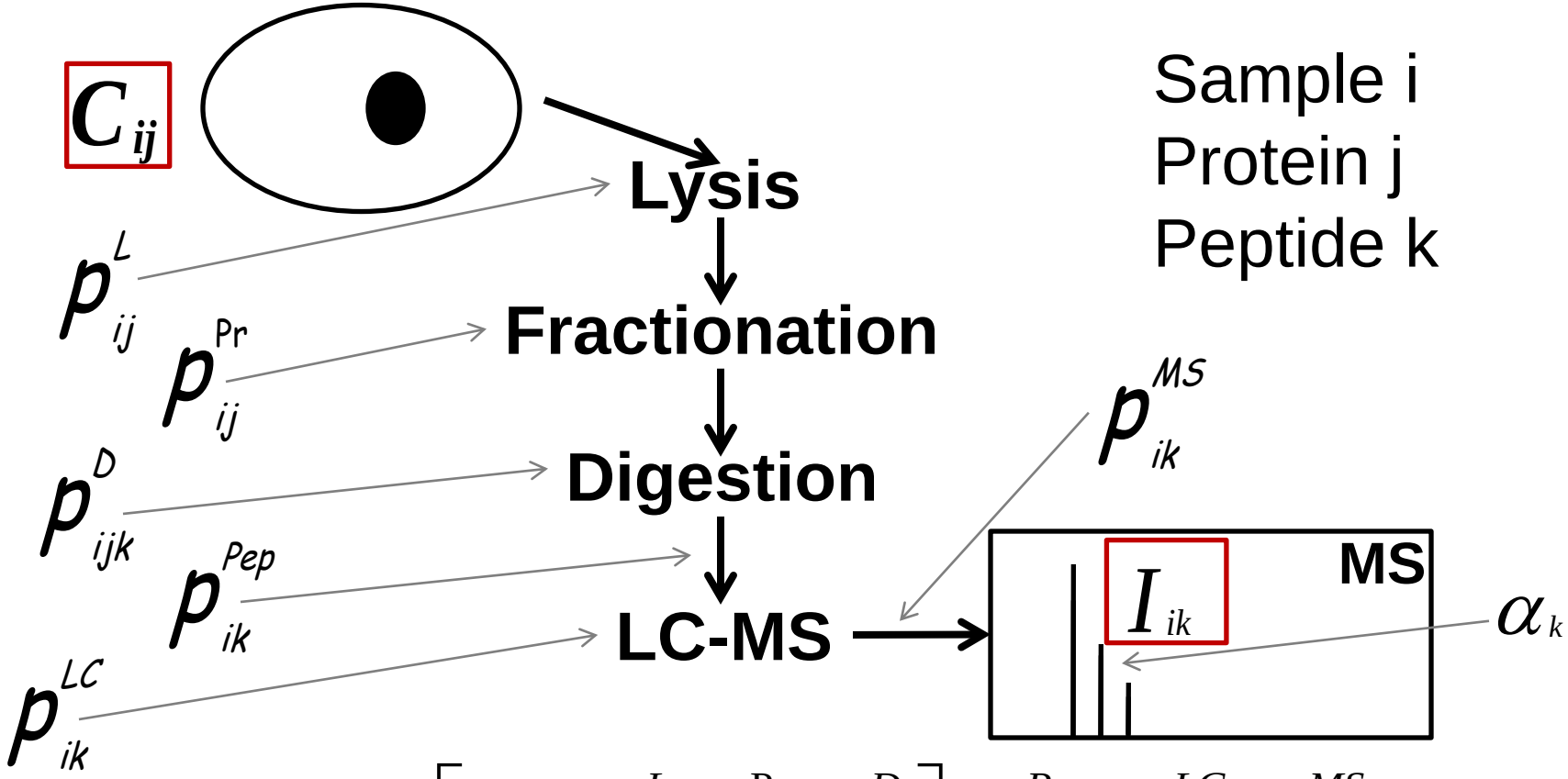
Protein quantitation I: Overview (Week 8)

Proteomic Bioinformatics – Quantitation

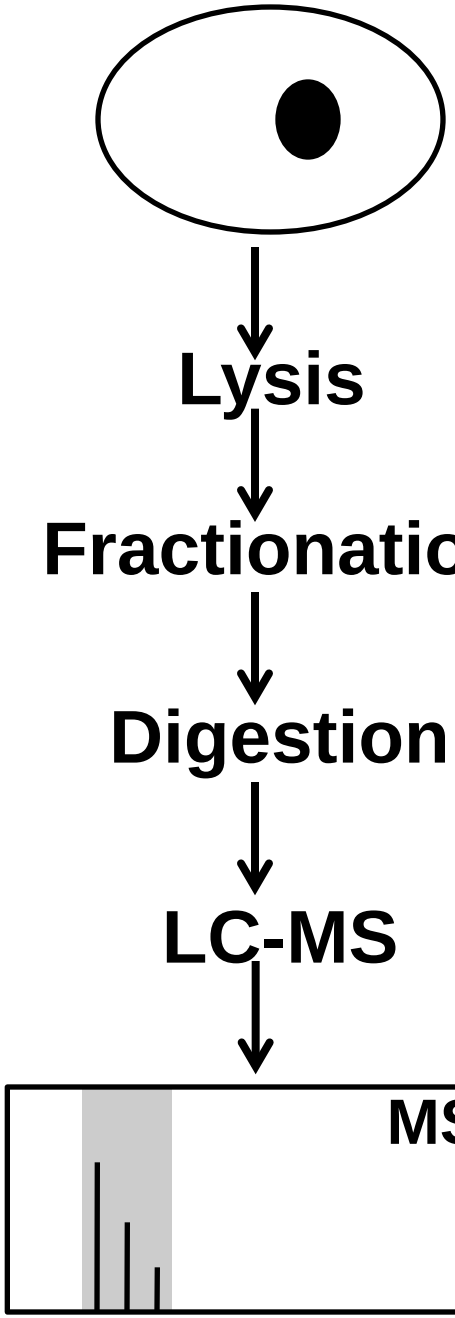


Sample i
Protein j
Peptide k

$$I_{ik} = \alpha_k \sum_j \left[C_{ij}\, p^{L}_{ij}\, p^{Pr}_{ij}\, p^{D}_{ijk} \right] p^{Pep}_{ik}\, p^{LC}_{ik}\, p^{MS}_{ik}$$

$$C^{k}_{ij} = \frac{I_{ik}}{\alpha_k\, p^{L}_{ij}\, p^{Pr}_{ij}\, p^{D}_{ijk}\, p^{Pep}_{ik}\, p^{LC}_{ik}\, p^{MS}_{ik}}$$

Quantitation – Label-Free (Standard Curve)



Sample i
Protein j
Peptide k

$$C_{ij}^{k} = f(I_{ik}) \approx \beta I_{ik}$$

Quantitation – Label-Free (MS)

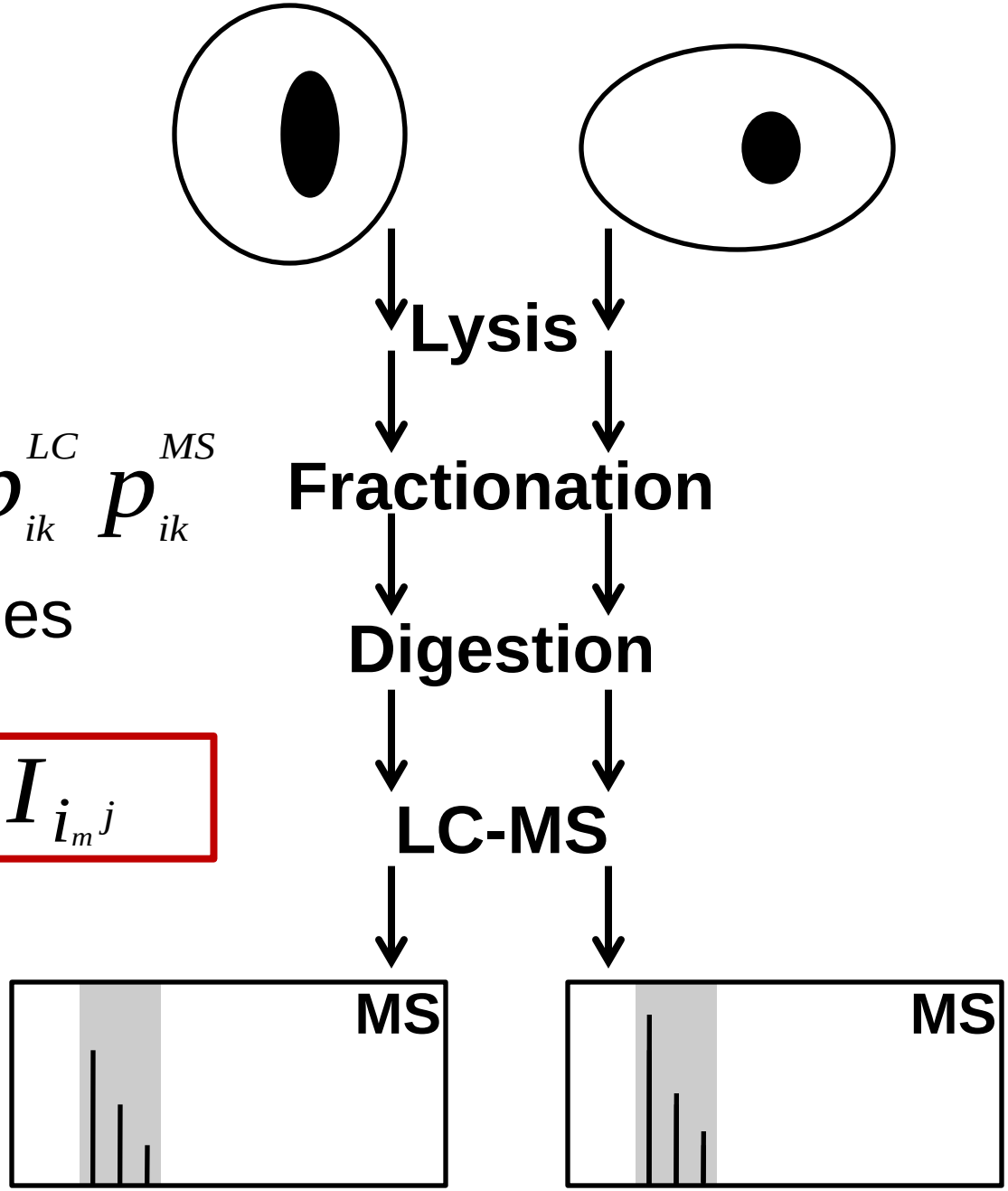
Sample i
Protein j
Peptide k

Assumption:

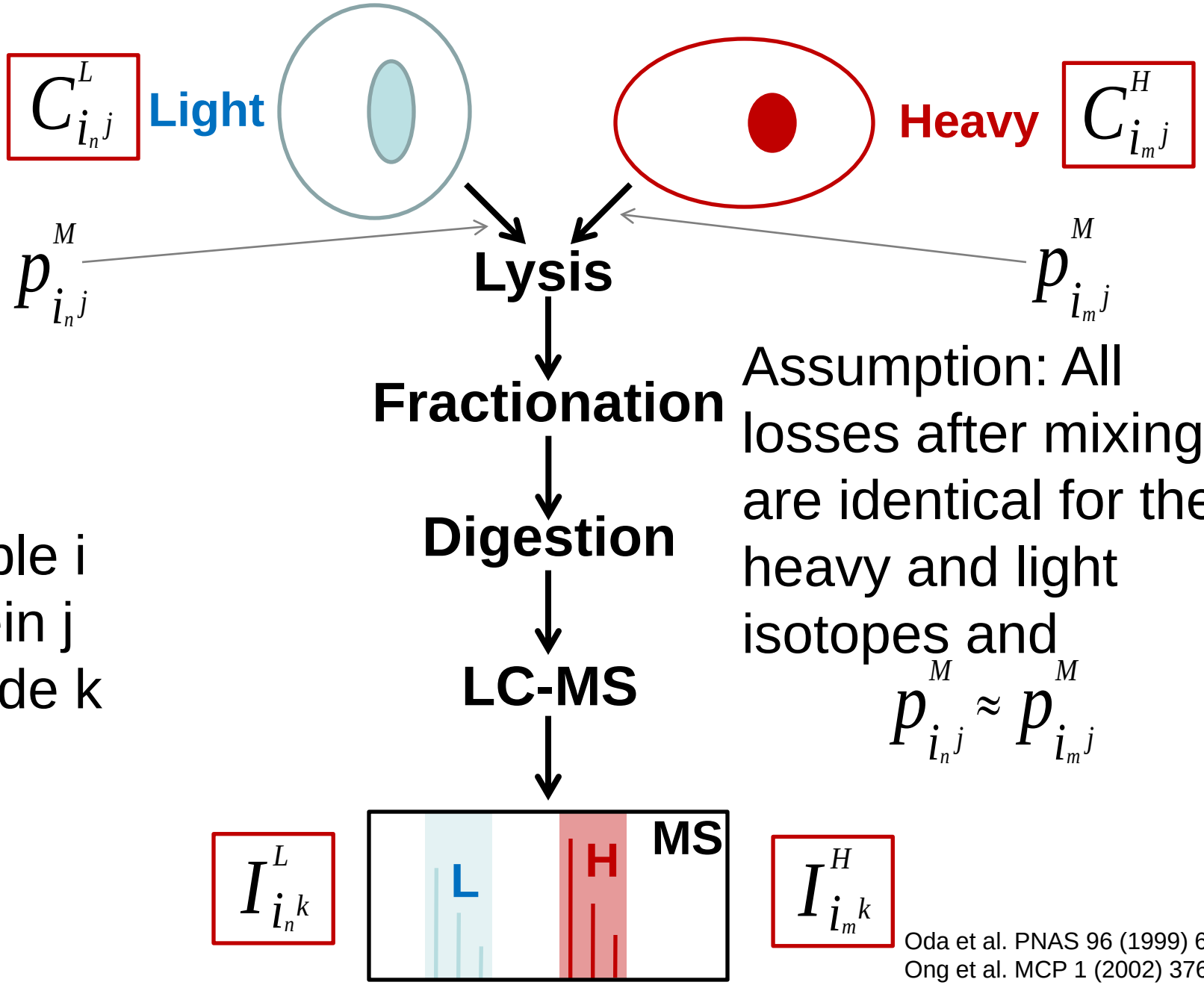
$$\alpha_k \, p_{ij}^{L} \, p_{ij}^{\Pr} \, p_{ijk}^{D} \, p_{ik}^{Pep} \, p_{ik}^{LC} \, p_{ik}^{MS}$$

