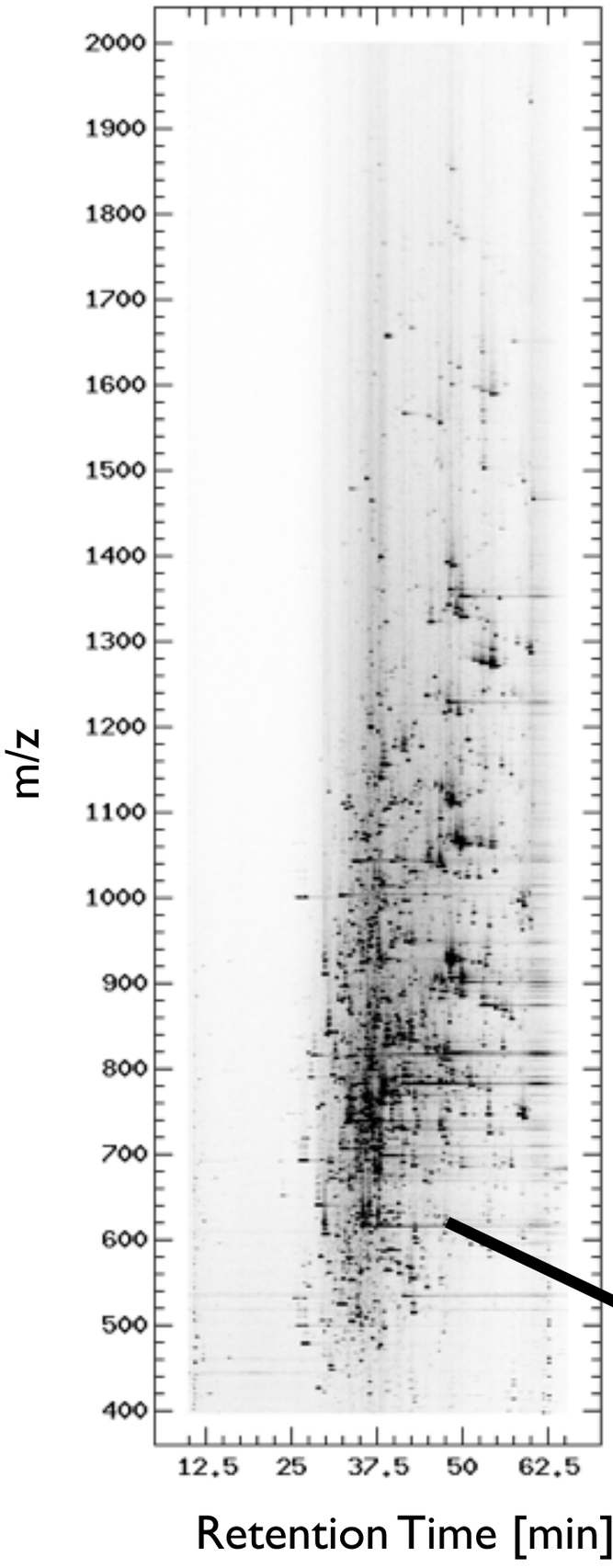
Chromatograms and Fragmentation spectra

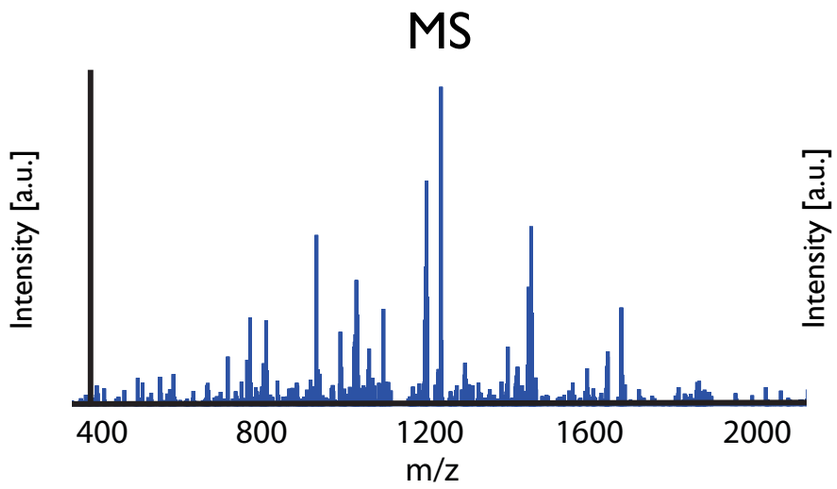


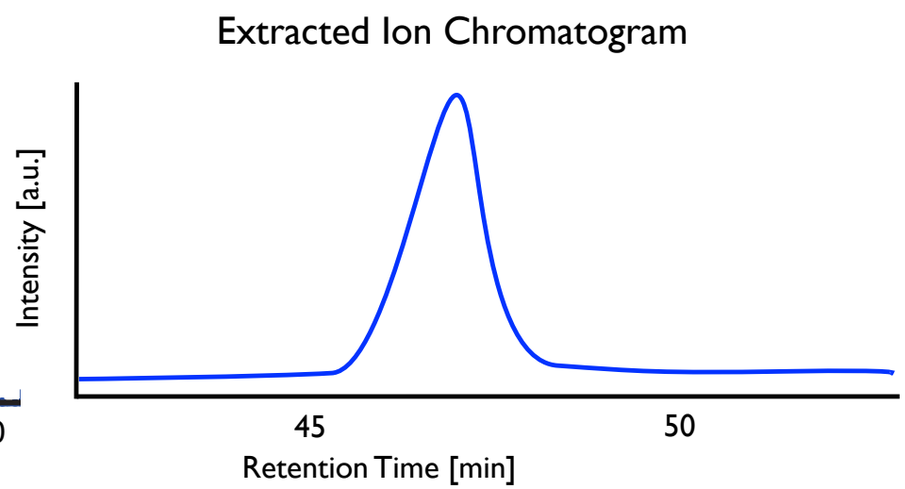


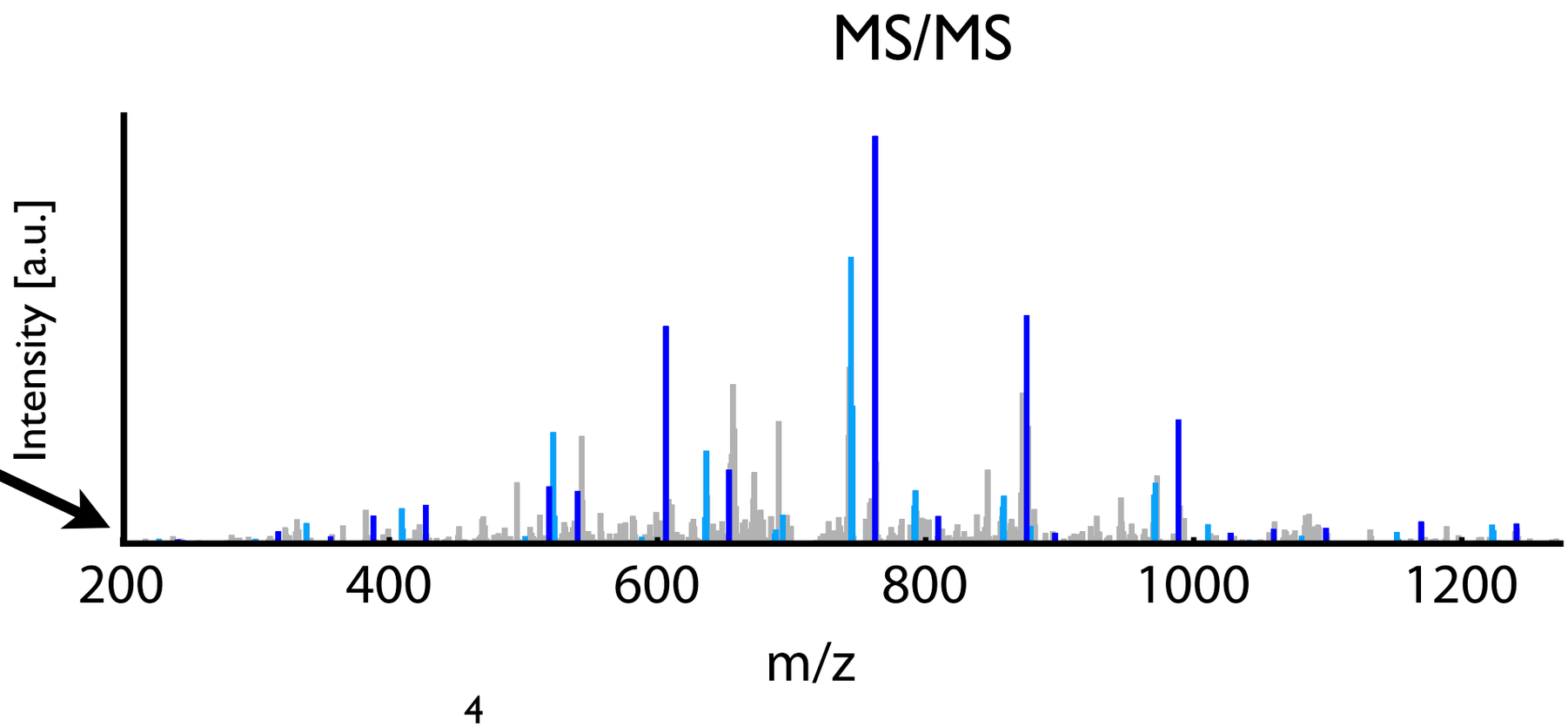


Chromatograms and Fragmentation spectra

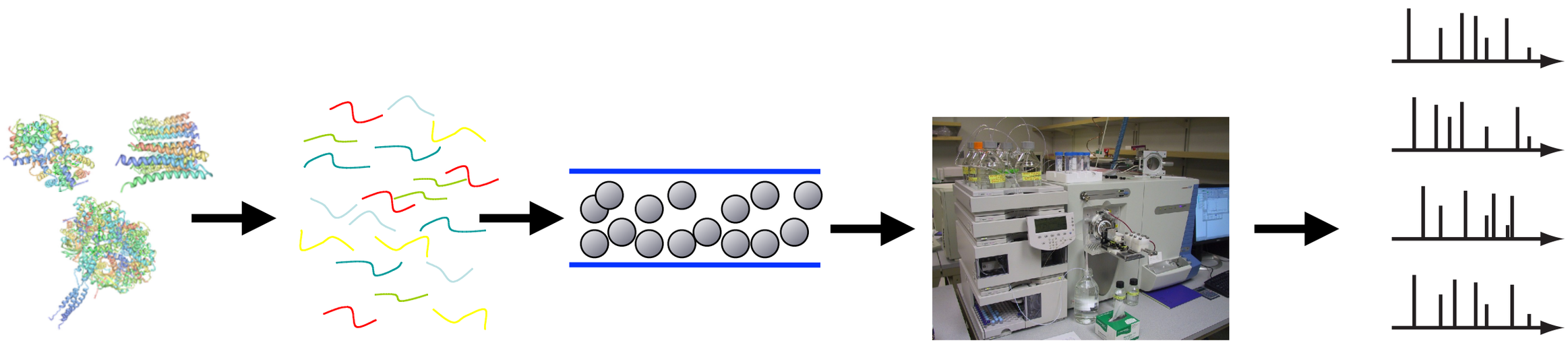




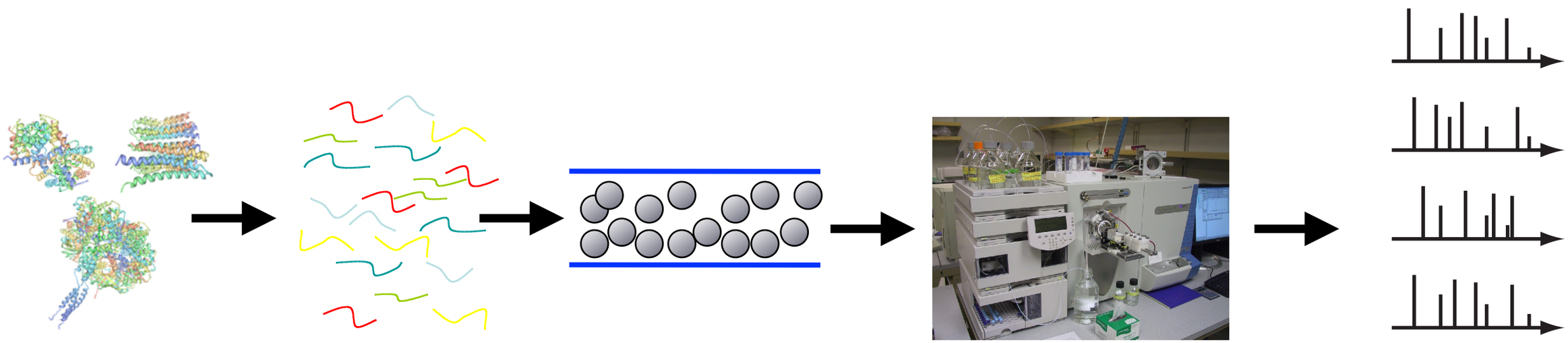




Peptide Identification



Peptide Identification

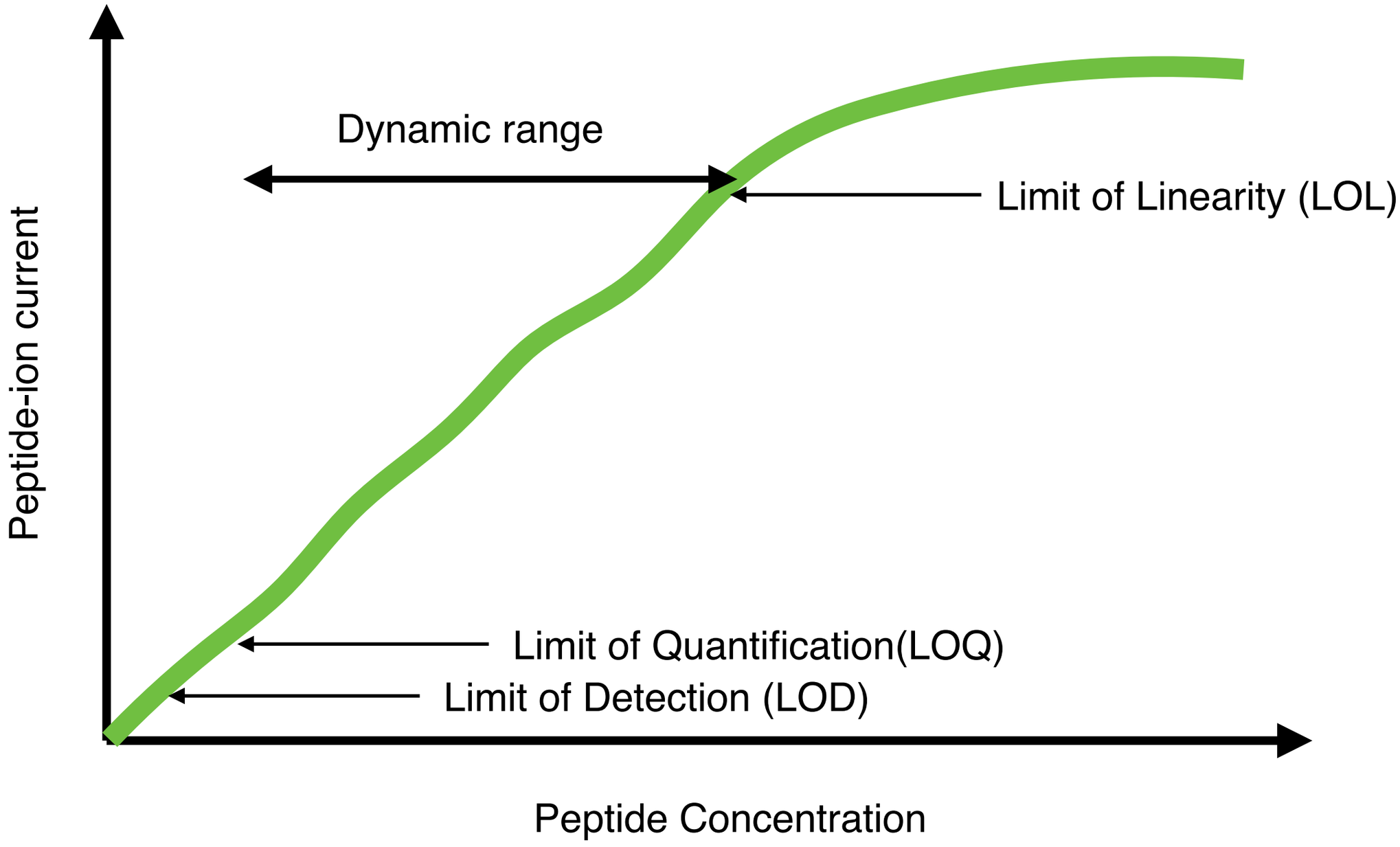


```
>REG3A_HUMAN
MLPPMALPSVSWMLLSCLMLLKQVQGEEPQRELPSARI
RCPKGSKAYGSHCYALFLSPKSWTDADLACQKRPSGNL
VSVLSGAEGSFVSSLVKSIGNSYSYVWIGKHDPTQGTE
PNGRGWEWSSSDVMNYFAWERNPSTISSPGHCASLSRS
TAFLRWKDYNCNVRLPYVCKFTD
>NPC2_HUMAN
MRFLAATFLLLALSTAAQAEPVQFKDCGSVDGVIKEVN
VSPCPTQPCQLSKGQSYSVNVTFTSNIQSKSSKAVVHG