constant for all samples

$$C_{i_n j} / C_{i_m j} = I_{i_n j} / I_{i_m j}$$



Quantitation – Metabolic Labeling



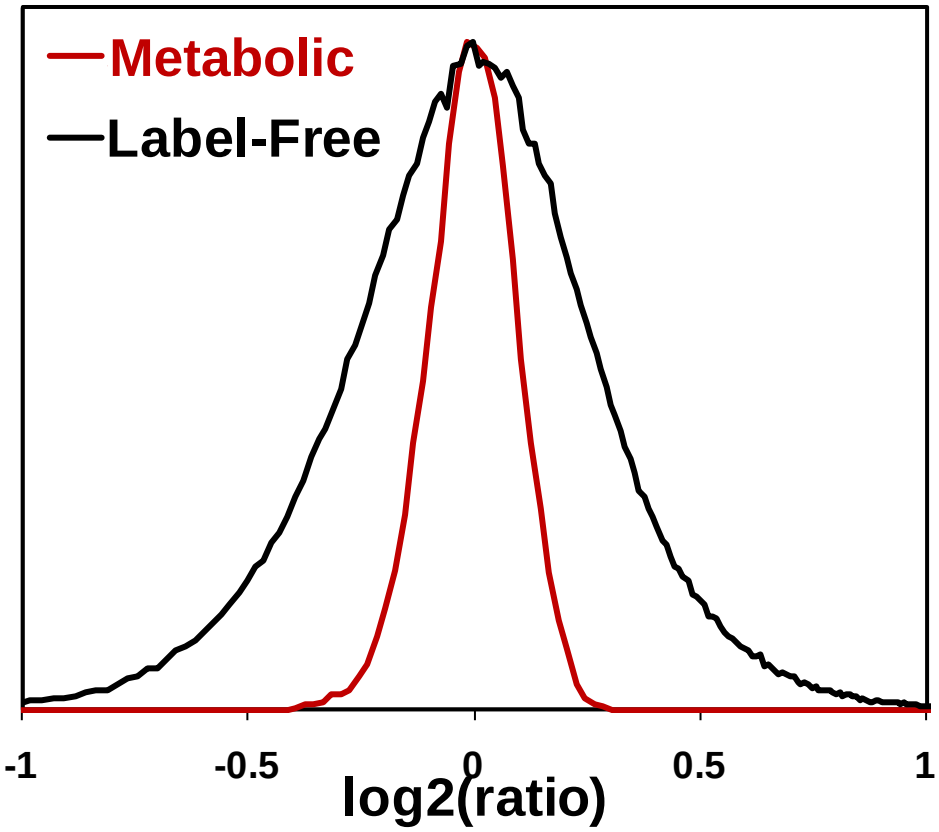
Sample i
Protein j
Peptide k

Assumption: All losses after mixing are identical for the heavy and light isotopes and

$$p^{M}_{i_n j} \approx p^{M}_{i_m j}$$

Oda et al. PNAS 96 (1999) 6591
Ong et al. MCP 1 (2002) 376

Comparison of metabolic labeling and label-free quantitation



Label free assumption:

$$\alpha_k \, p_{ij}^{L} \, p_{ij}^{\mathrm{Pr}} \, p_{ijk}^{D} \, p_{ik}^{Pep} \, p_{ik}^{LC} \, p_{ik}^{MS}$$

constant for all samples

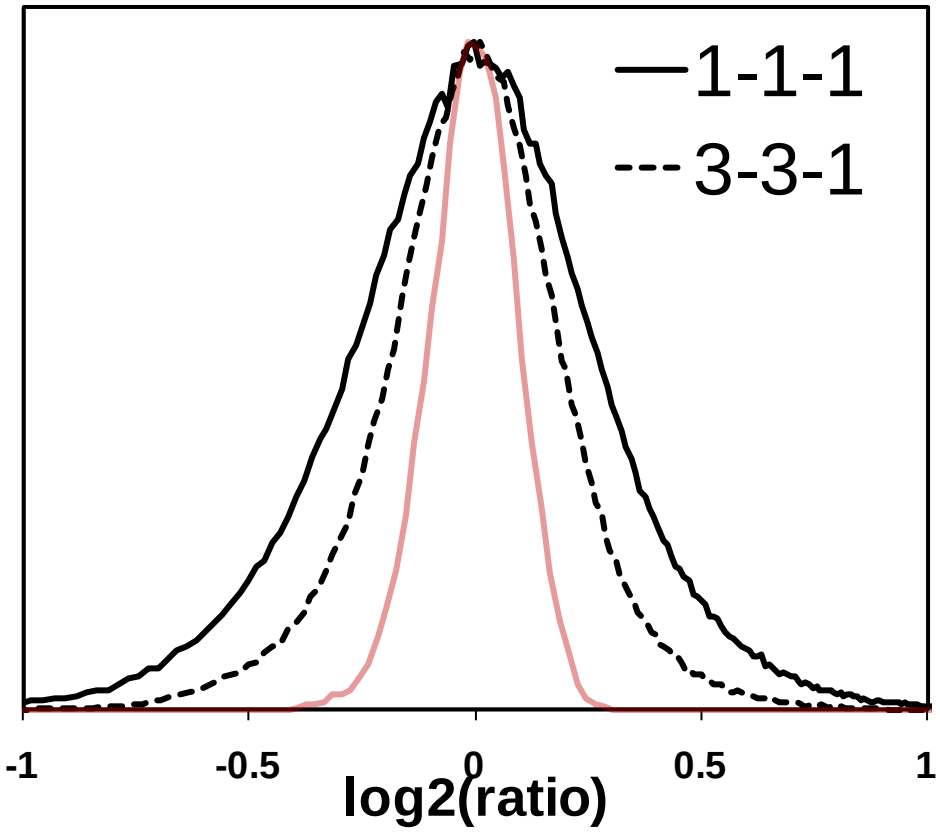
Metabolic labeling assumption:

$$p_{ij}^{M}$$

constant for all samples and the behavior of heavy and light isotopes is identical

G. Zhang et al., JPR 8 (2008) 1285-1292

Intensity variation between runs

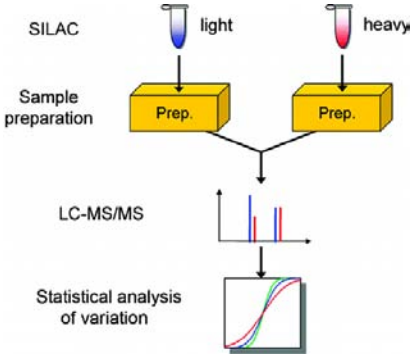


Replicates

1 IP
1 Fractionation
1 Digestion

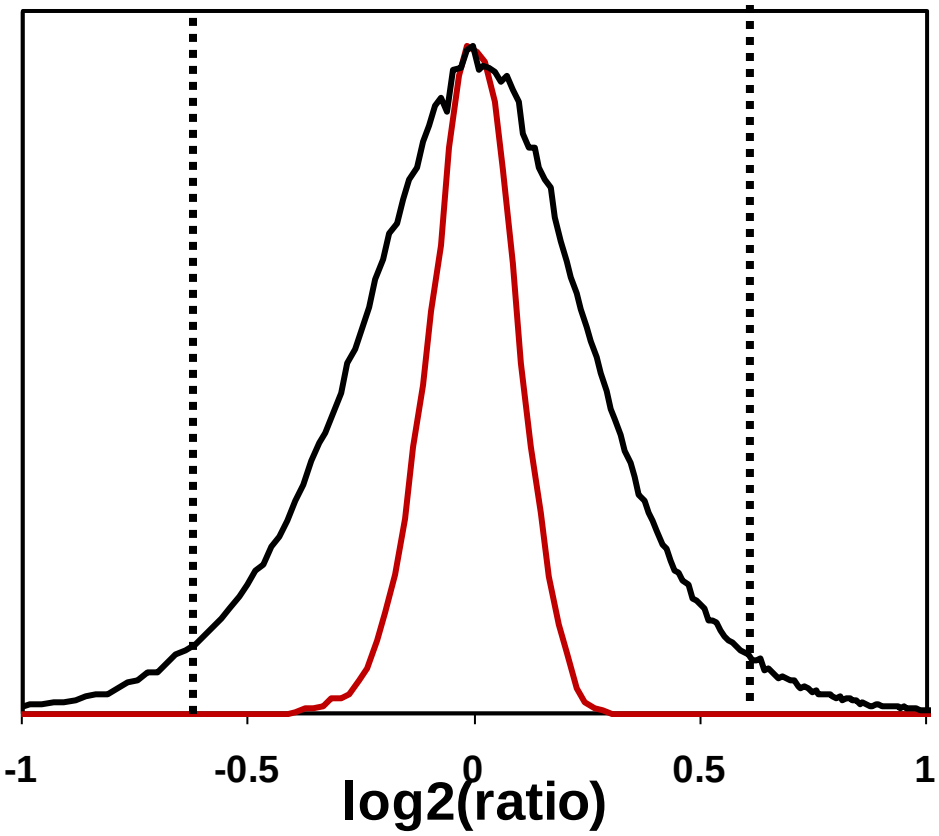
vs

3 IP
3 Fractionations
1 Digestion



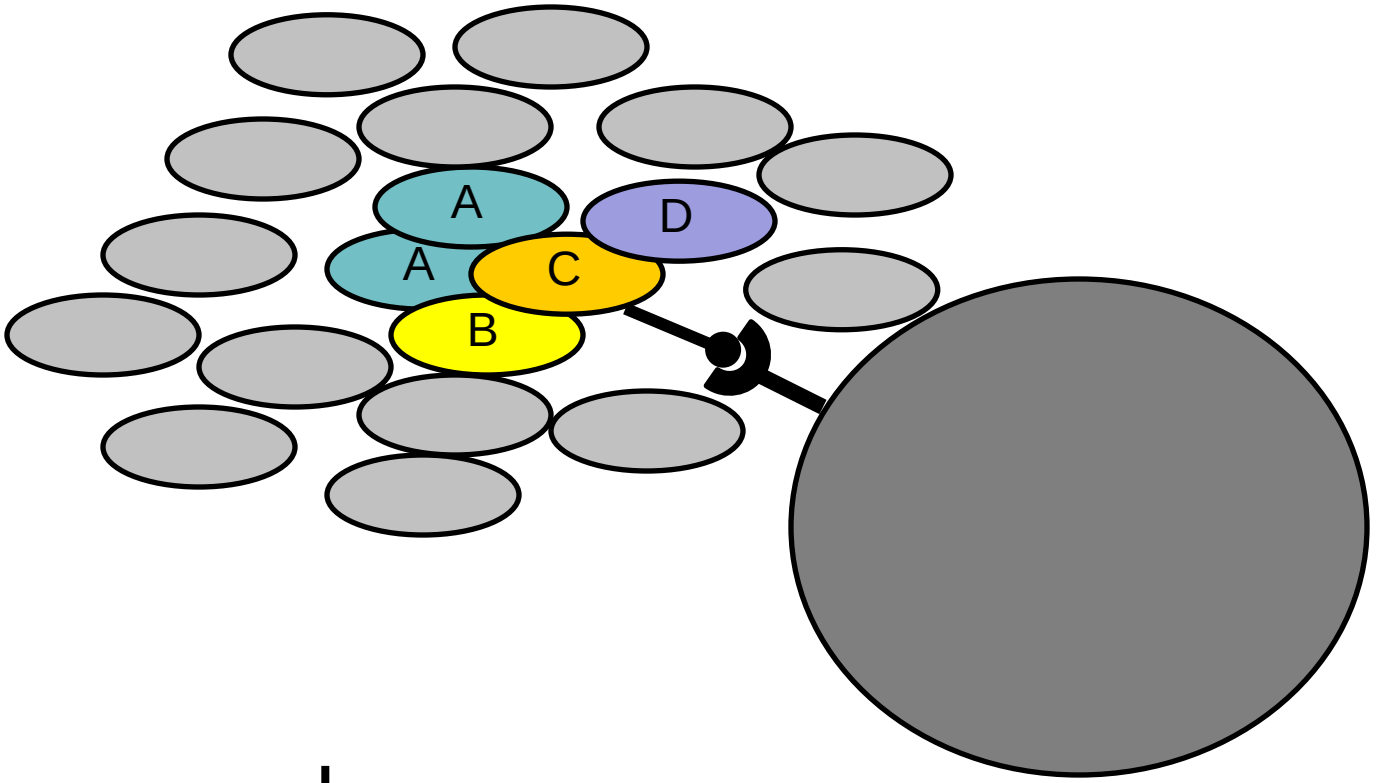
G. Zhang et al., JPR 8 (2008) 1285-1292

How significant is a measured change in amount?



It depends on the size of the random variation of the amount measurement that can be obtained by repeat measurement of identical samples.

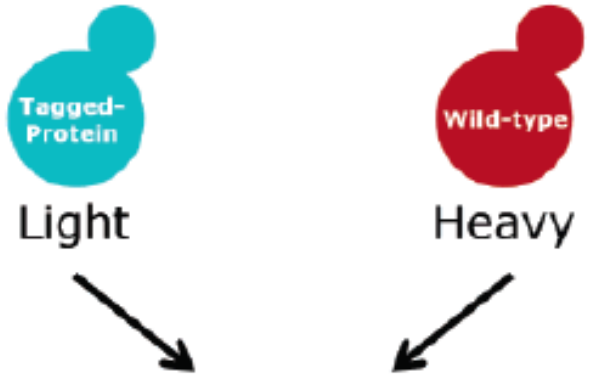
Protein Complexes



Digestion

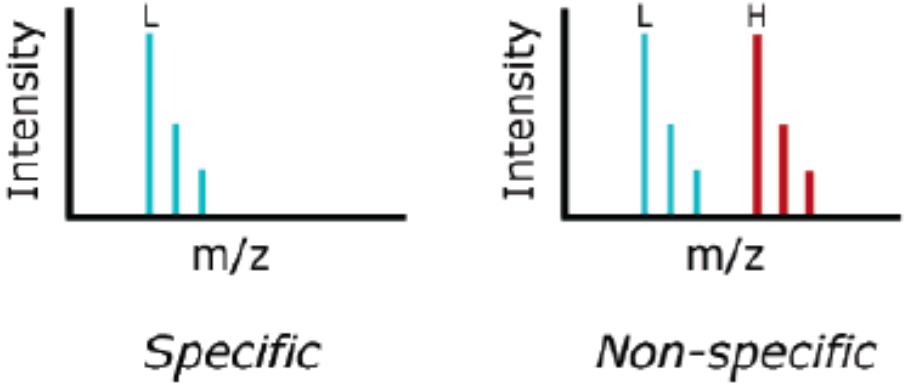
Mass spectrometry

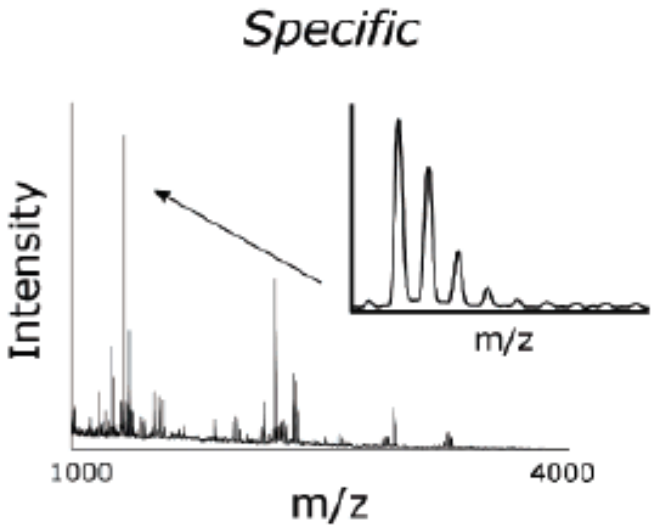
Protein Complexes – specific/non-specific binding