ILMGVPVPFPIPEPDGCKSGINCPIQKDKTYSYLNKLP
VKSEYPSIKLVVEWQLQDDKNQSLFCWEIPVQIVSHL
>ATP6_HUMAN
MNENLFASFIAPTILGLPAAVLIILFPPLLIPTSKYLI
NNRLITTQQWLIKLTSKQMMTMHNTKGRTWSLMLVSLI
IFIATTNLLGLLPHSFTPTTQLSMNLAMAIPLWAGTVI
MGFRSKIKNALAHFLPQGTPTPLIPMLVIIETISLLIQ
PMALAVRLTANITAGHLLMHLIGSATLAMSTINLPSTL
IIFTILILLTILEIAVALIQAYVFTLLVSLYLHDNT
```

```
MLPPMALPSVSWMLLK
CLMLLSQVQGEEPQR
ELPSAR
IR
CPK
GSK
AYGSHCYALFLSPK
SWTDADLACQK
PSGNLVSVLSGAEGSFVSSLVK
SIGNSYSYVWIGK
HDPTQGTEPNGR
GWEWSSSDVMNYFAWER
NPSTISSPGHCASLSR
…
```

Quantitative Proteomics

Calibration curve



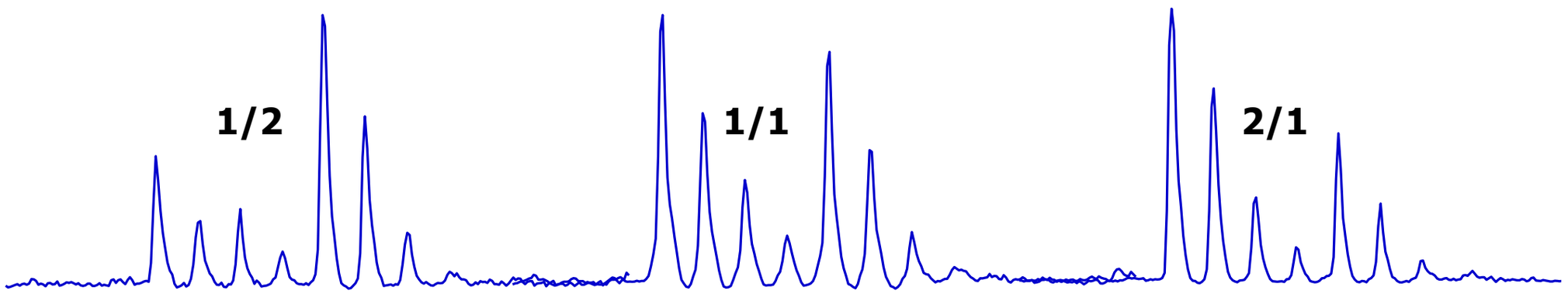
Different peptides have different slope

Why is protein quantification with mass spectrometry non-trivial?

Many of the steps from a protein to an ion current are non-predictively non-linear like:

- The efficiency of the enzyme
- The yield of the chromatography (some peptides do precipitate)
- The efficiency of the electrospray

Hence, relative quantification by mass spectrometry is easier than absolute quantification

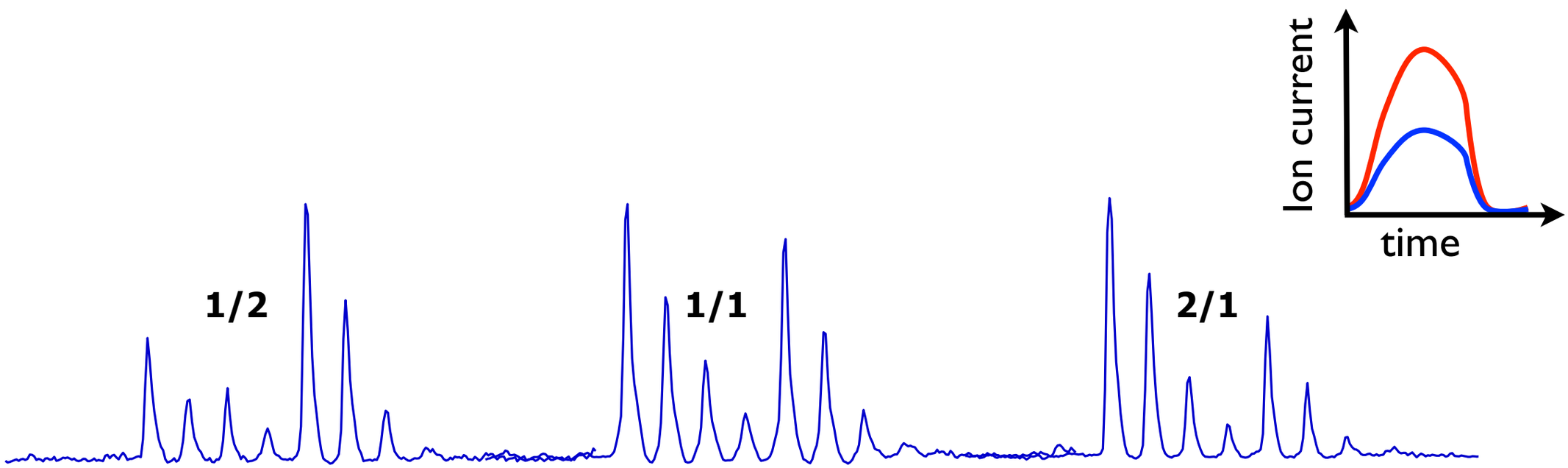


The primary principles in quantitation

- Make each sample distinguishable by either:

(1) introduce mass differences between the samples

(2) perform distinct experimental runs for each sample



Protein quantitation

<u>LC-MS(/MS) - protein or culture level labeling</u>

- SILAC (2), cell cultures, relative

<u>LC-MS/MS - peptide level labeling</u>

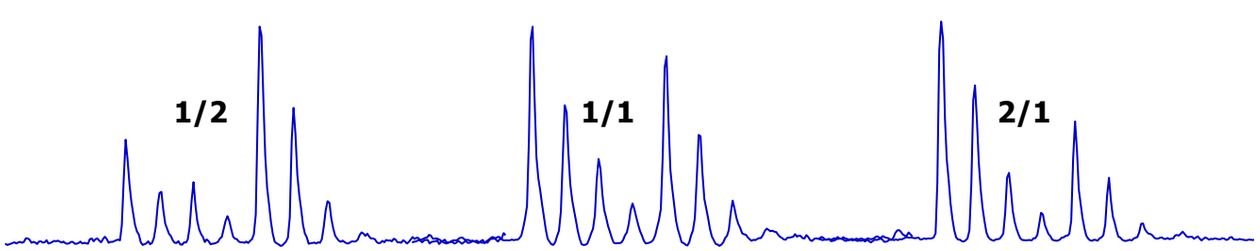
- iTRAQ (2), peptides, relative
- Spiked peptides (eg. AQUA) (2), peptides, absolute

<u>LC-MS(/MS) - label free</u>

- MS Alignments (3), peptides, relative or absolute
- Spectral counting (3), peptide fragments, relative

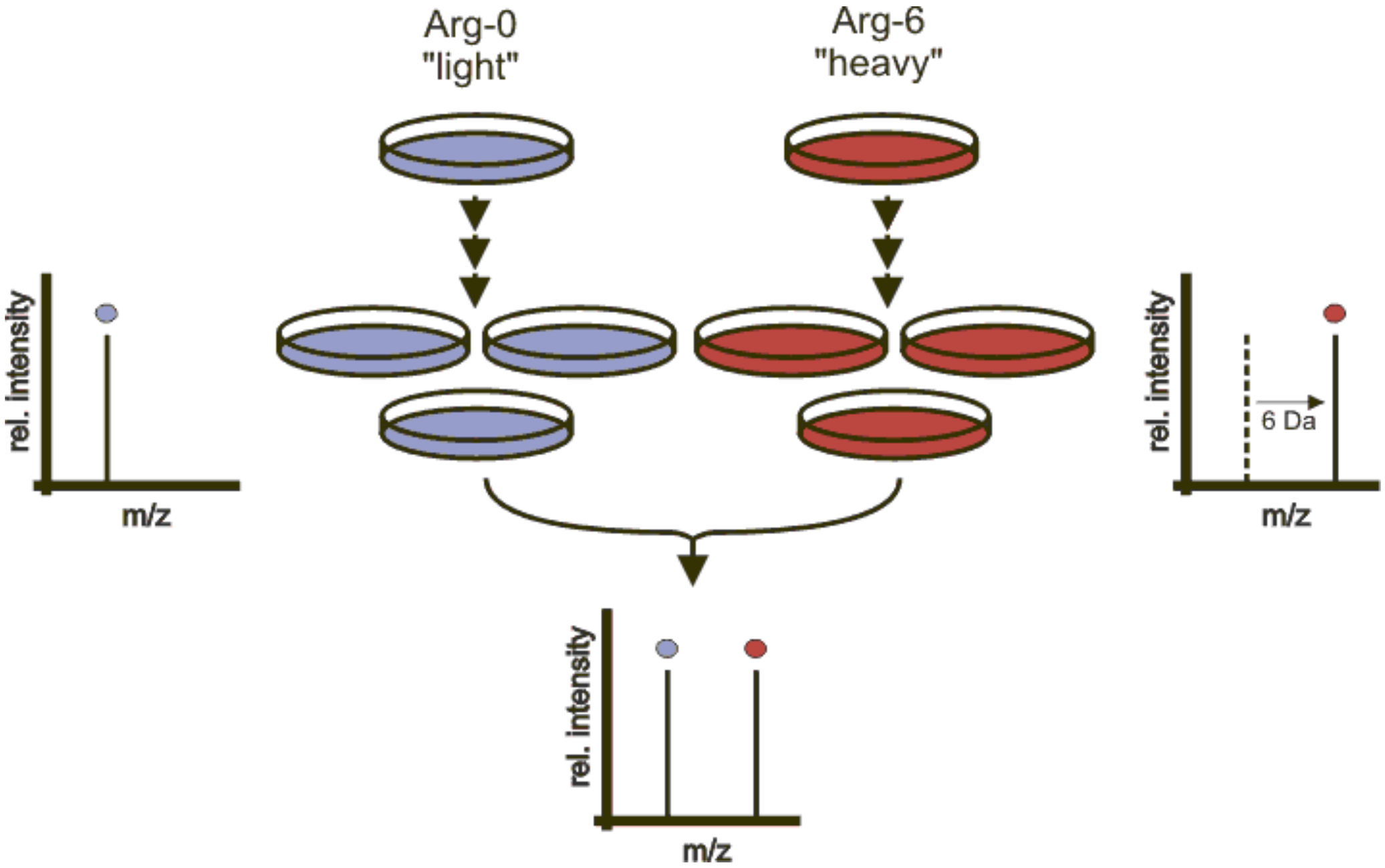
<u>Selected Reaction Monitoring</u>

- With or without labels

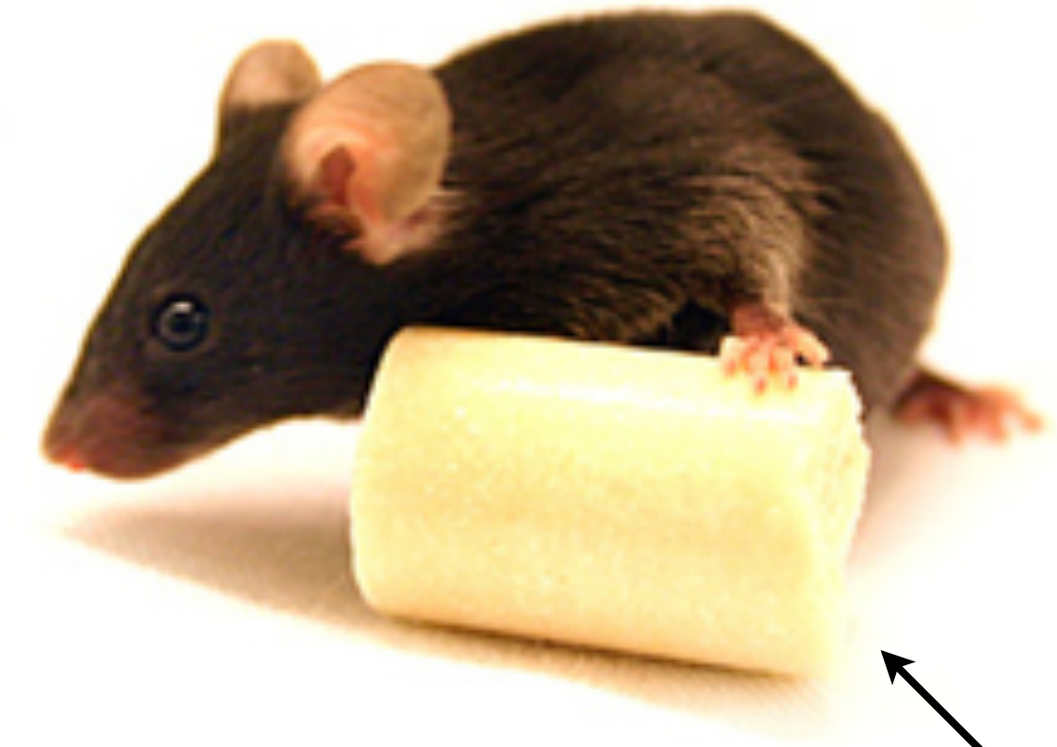


Culture-level labeling

SILAC (in-vivo heavy isotope label)