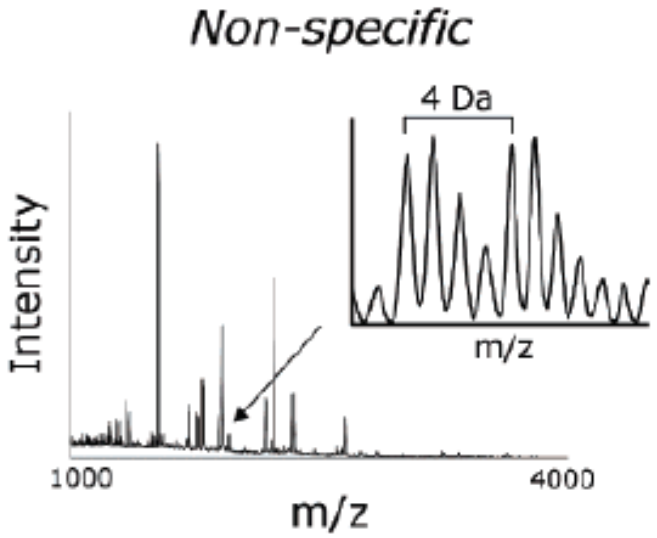


Immunoisolation

Mass Spectrometry

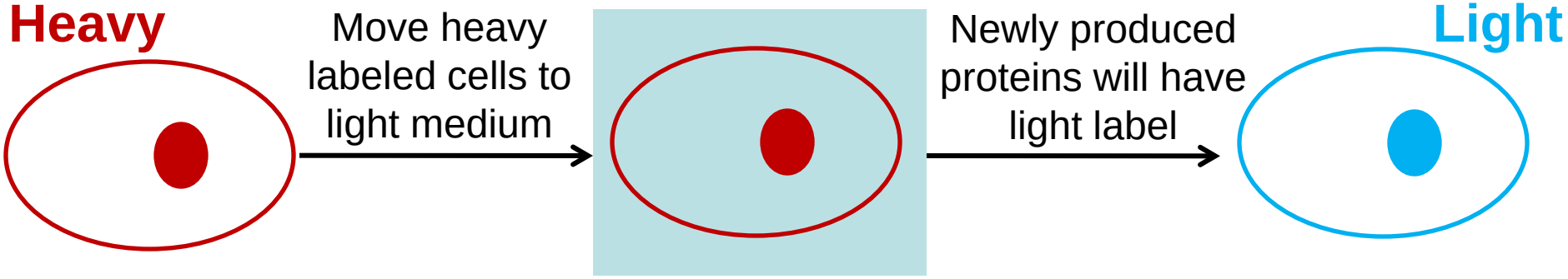






Tackett et al. JPR 2005

Protein Turnover



$$\frac{dC_j^H(t)}{dt} = -(K_C + K_T)C_j^H(t)$$

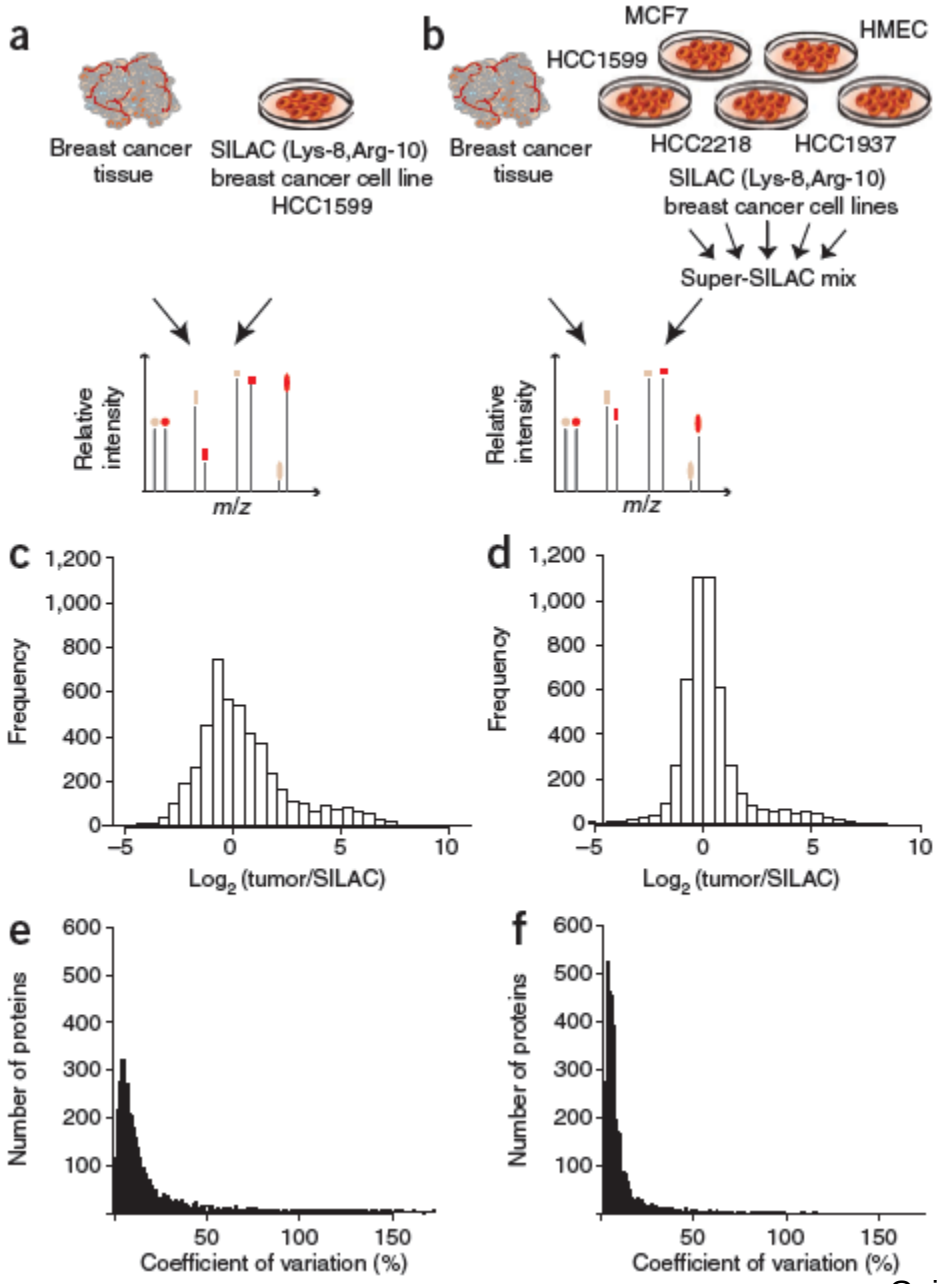
$$C_j^L(t) + C_j^H(t) = C_j^H(0)$$

$$\Rightarrow C_j^H(t) = C_j^H(0)e^{-(K_C+K_T)t}$$

$K_C=\log(2)/t_C$, t_C is the average time it takes for cells to go through the cell cycle, and $K_T=\log(2)/t_T$, t_T is the time it takes for half the proteins to turn over.

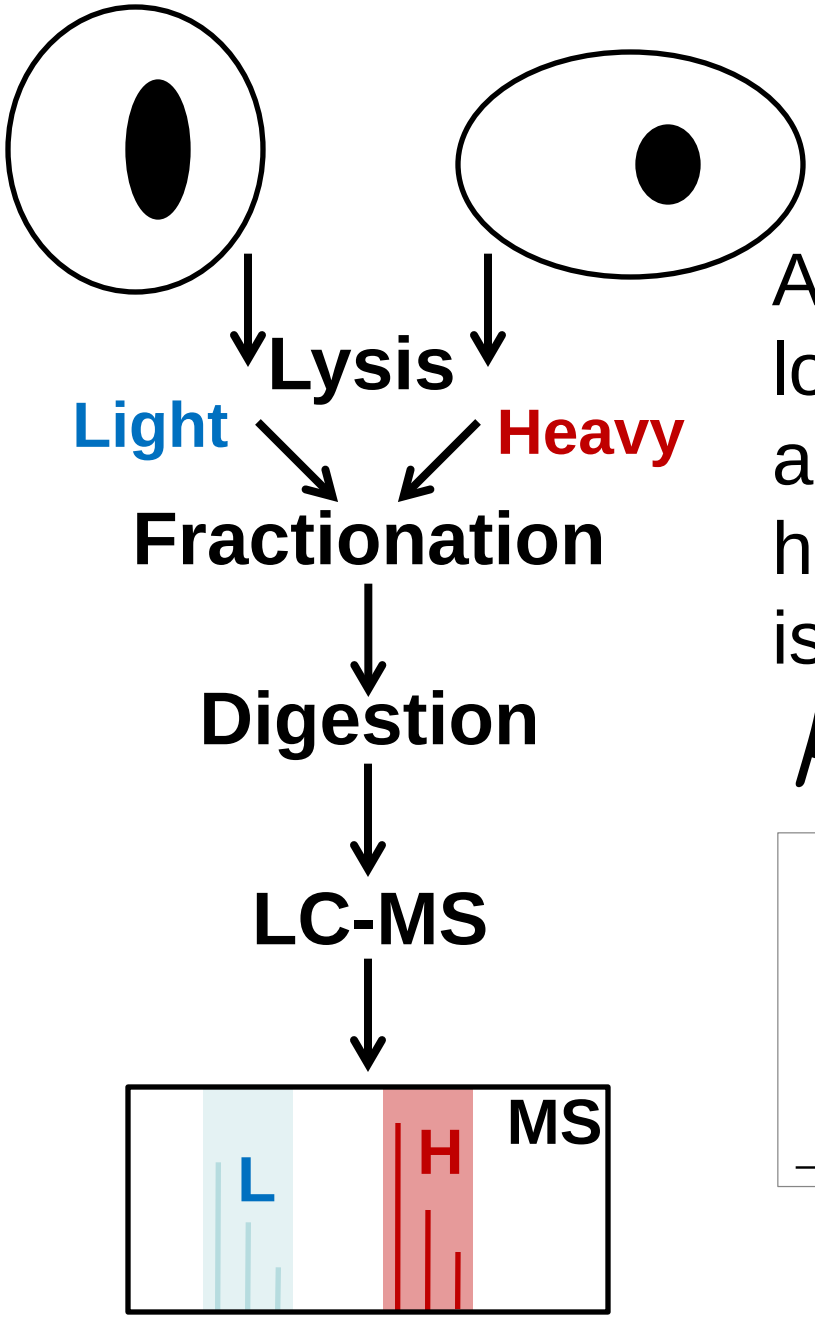
$$\log\left(\frac{I_j^H(t) + I_j^L(t)}{I_j^H(t)}\right) = t\left(\frac{1}{t_C} + \frac{1}{t_T}\right)\log(2)$$

Super-SILAC



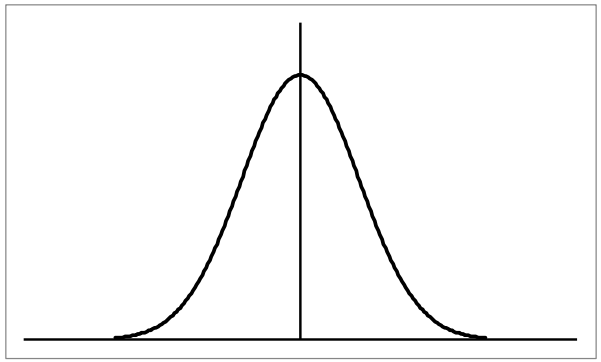
Geiger et al., Nature Methods 2010

Quantitation – Protein Labeling



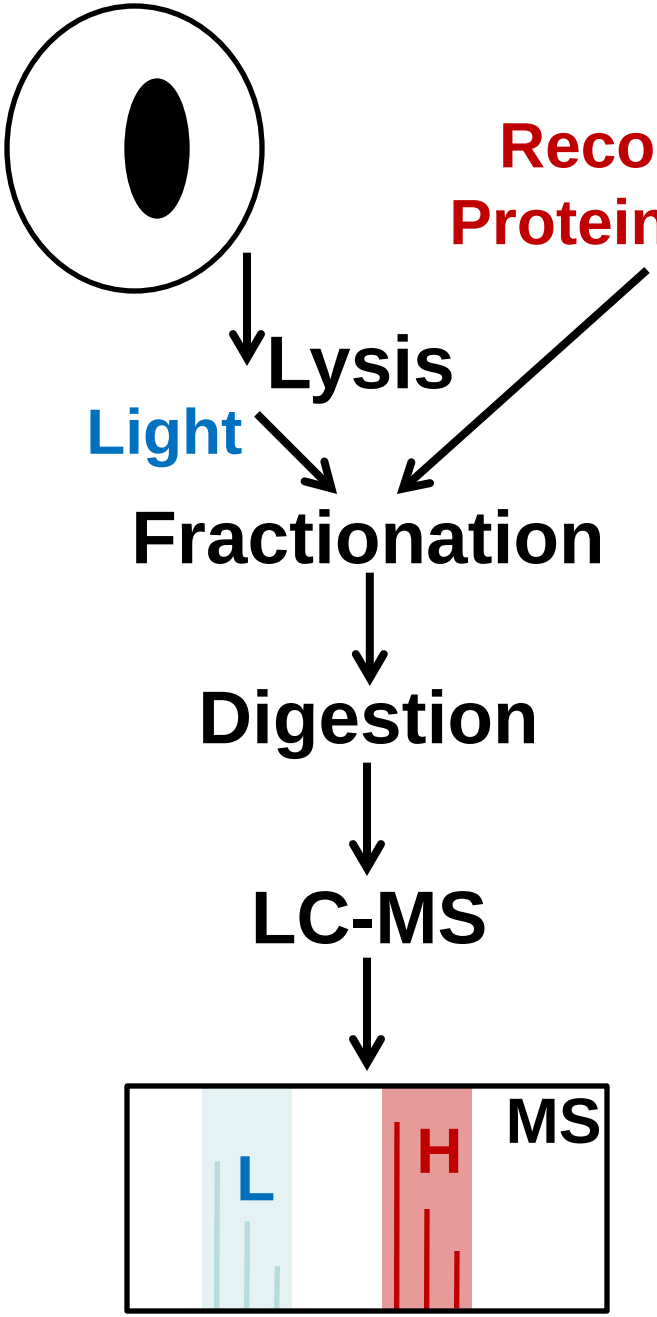
Assumption: All losses after mixing are identical for the heavy and light isotopes and

$$p^{L}_{i_n j}\, p^{M}_{i_n j} \approx p^{L}_{i_m j}\, p^{M}_{i_m j}$$



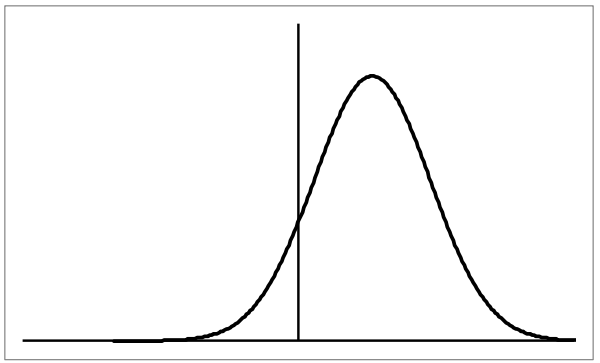
Gygi et al. Nature Biotech 17 (1999) 994