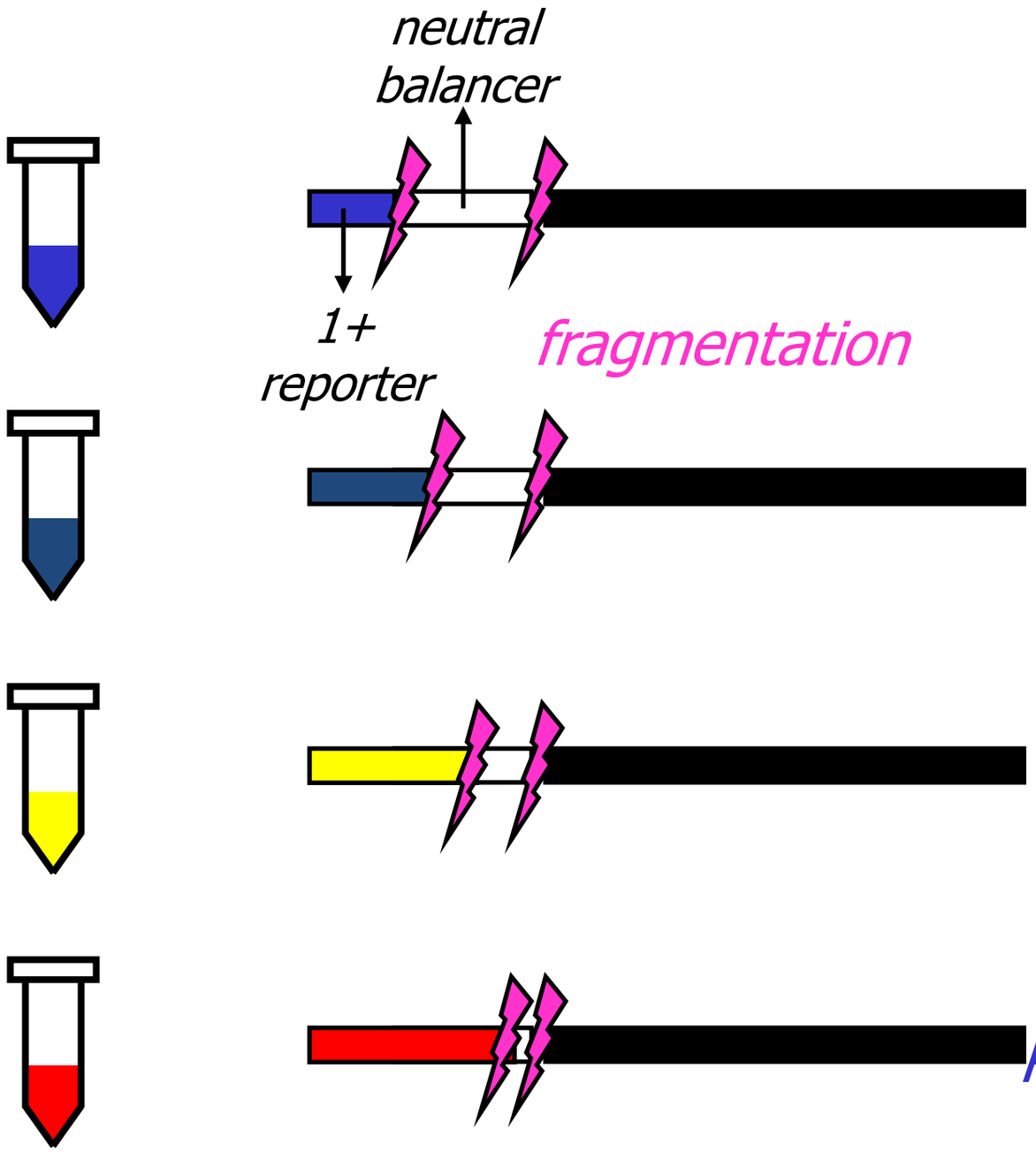
SILAC Mouse



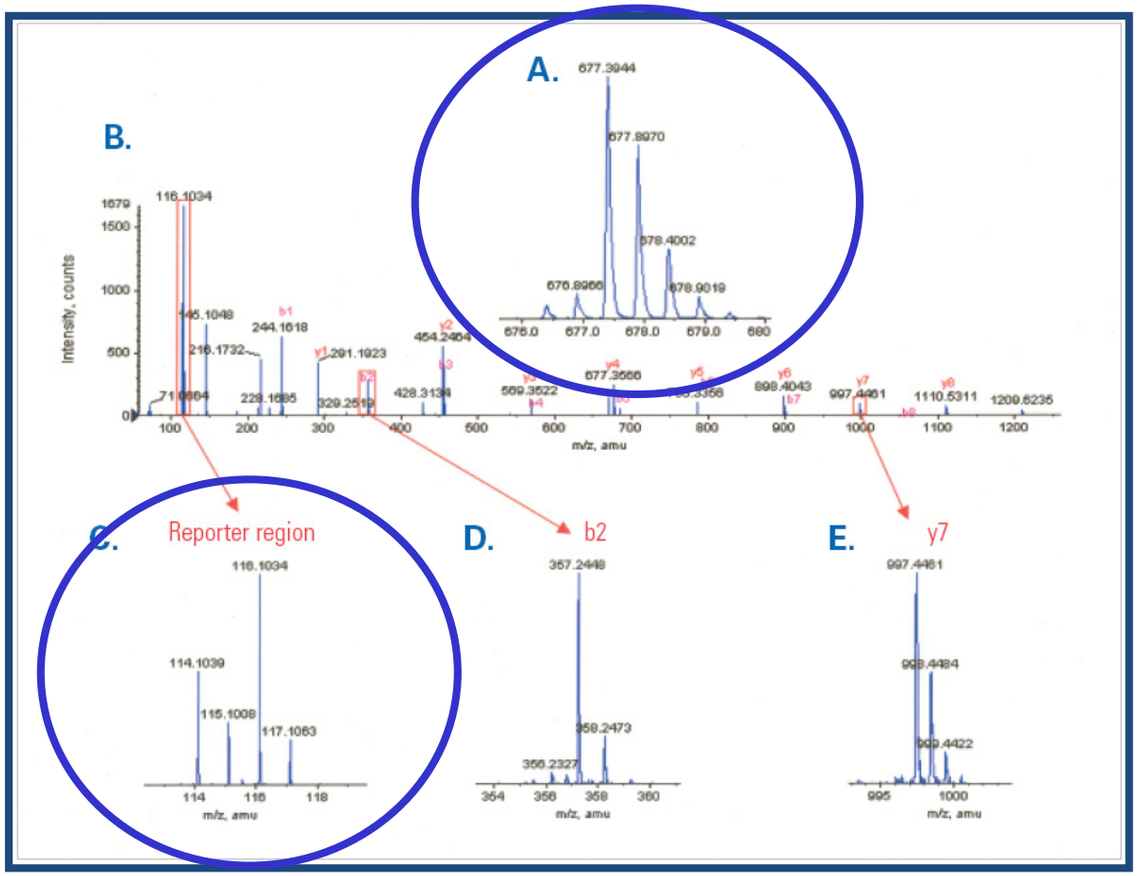
Heavy Cheese

Petide-level chemical modifications

iTRAQ - MS/MS fragment (isobaric) labeling of peptides



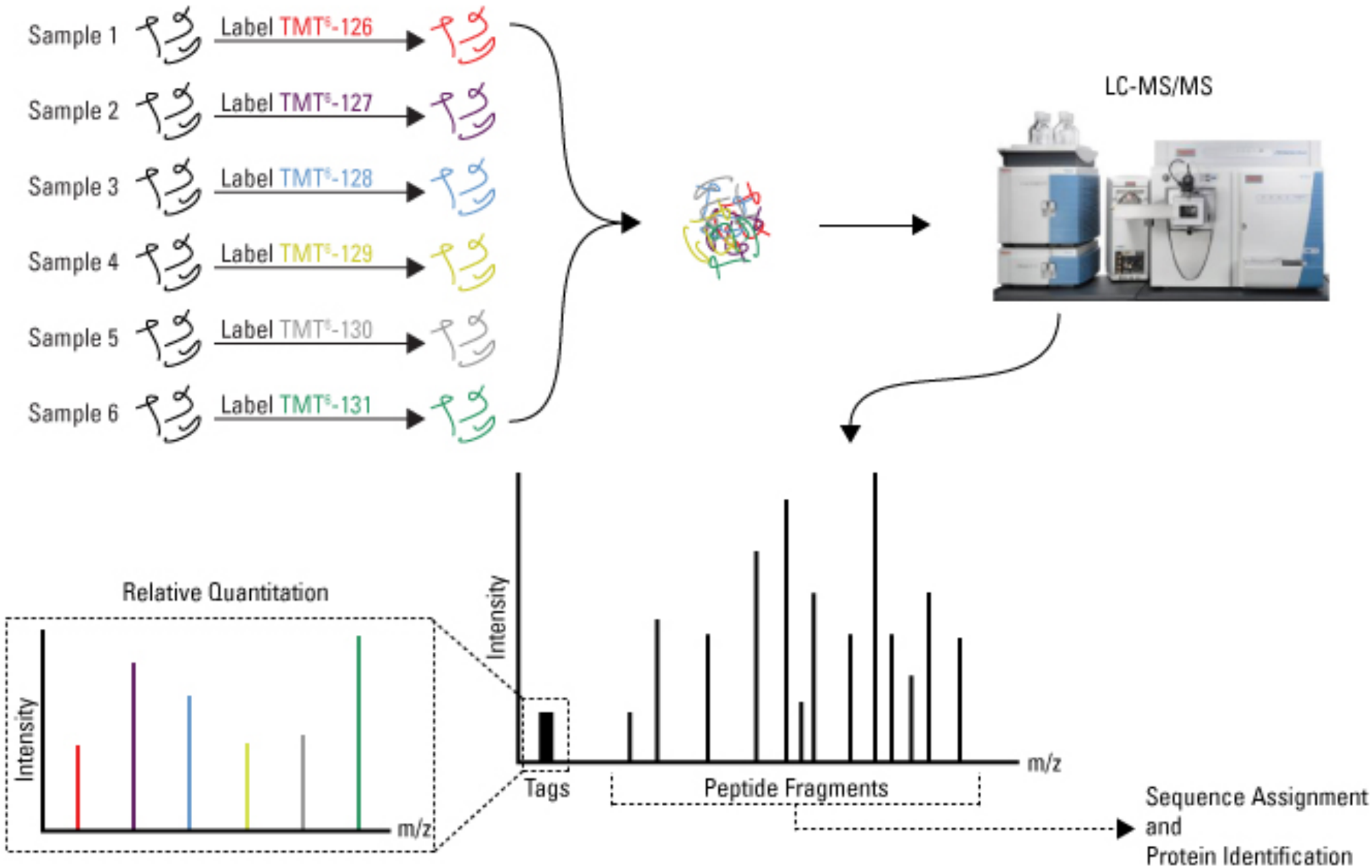
precursors



Reporter fragments

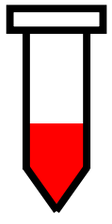
Relative sample abundance visible in fragmentation spectra