Quantitation – Labeled Proteins

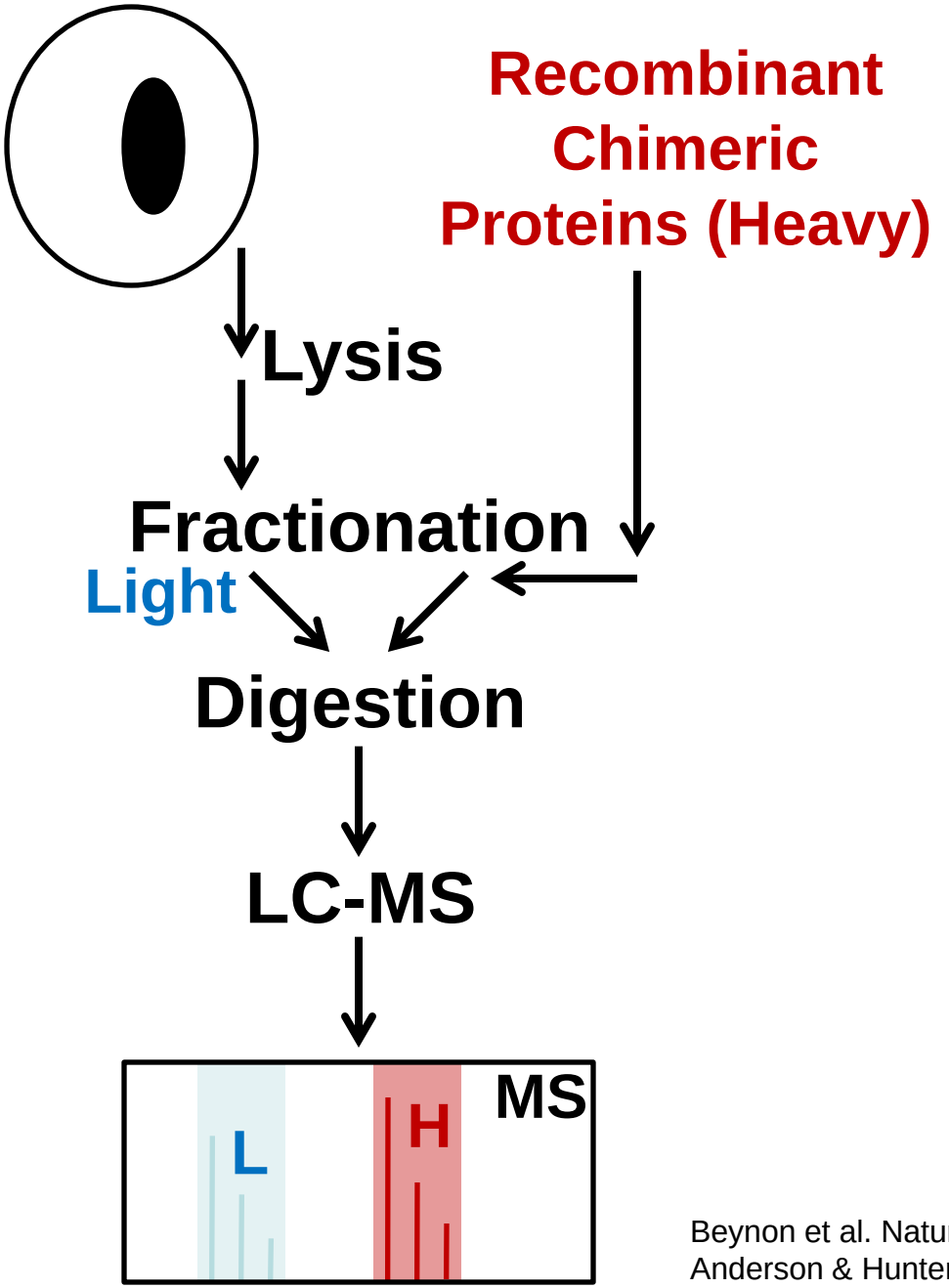


Assumption: All losses after mixing are identical for the heavy and light isotopes and

$$p^{L}_{i_n j}\, p^{M}_{i_n j} \approx p^{M}_{i_m j}$$

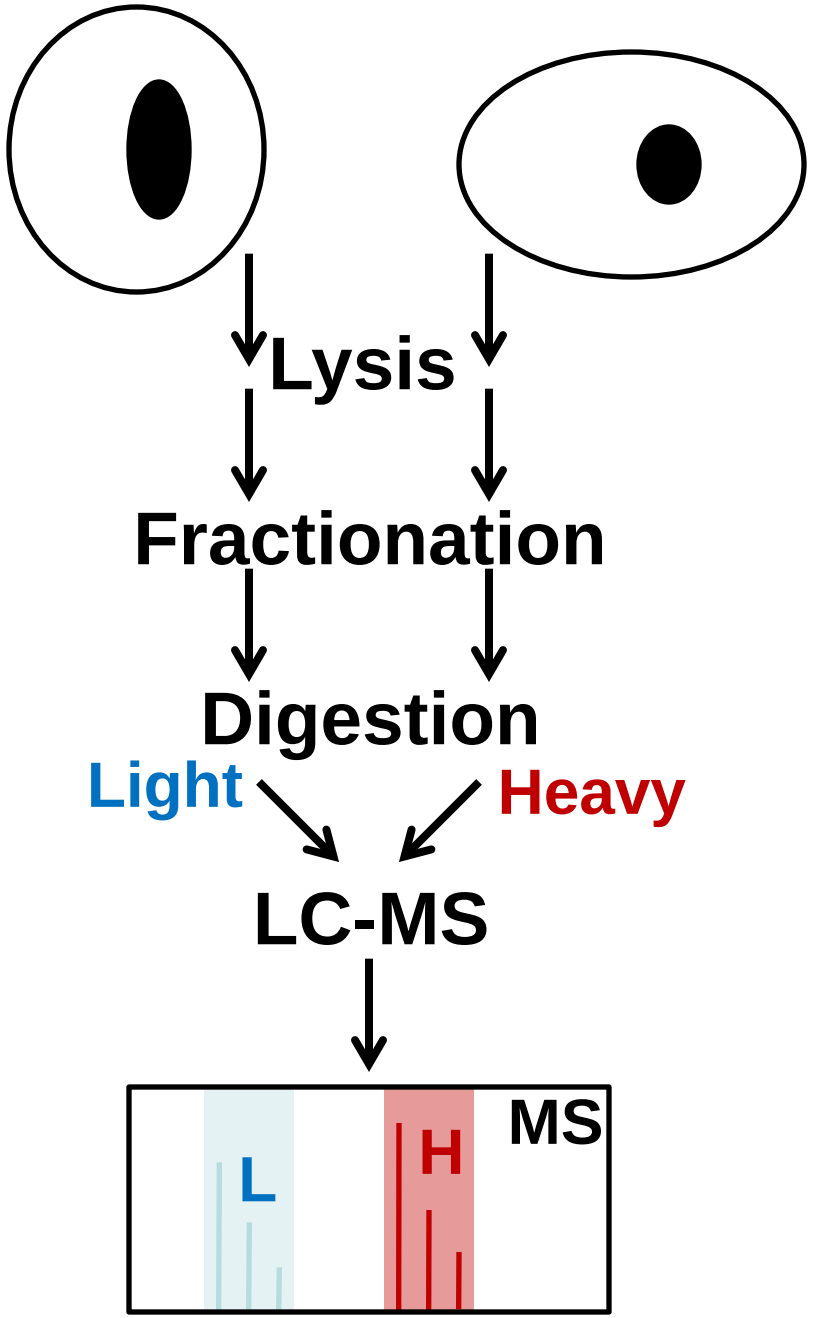


Quantitation – Labeled Chimeric Proteins



Beynon et al. Nature Methods 2 (2005) 587
Anderson & Hunter MCP 5 (2006) 573

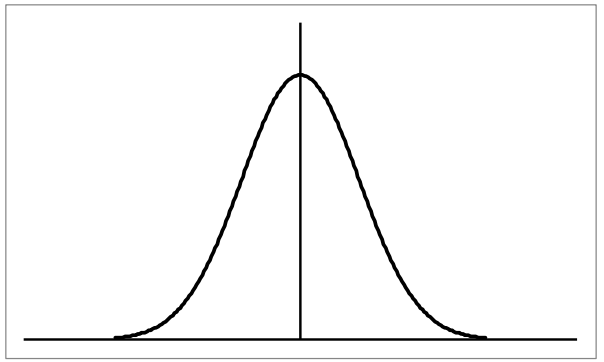
Quantitation – Peptide Labeling



Assumption: All losses after mixing are identical for the heavy and light isotopes and

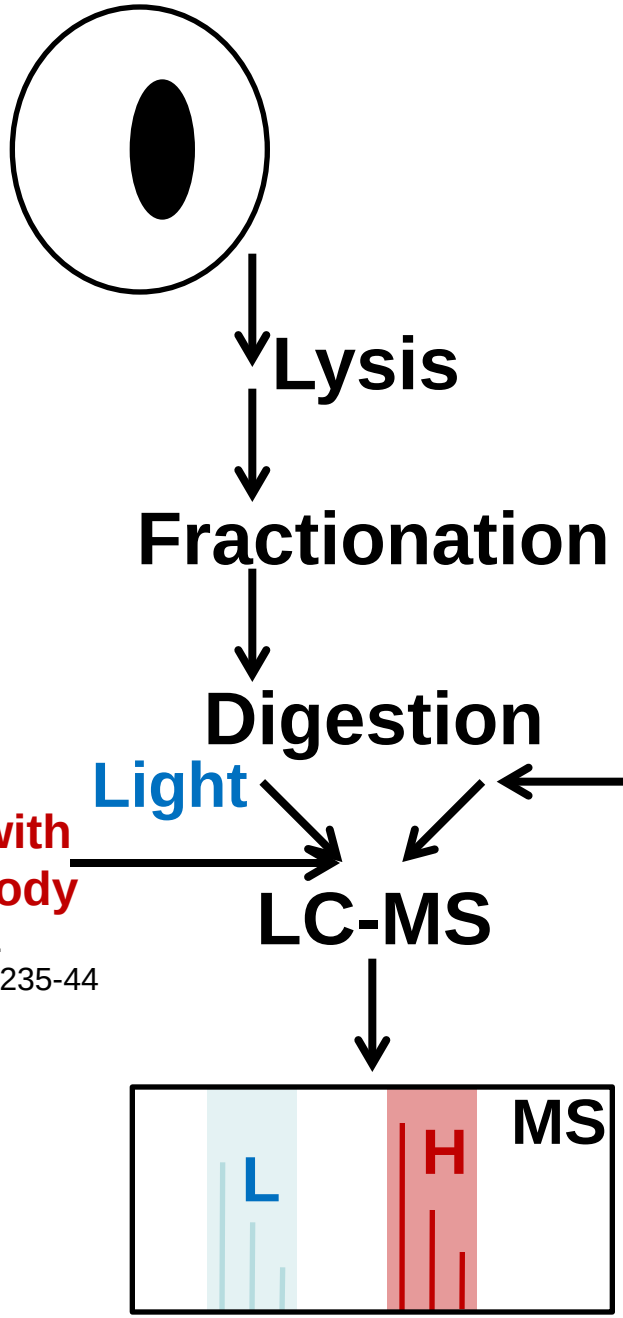
$$p_{i_n}^{L}{}_{j}\, p_{i_n}^{\mathrm{Pr}}{}_{j}\, p_{i_n}^{D}{}_{jk}\, p_{i_n}^{M}{}_{k} \approx$$

$$\approx p_{i_m}^{L}{}_{j}\, p_{i_m}^{\mathrm{Pr}}{}_{j}\, p_{i_m}^{D}{}_{jk}\, p_{i_m}^{M}{}_{k}$$



Gygi et al. Nature Biotech 17 (1999) 994
Mirgorodskaya et al. RCMS 14 (2000) 1226

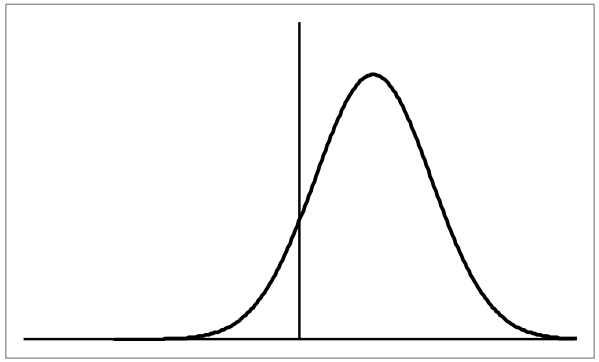
Quantitation – Labeled Synthetic Peptides



Anderson, N.L., et al.
Proteomics 3 (2004) 235-44

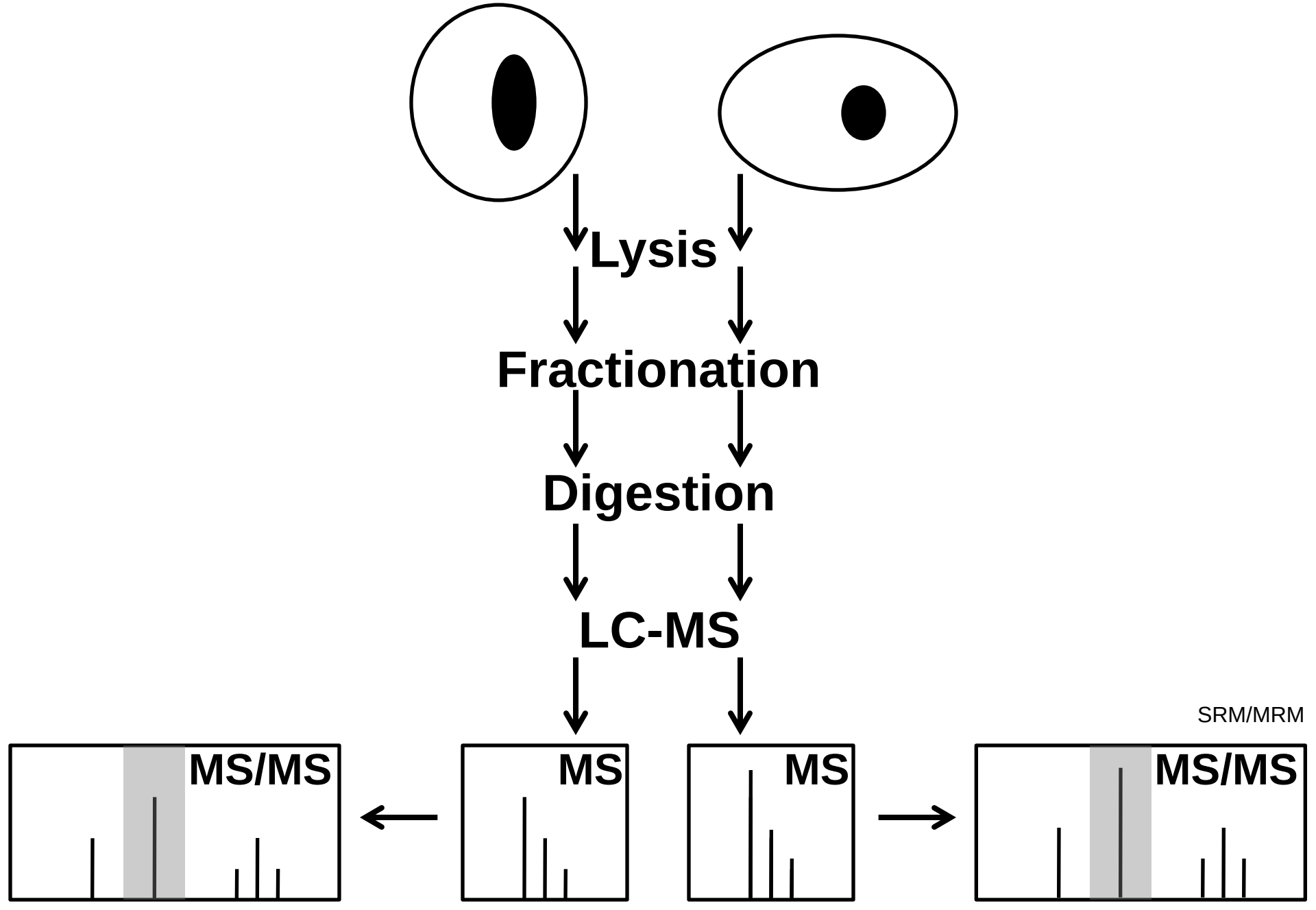
Assumption: All losses after mixing are identical for the heavy and light isotopes and

$$p_{i_n}^{L\,j}\, p_{i_n}^{\mathrm{Pr}\,j}\, p_{i_n}^{D\,jk}\, p_{i_n}^{M\,k} \approx p_{sk}^{M}$$