TMT - Yet another iso-baric label

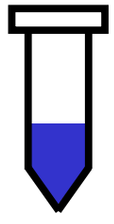


AQUA (peptide heavy isotope label)

Aimed at absolute quantitation



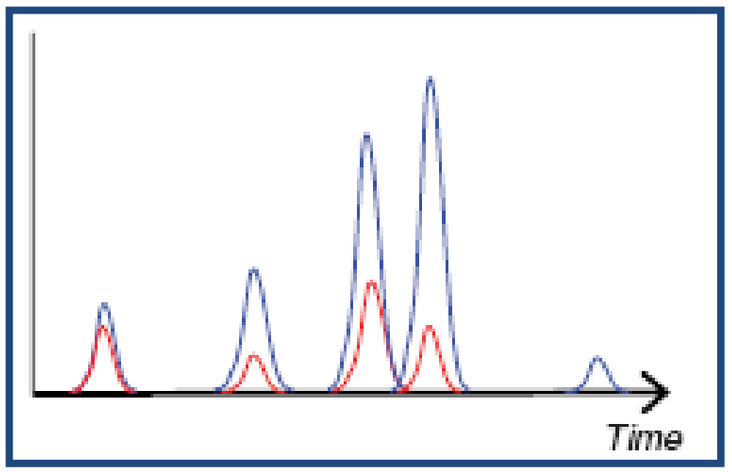
sample
unlabelled peptides
unknown abundance



internal standard
labelled, synthesized peptides
known abundance

Mix in desired ratio

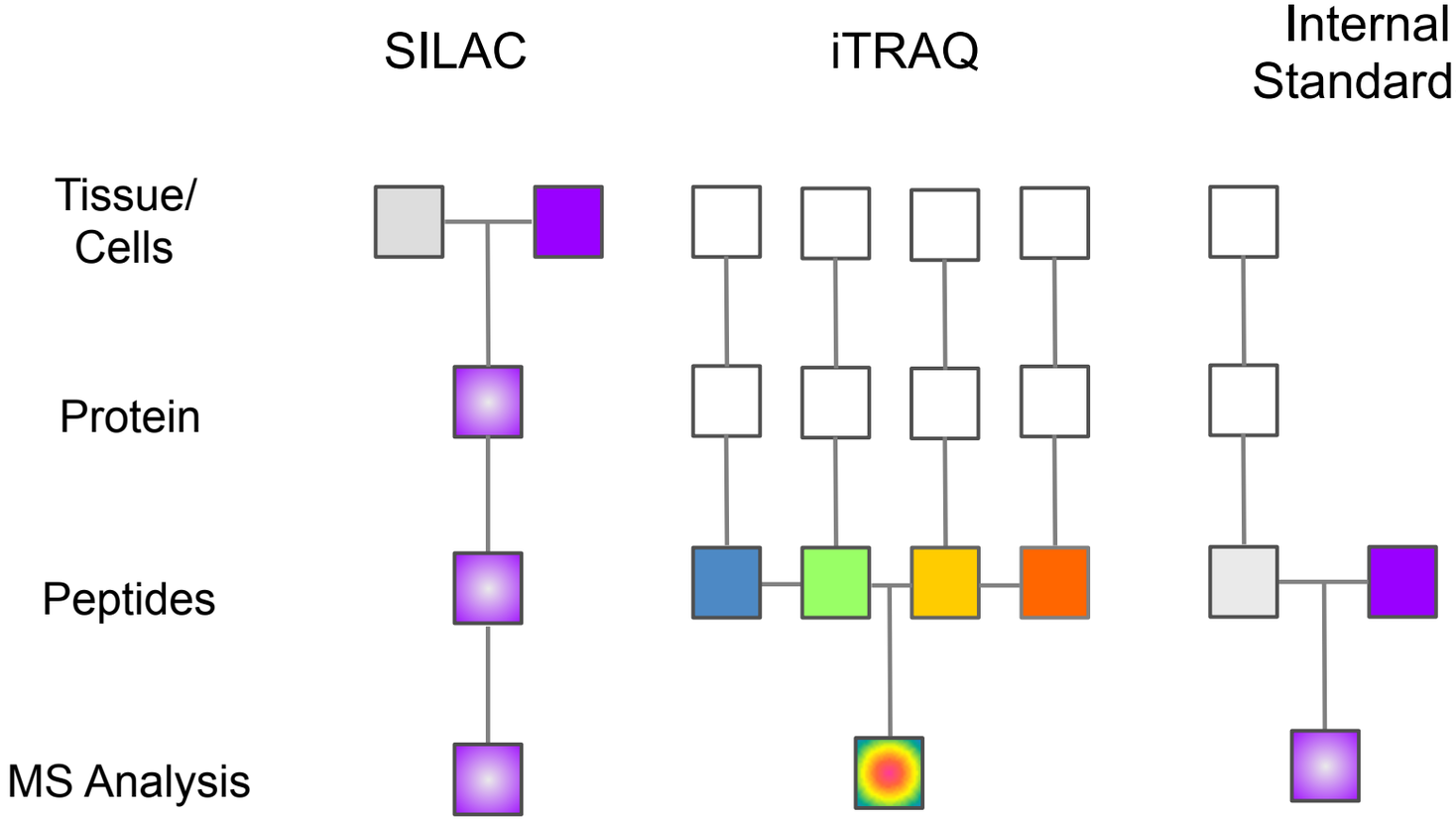
MS analysis

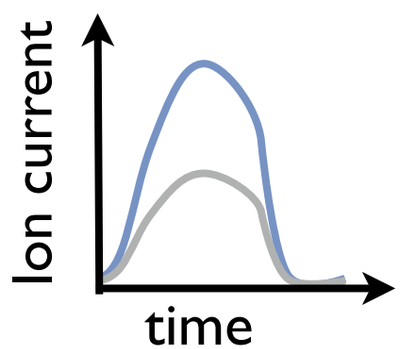


Compare signal intensities
Derive absolute quantitation

Labeled Quantification methods

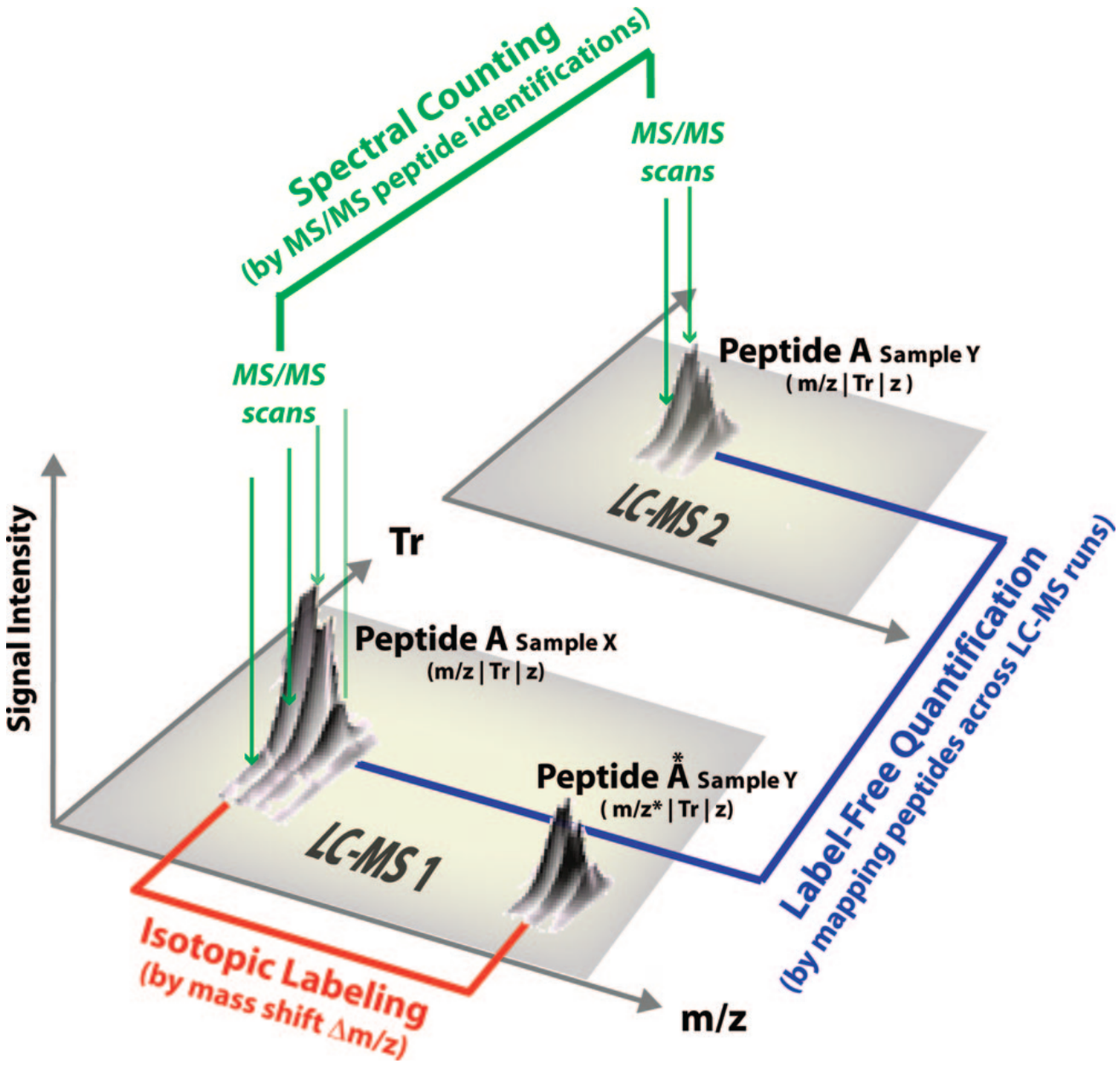
At what point are the samples combined?