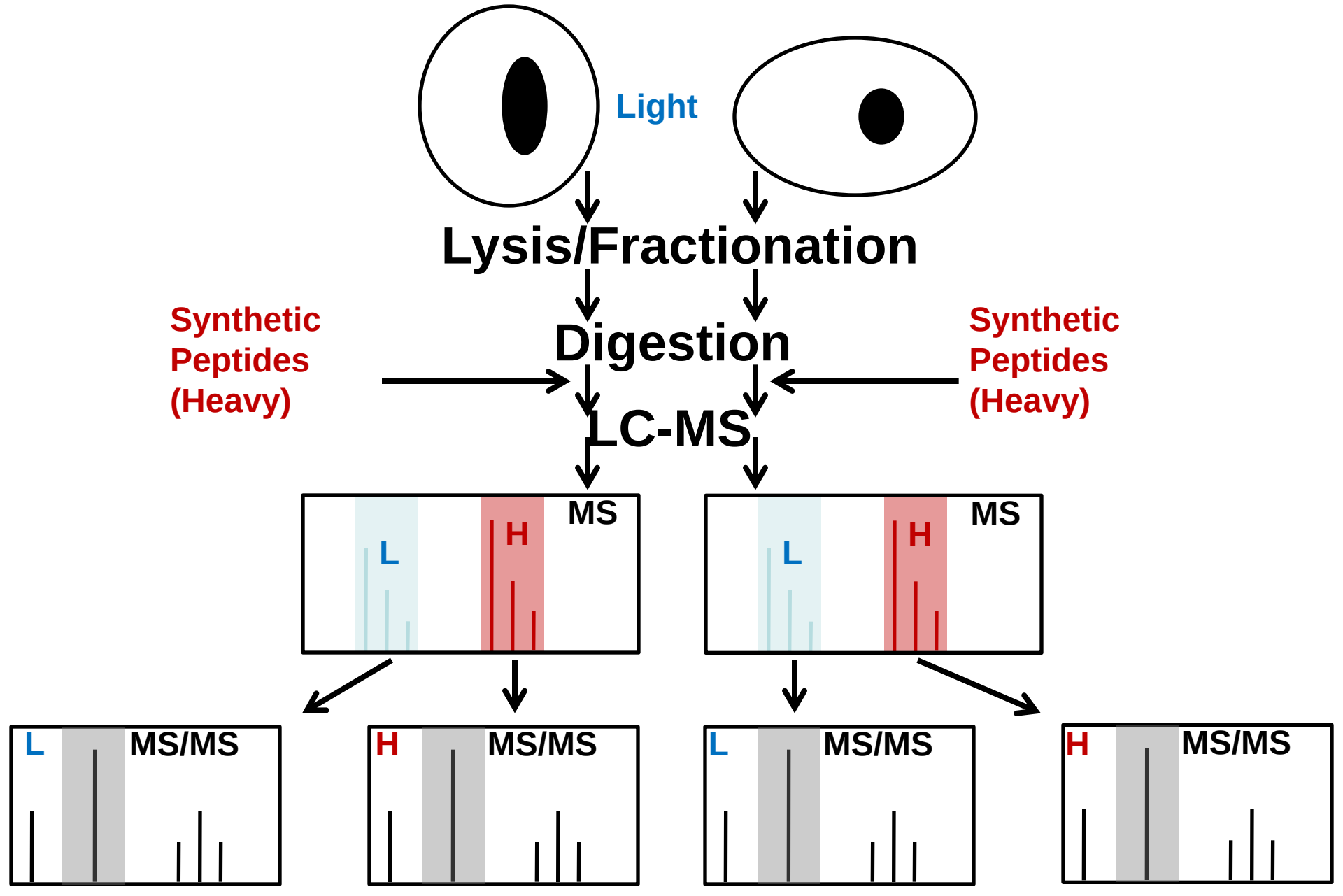


Gerber et al. PNAS 100 (2003) 6940

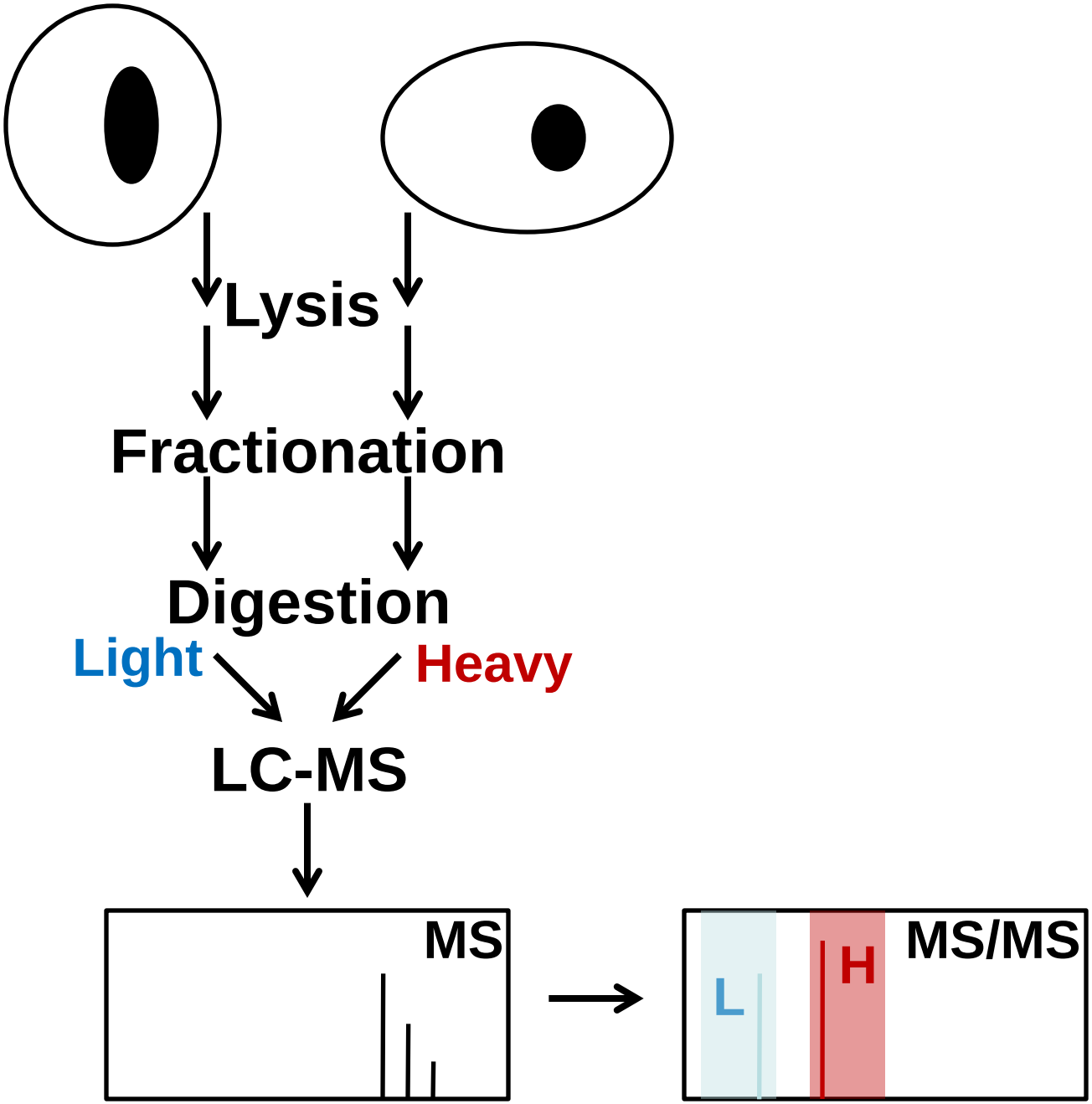
Quantitation – Label-Free (MS/MS)



Quantitation – Labeled Synthetic Peptides

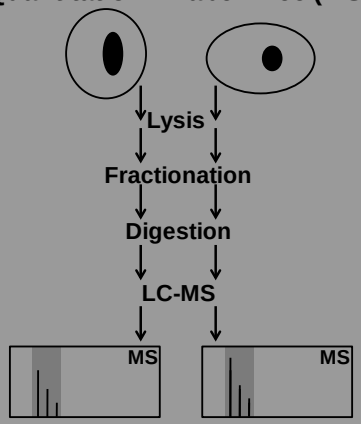


Quantitation – Isobaric Peptide Labeling

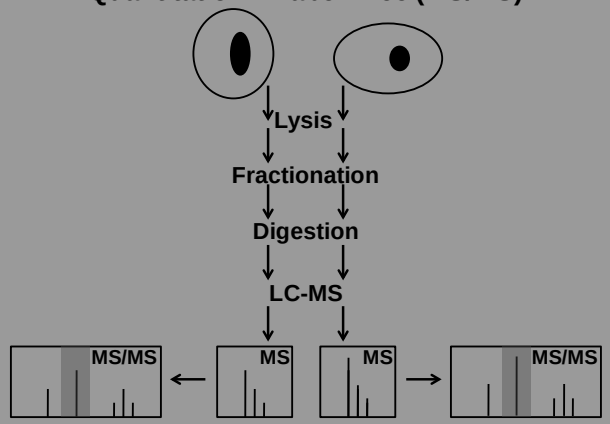


Ross et al. MCP 3 (2004) 1154

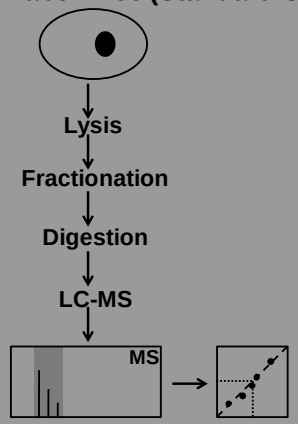
Quantitation – Label-Free (MS)



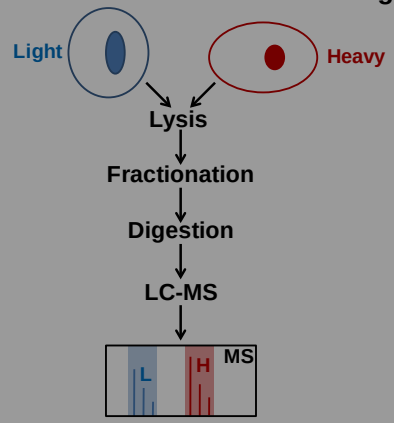
Quantitation – Label-Free (MS/MS)



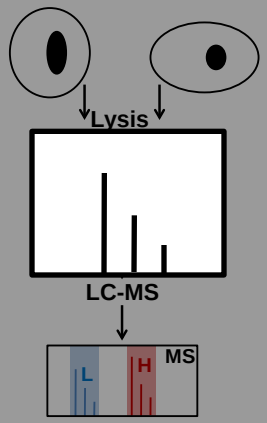
Quantitation – Label-Free (Standard Curve)



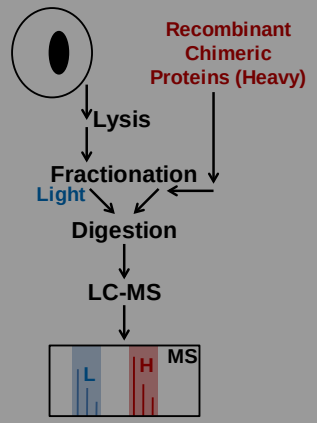
Quantitation – Metabolic Labeling



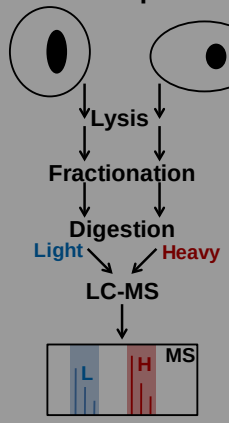
Quantitation – Protein Labeling



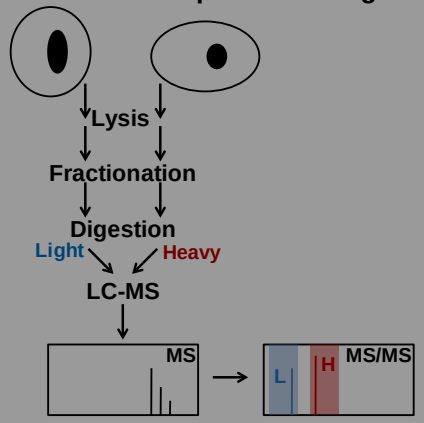
Quantitation – Labeled Chimeric Proteins



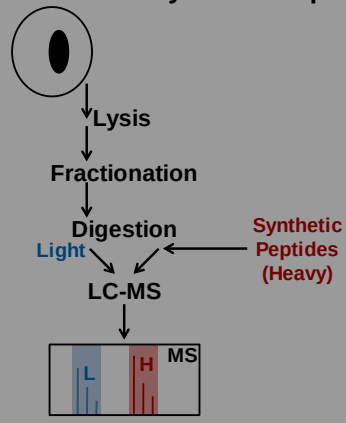
Quantitation – Peptide Labeling



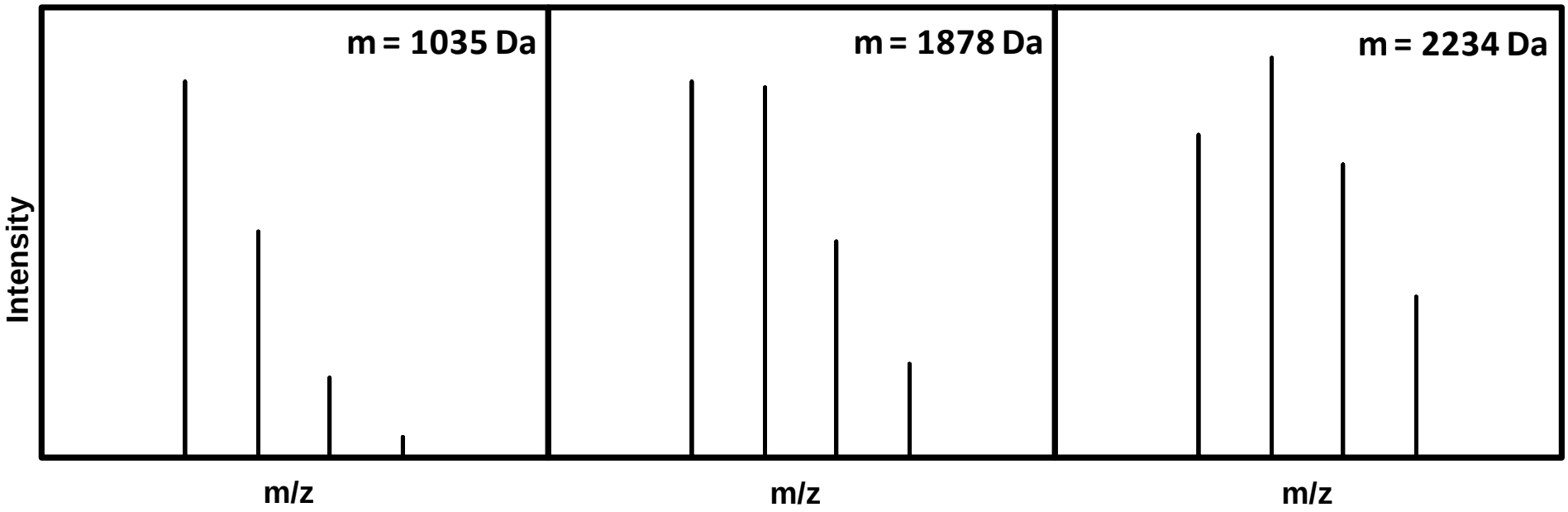
Quantitation – Isobaric Peptide Labeling



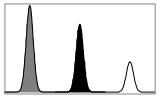
Quantitation – Labeled Synthetic Peptides

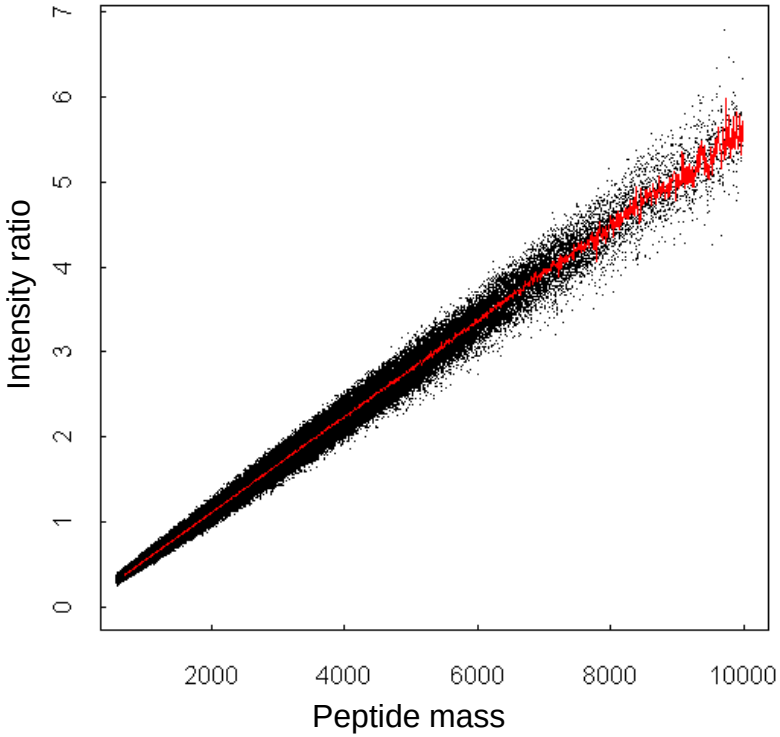


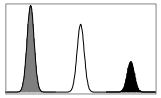
Isotope distributions

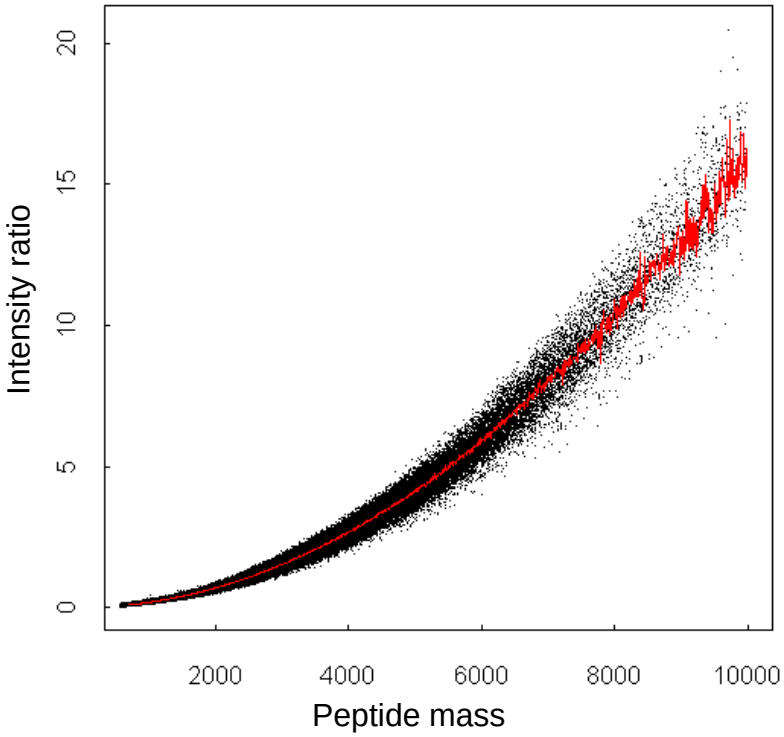


Isotope distributions

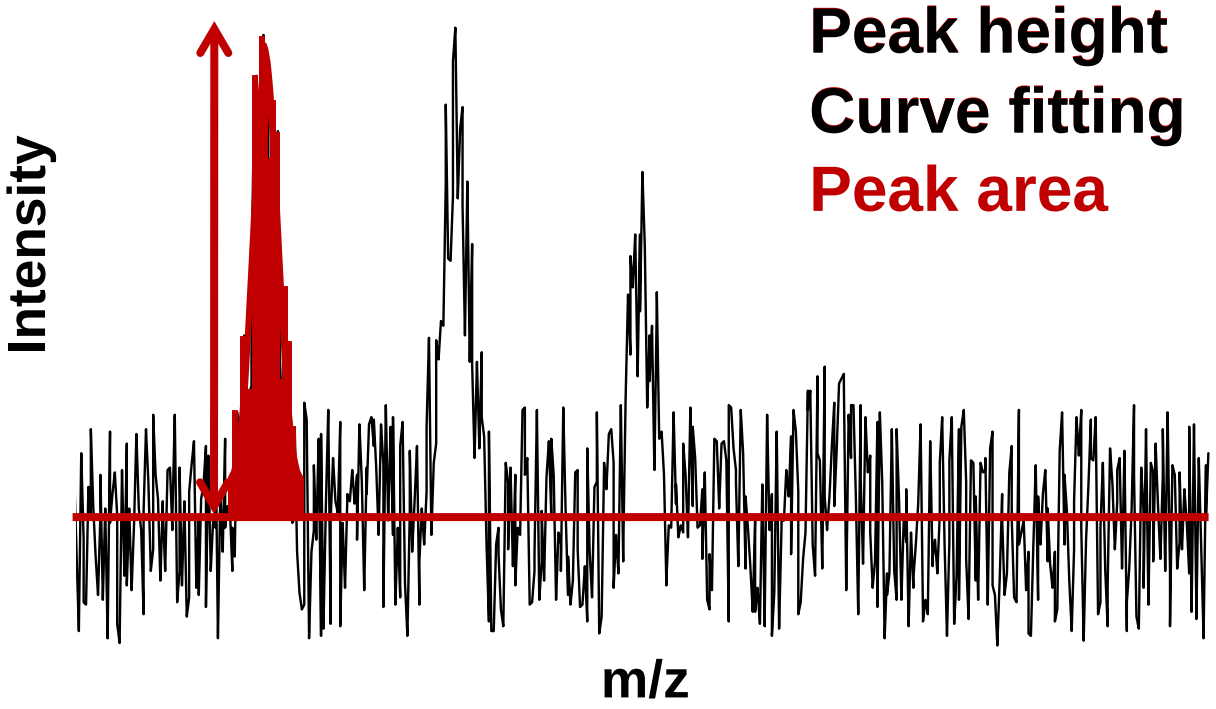




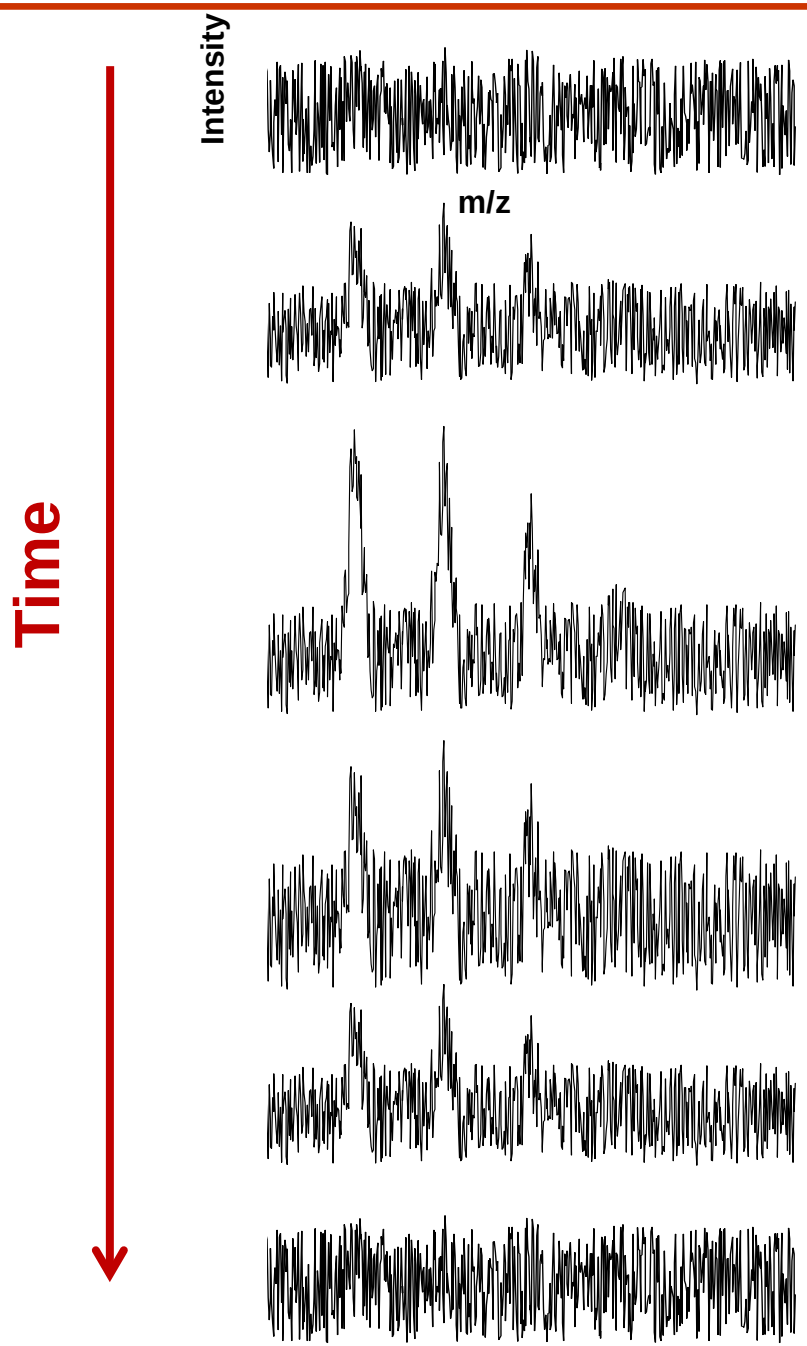


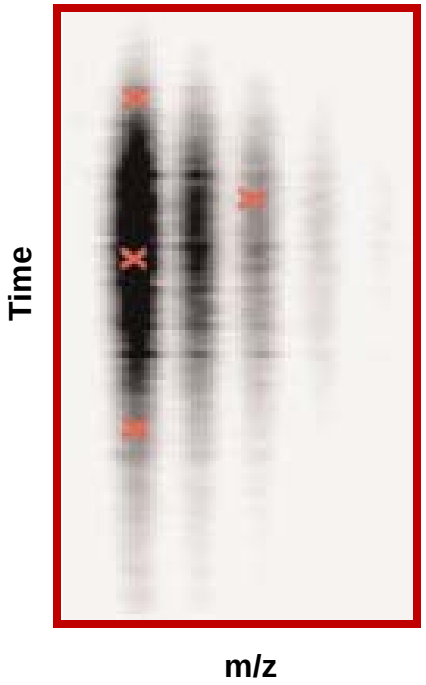


Estimating peptide quantity

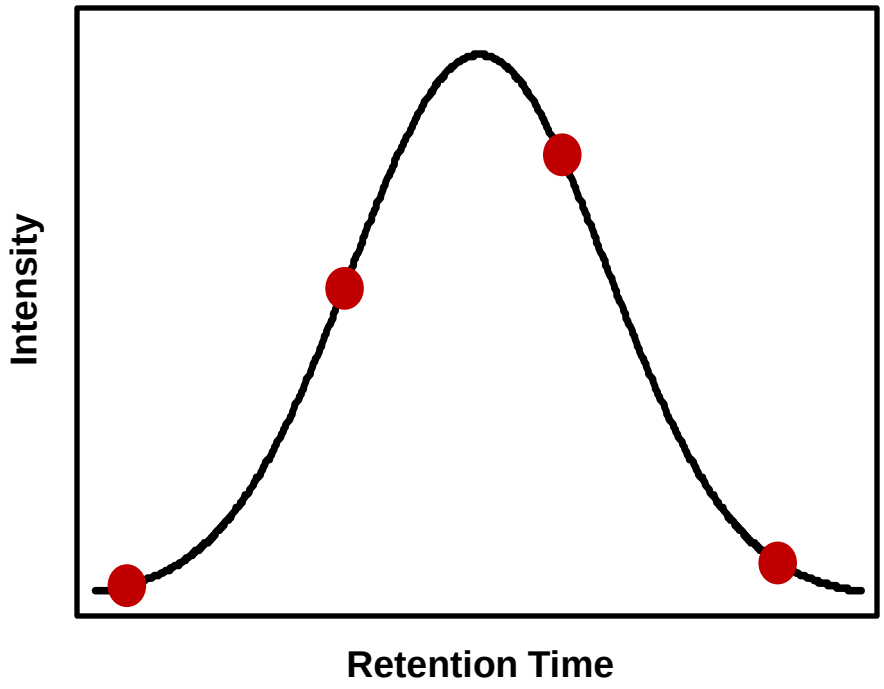


Time dimension

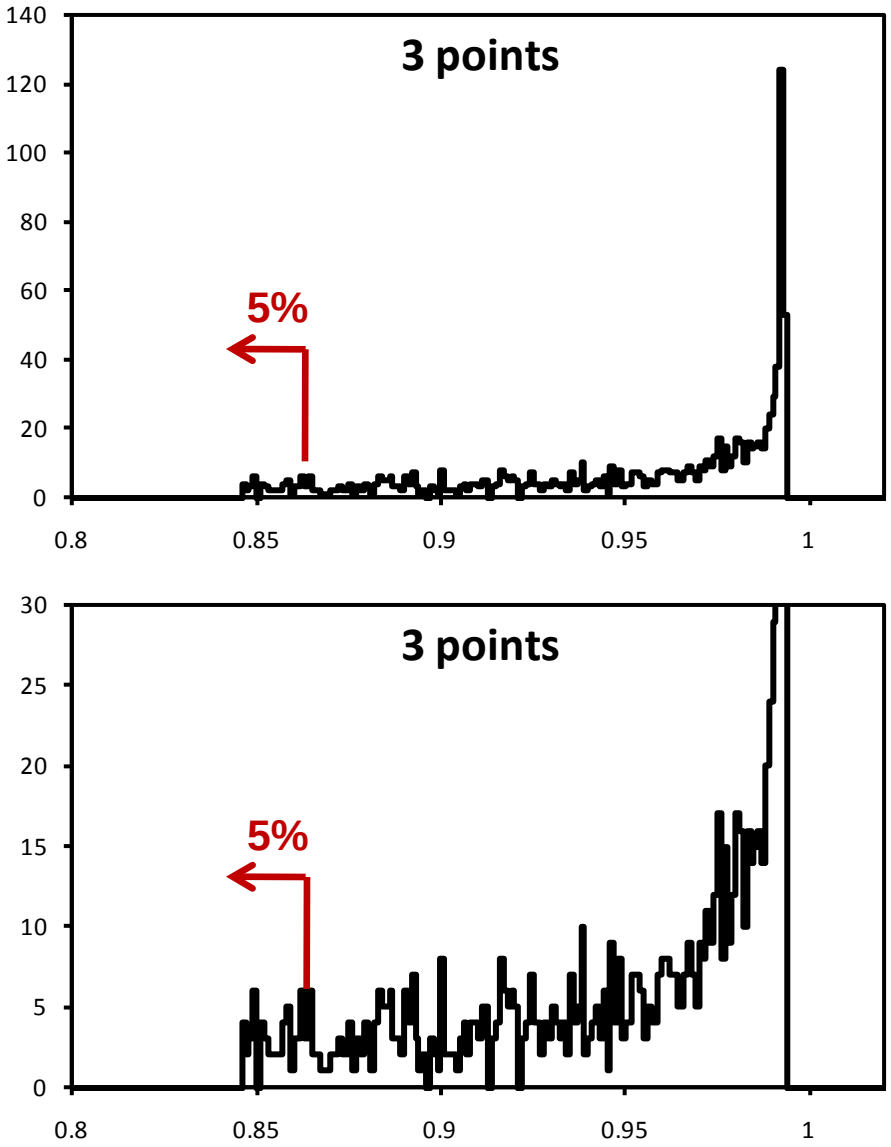




Sampling

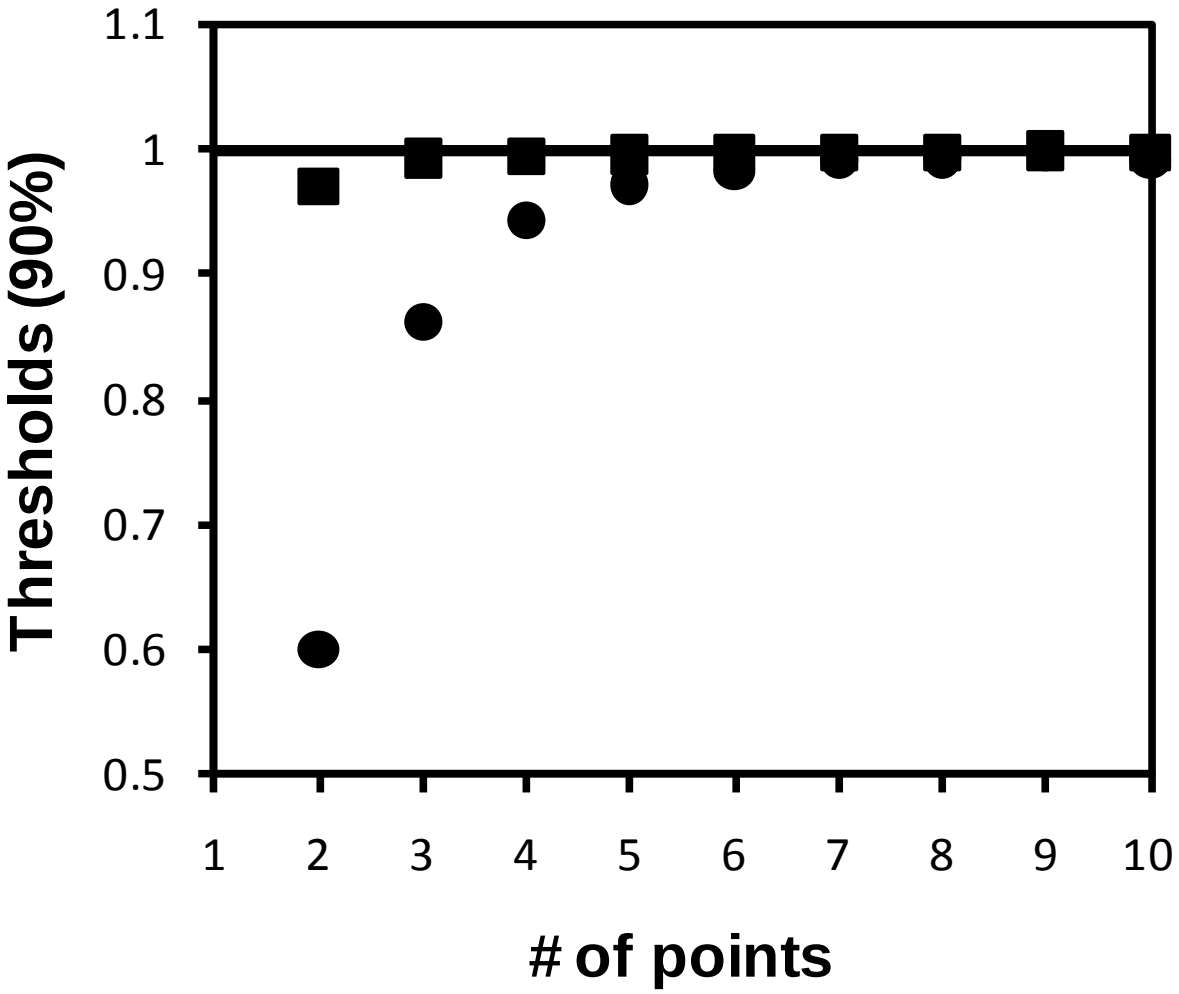


Sampling

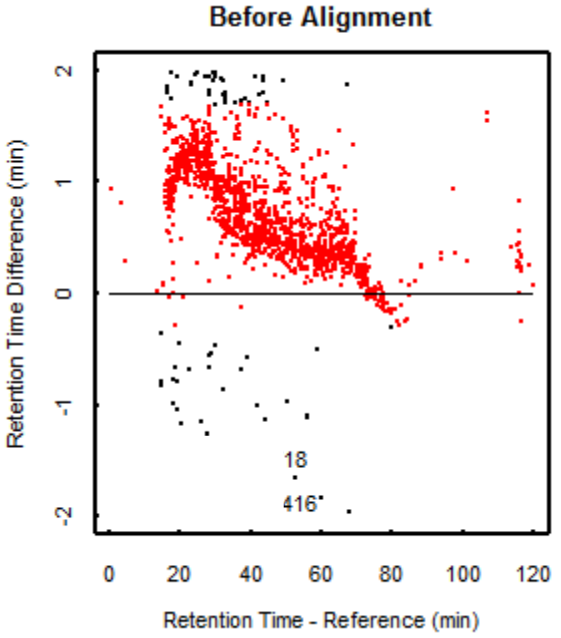