Label Free Quantification

Spectral counting



(1) Count the number of fragmentation events that are triggered

Mueller *et al.*, *J. Proteome Research*, 7, p. 51, 2008

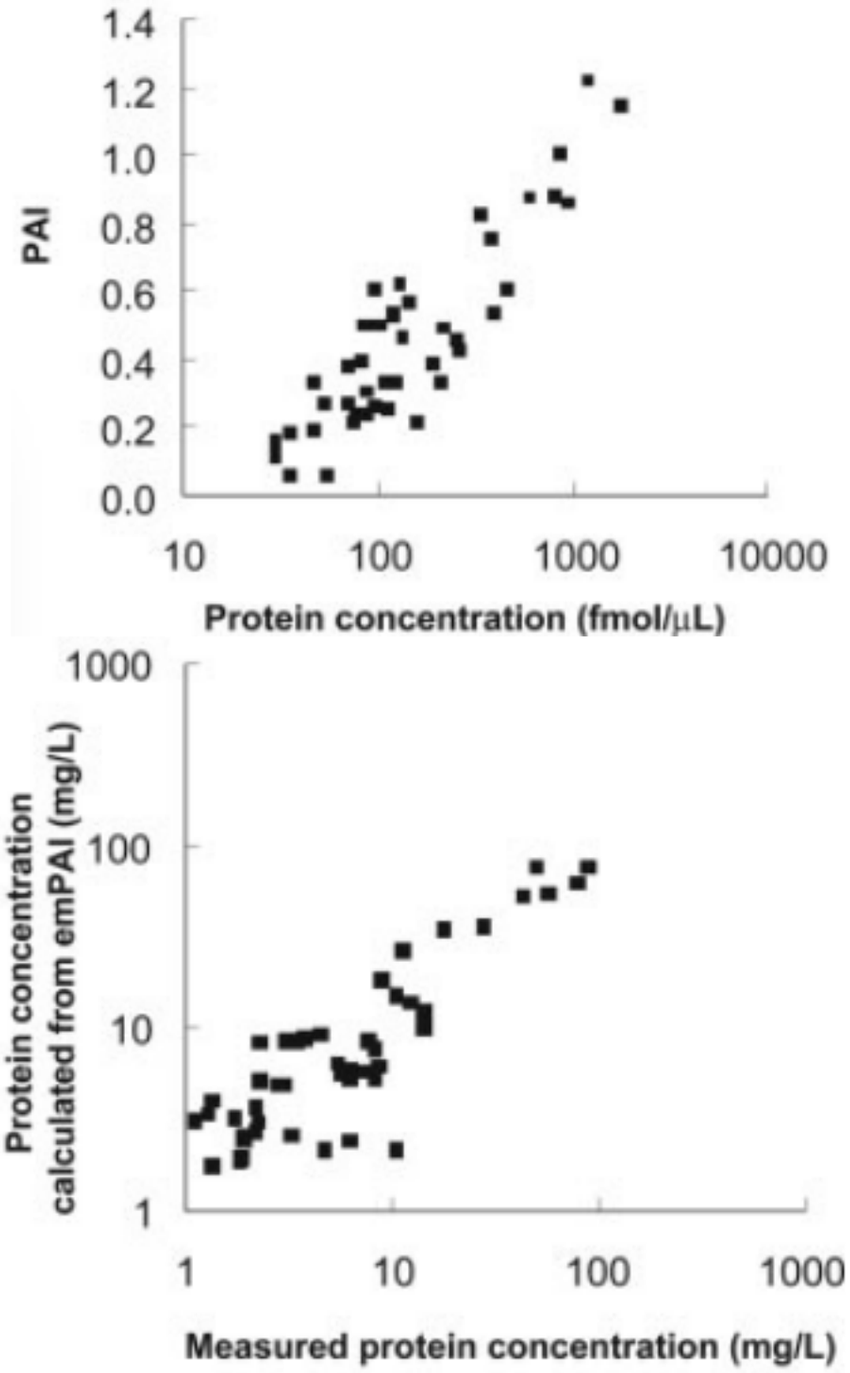
emPAI (Exponentially Modified Protein Abundance Index)

$emPAI = 10^{PAI} - 1$

Where $PAI = N_{observed} / N_{observable}$

What is an 'observable' peptide?

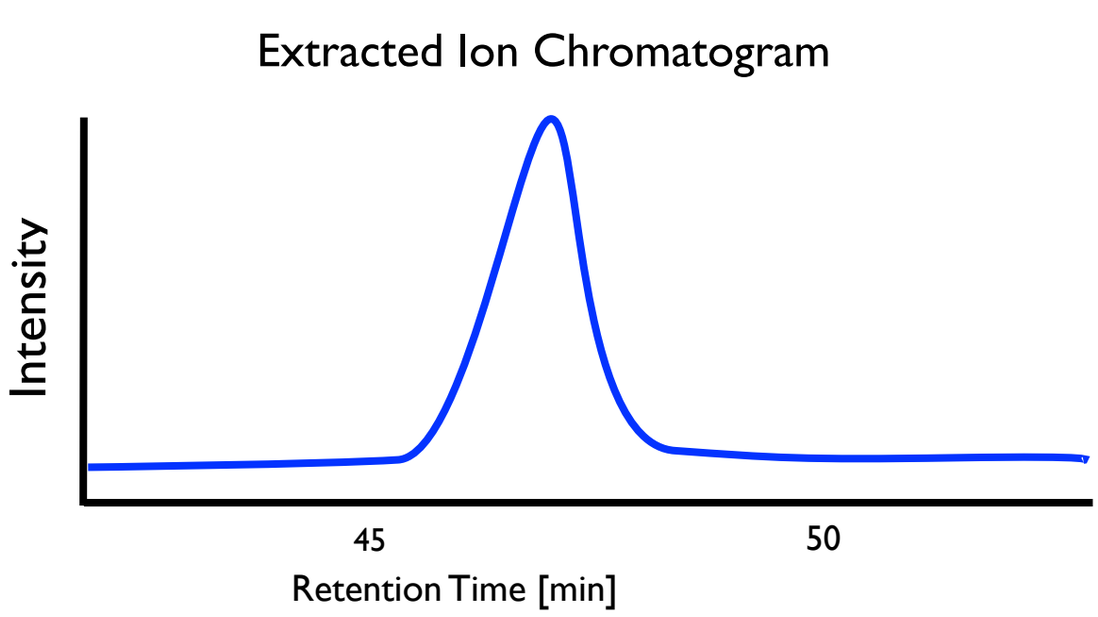
- Peptides with a precursor mass between 800-2400Da.
- There is a roughly linear relationship between log protein concentration and the ratio of 'observable' peptides observed in range of 3-500 fmoles.
- If you know how much total protein you analyzed you can derive absolute abundances.