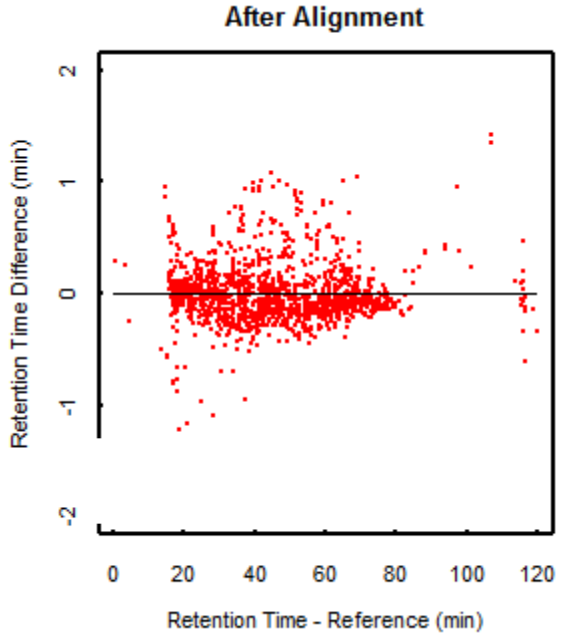
Acquisition time = 0.05σ

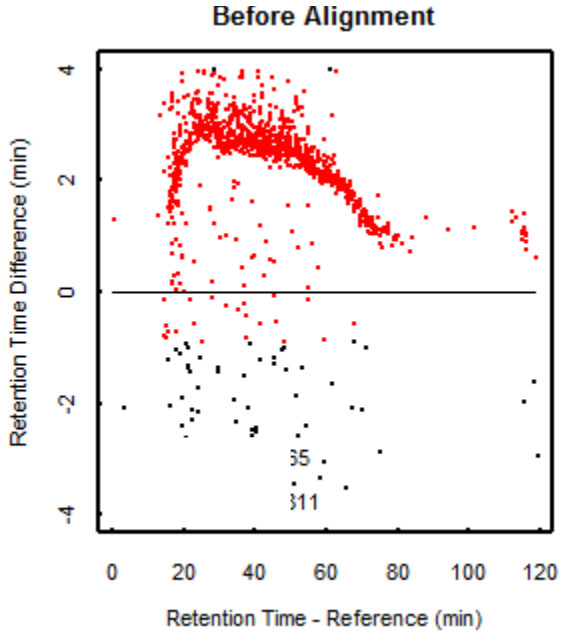
Sampling

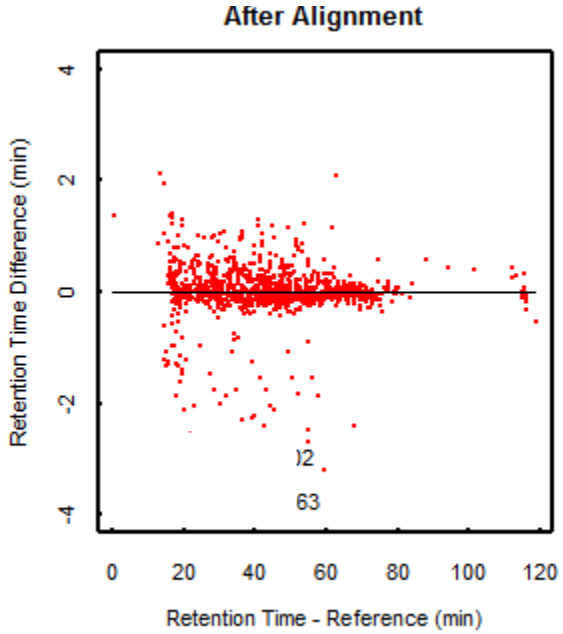


Retention Time Alignment

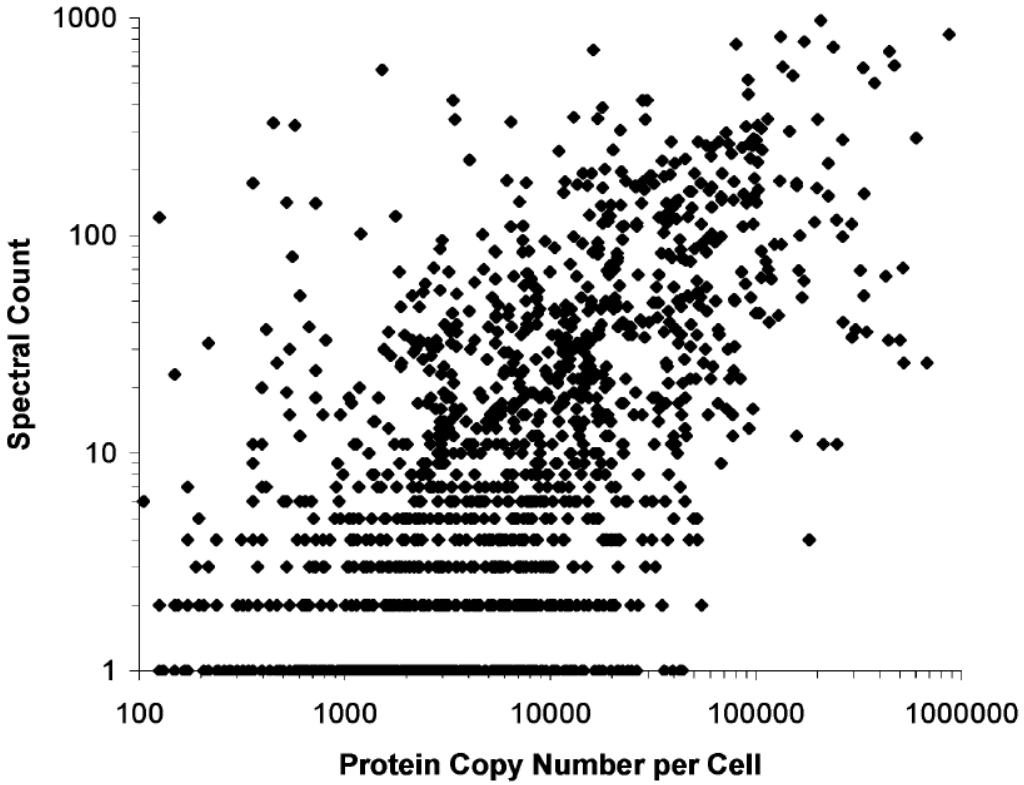


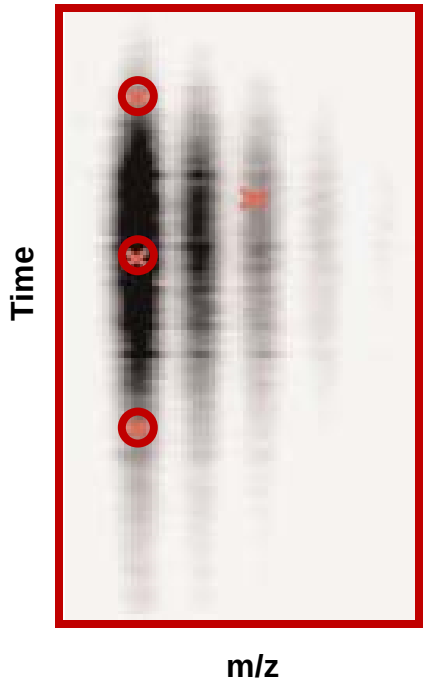






Estimating peptide quantity by spectrum counting



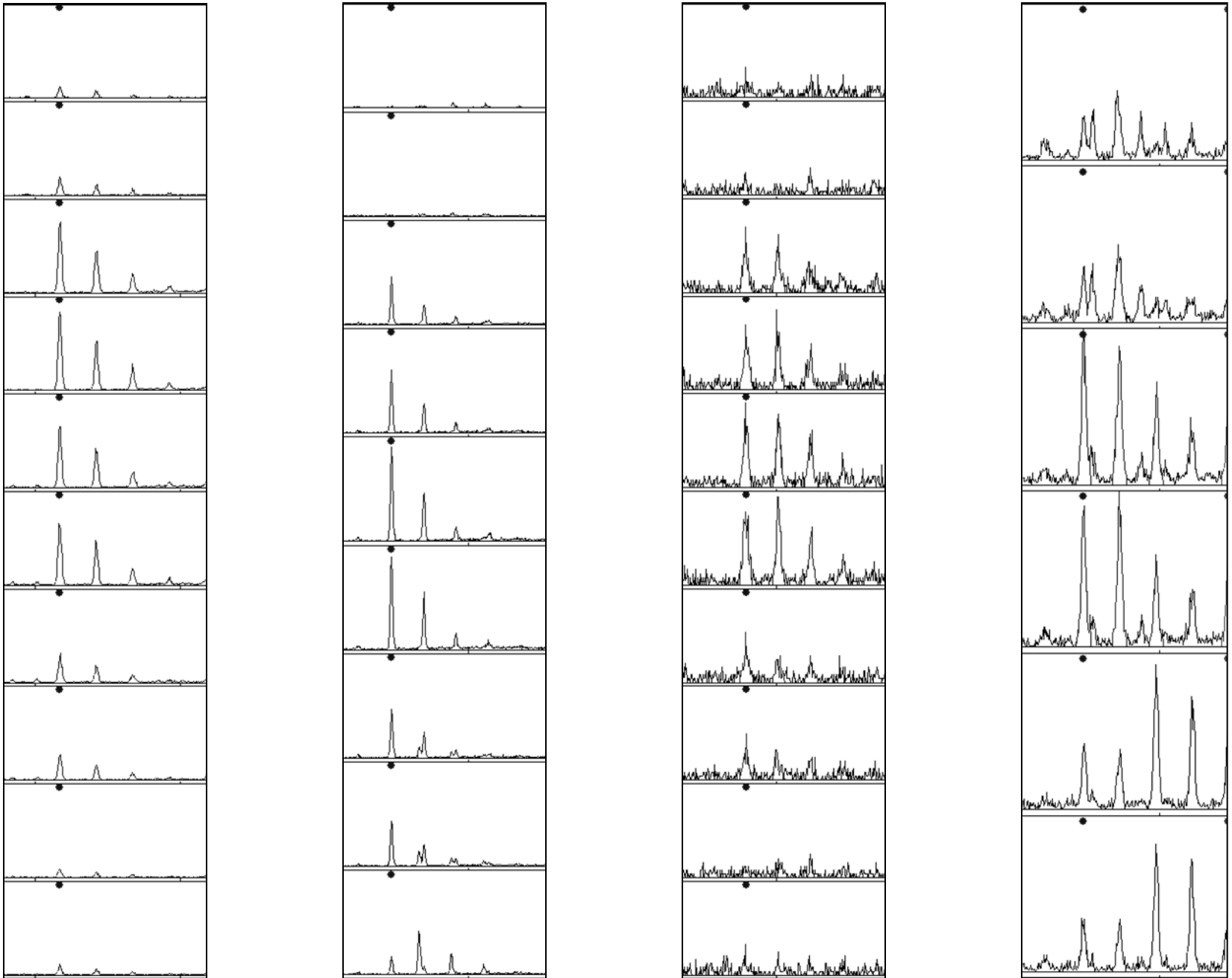


Liu et al., Anal. Chem. 2004, 76, 4193

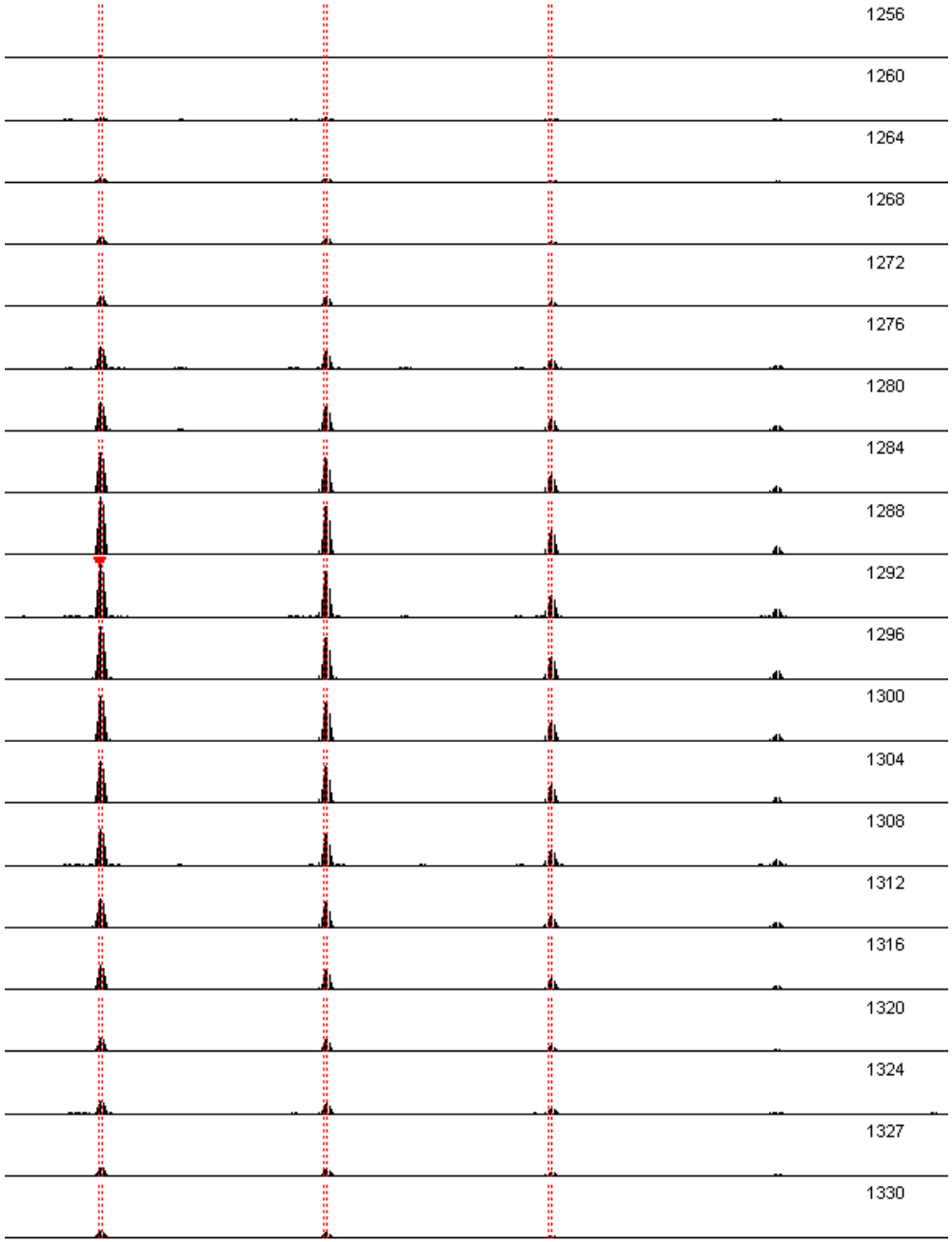
What is the best way to estimate quantity?