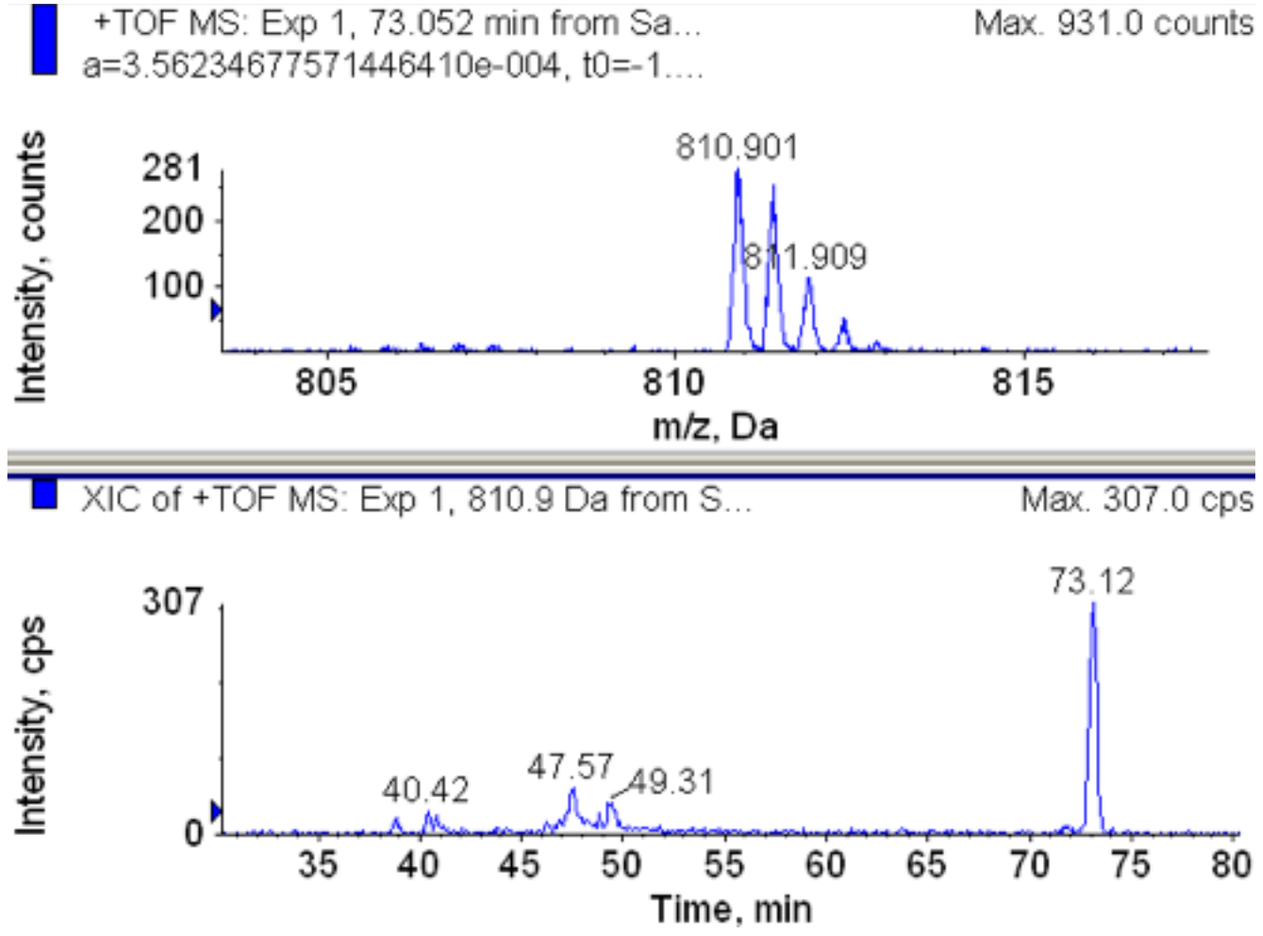


Ishihama et al. *Mol Cell Proteomics* (2005) **4** 9 1265-1272

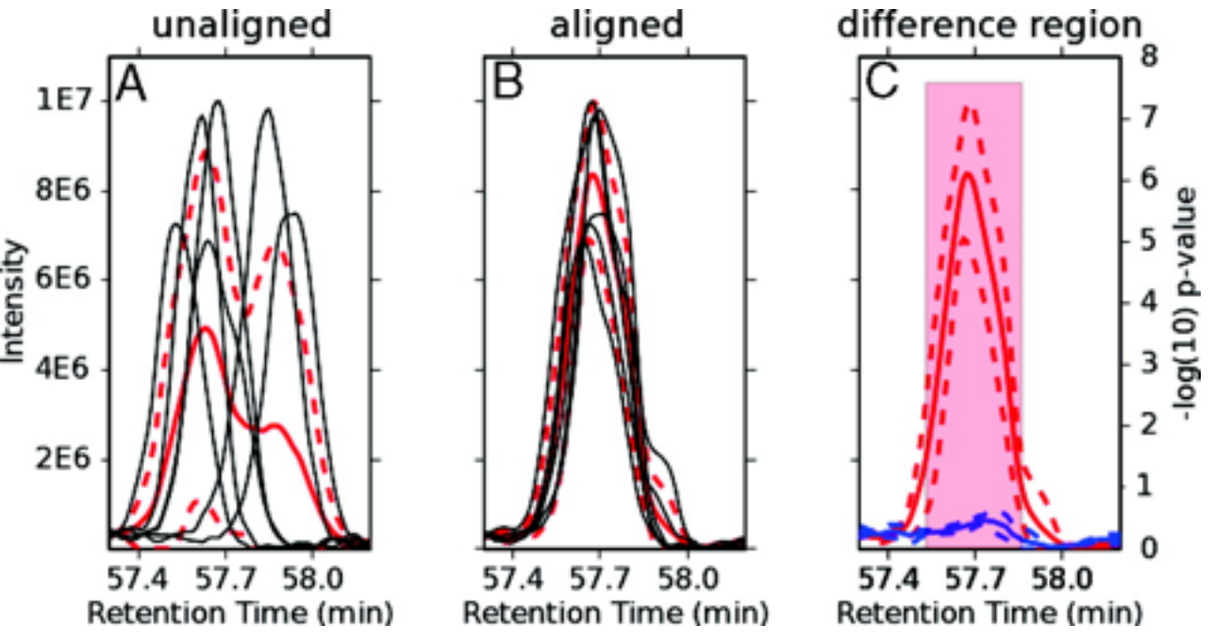
XIC (Extracted Ion Chromatogram)

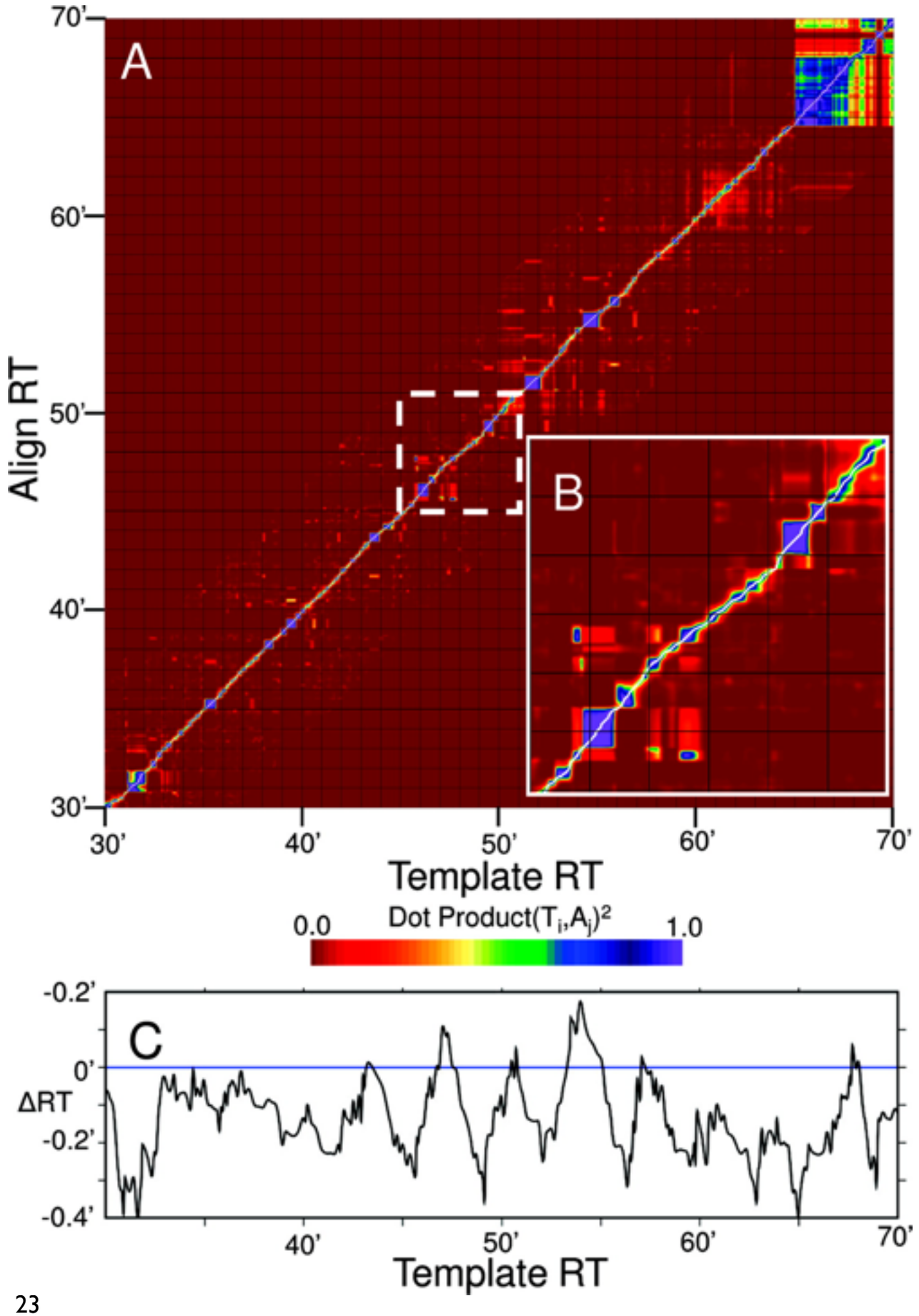
- Measure intensity of peak during its elution off HPLC column and into the mass spectrometer.
- Measure area of peak in XIC.
- More accurate than selecting peak intensity for one given scan.





Label Free: MS Alignments

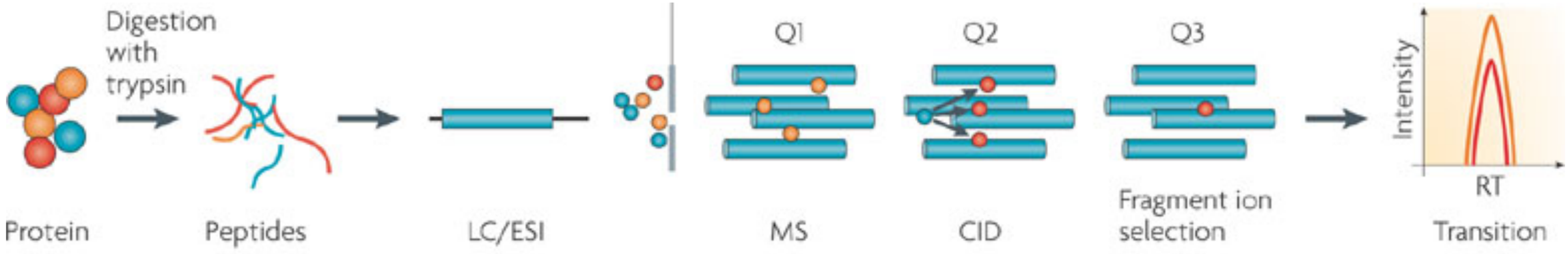




[Finney et al *JPR* 2008]

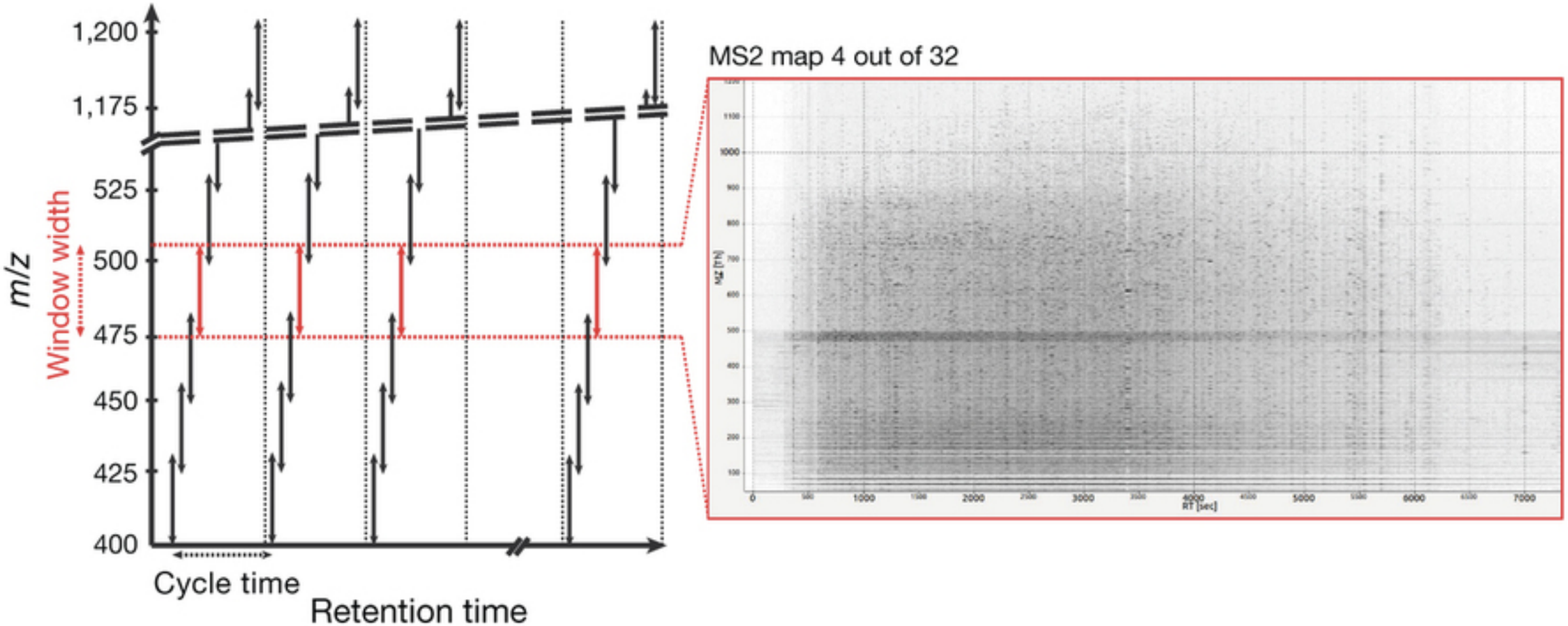
Other Acquisition Techniques

Selected Reaction Monitoring (SRM or MRM)



+ Higher sensitivity (low-attomolar)
+ Broader dynamic range (up to 5 orders of magnitude)
- Requires known protein targets and their characteristics
- max 200-300 measurements per run

Data Independent Analysis (DIA or SWATH)



+ Highly reproducible
+ Broad dynamic range (~4 orders of magnitude)
+ Measures a large number of proteins per run
- Requires known protein targets and their characteristics

Comparison of acquisition techniques

	Shotgun	**SRM**	**DIA**
Quantifiable Proteins per run	1000s	100s	1000s
Reproducibility	++	++++	+++
Sensitivity	+++	++++	+++
Dynamic Quantification Range	3-4 magnitudes	5 magnitudes	3-4 magnitudes

Selected Applications

Identification of components in protein complexes

Tandem Affinity Purification

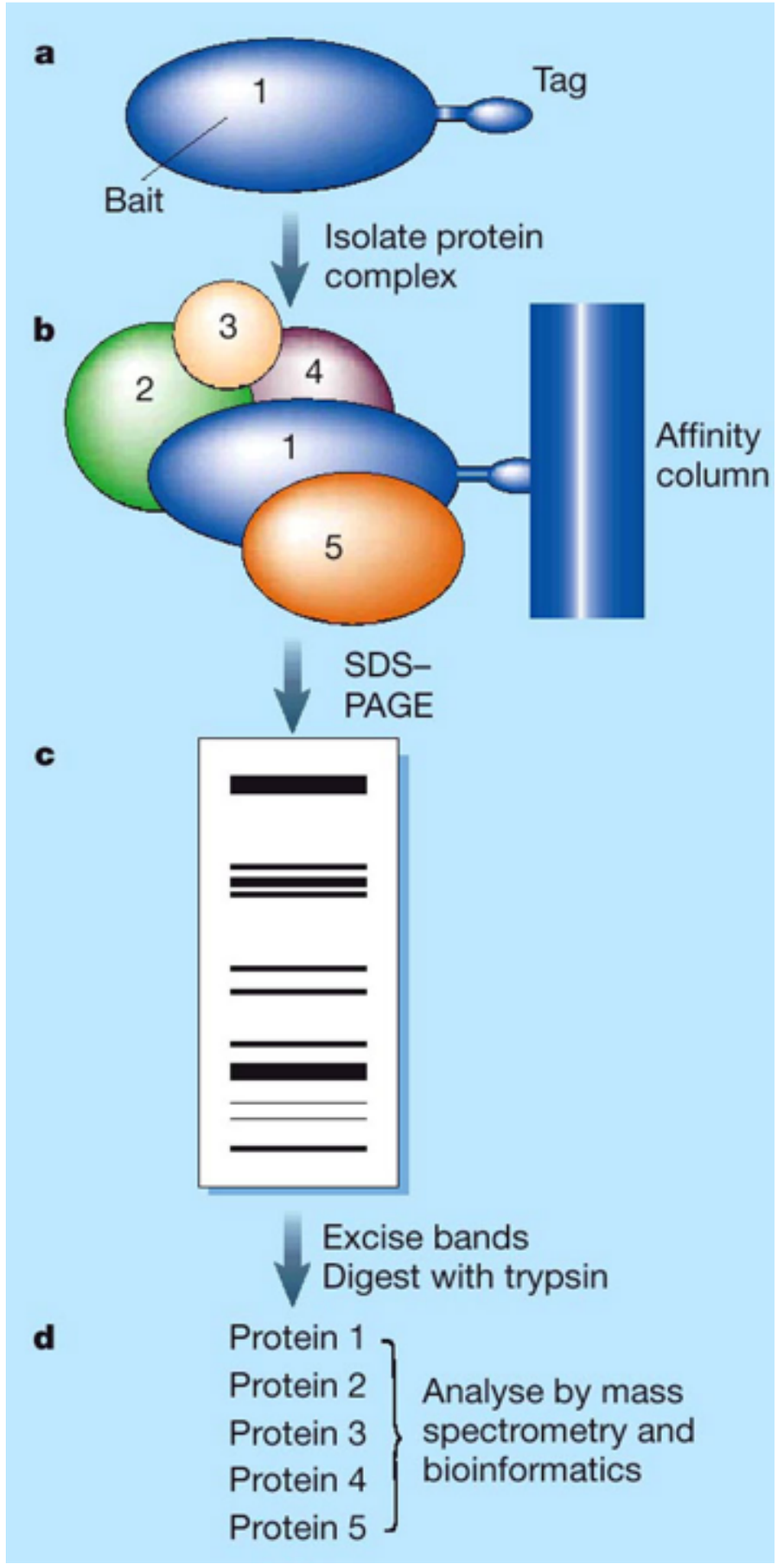
a) An 'affinity tag' is first attached to a target protein (the 'bait')

b) Bait proteins are precipitated, along with its protein complex

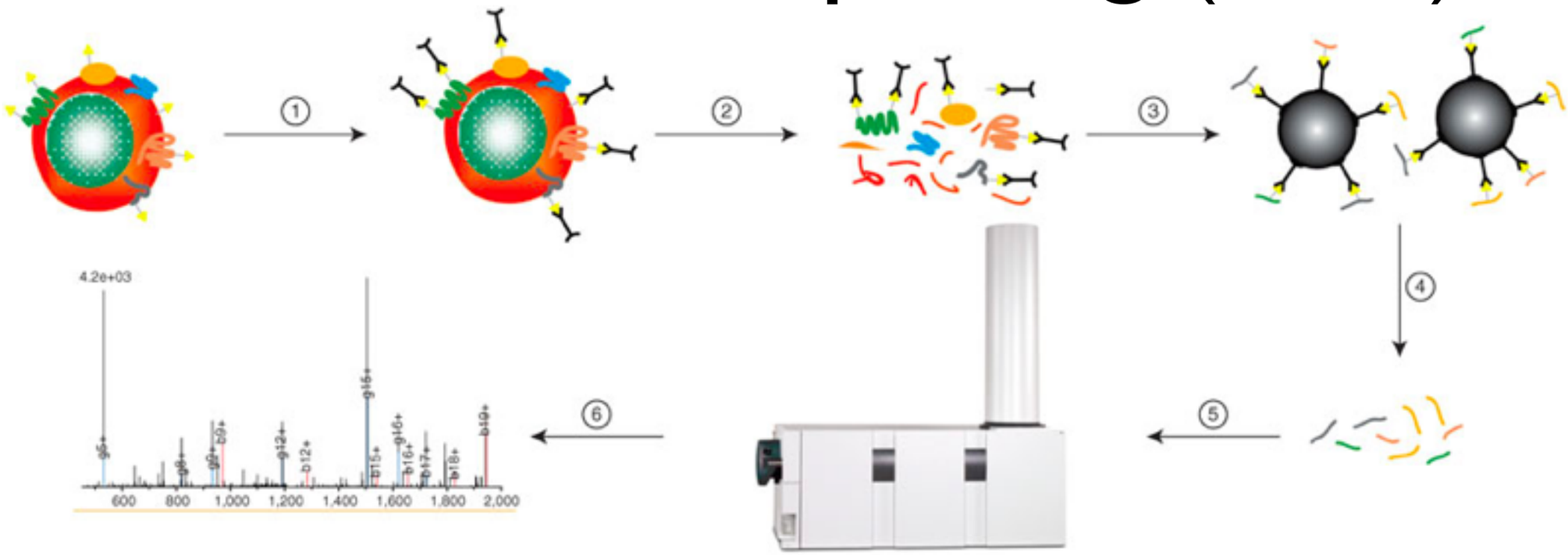
c) Purified protein complexes are resolved by one-dimensional SDS–PAGE

d) Proteins are excised from the gel, digested with trypsin, and analyzed by mass spectrometry.

[Kumar & Snyder, 2002]



Example: Barcoding cell types with Cell surface–capturing (CSC)



Identifies 110 cell surface proteins

Prof of principle: differentiate Ramos B cells from Jurkat T cells

[Wollscheid *et al.*, Nat Biotech 2009]