Peak height	**- resistant to interference** **- poor statistics**
Peak area	**- better statistics** **- more sensitive to interference**
Curve fitting	**- better statistics** **- needs to know the peak shape** **- slow**
Spectrum counting	**- resistant to interference** **- easy to implement** **- poor statistics for low-abundance proteins**

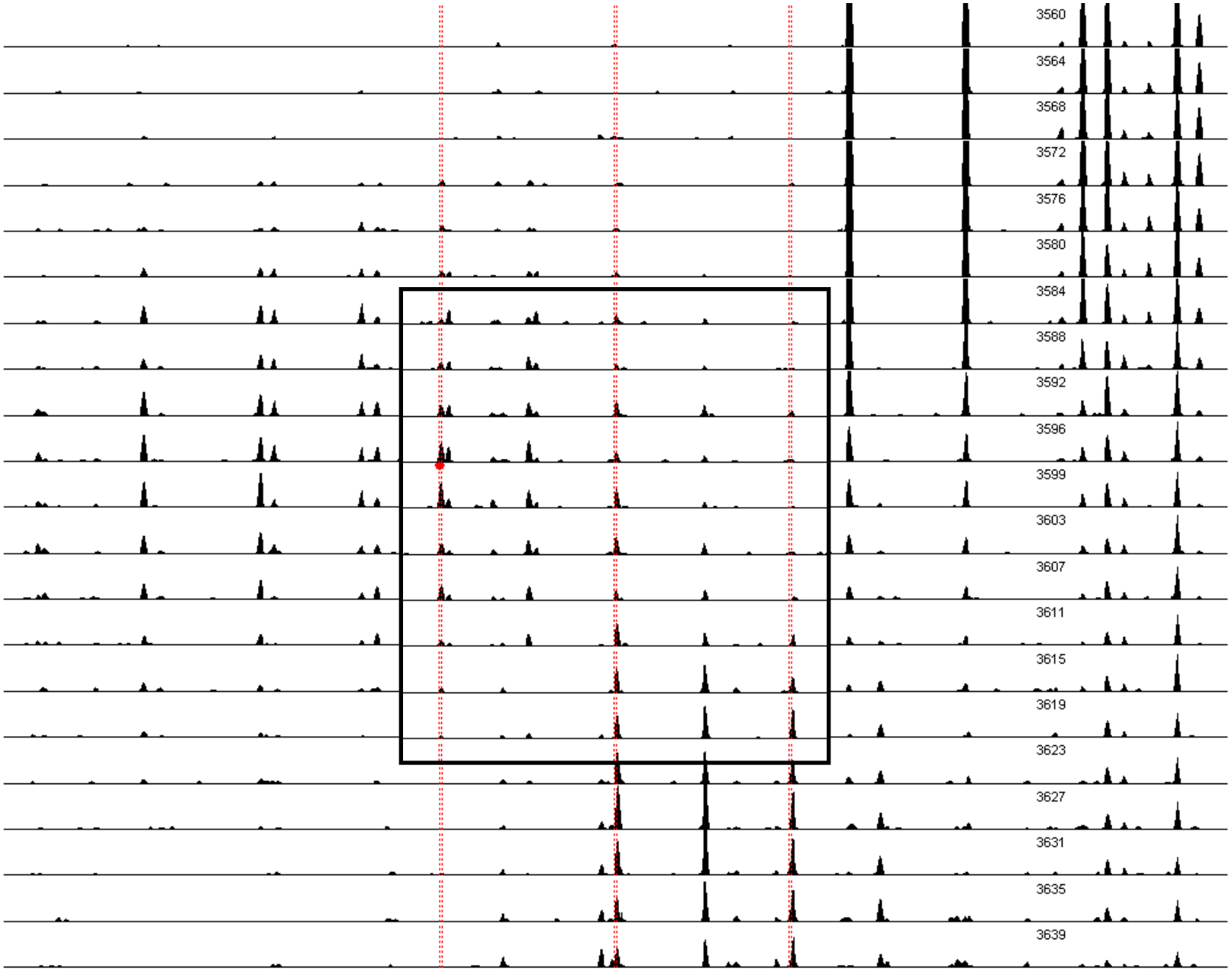
Examples - qTOF



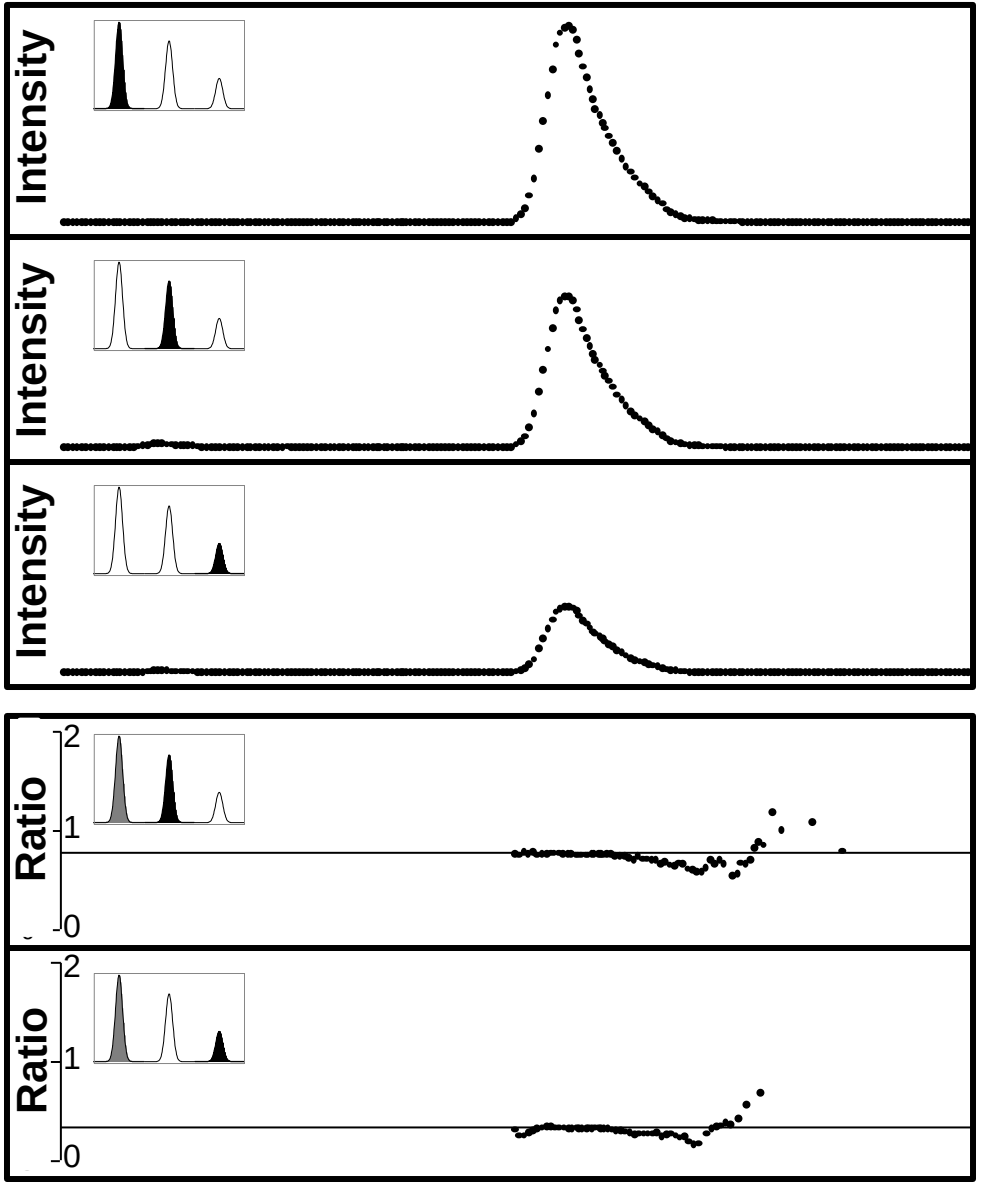
Examples - Orbitrap



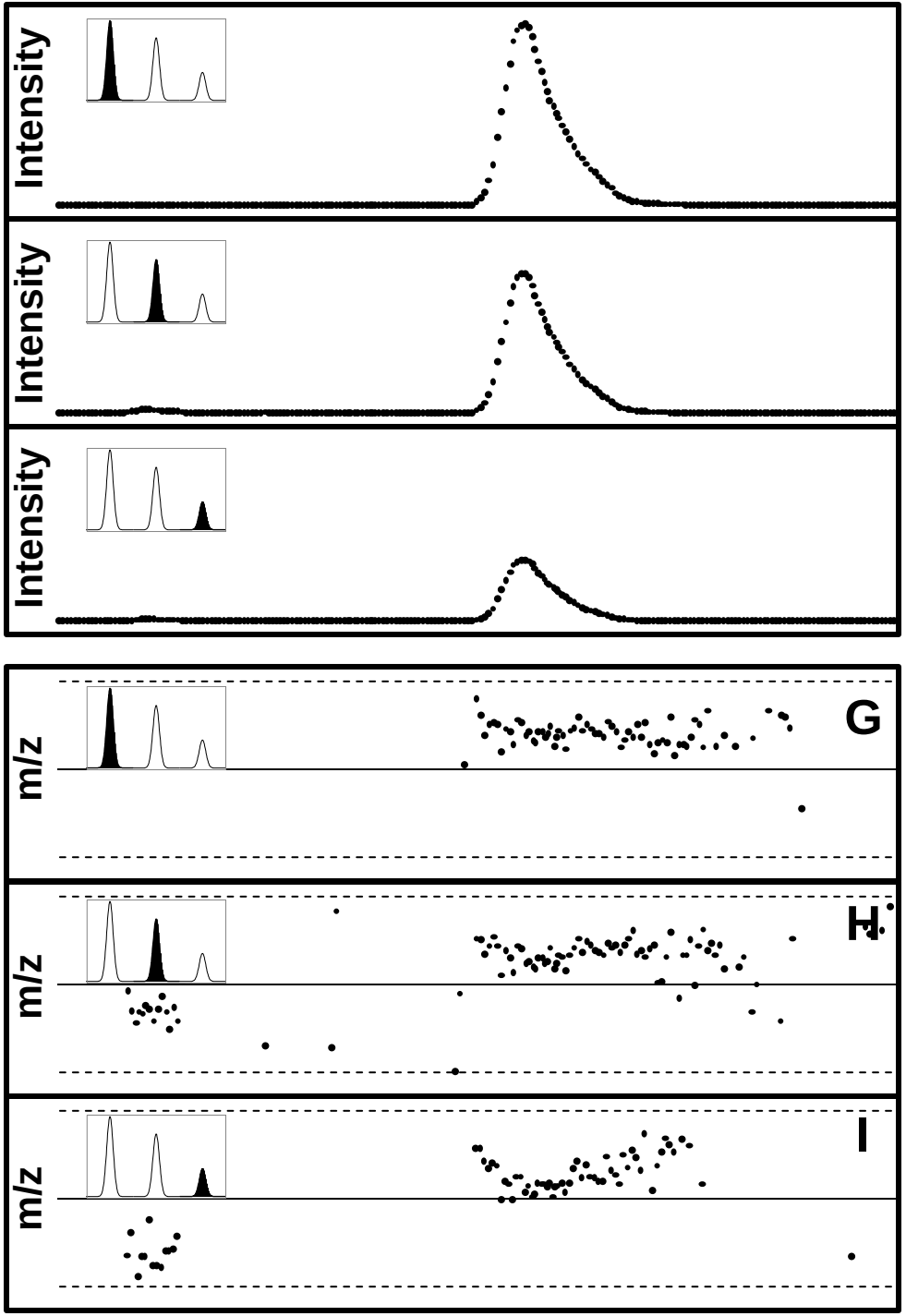
Examples - Orbitrap



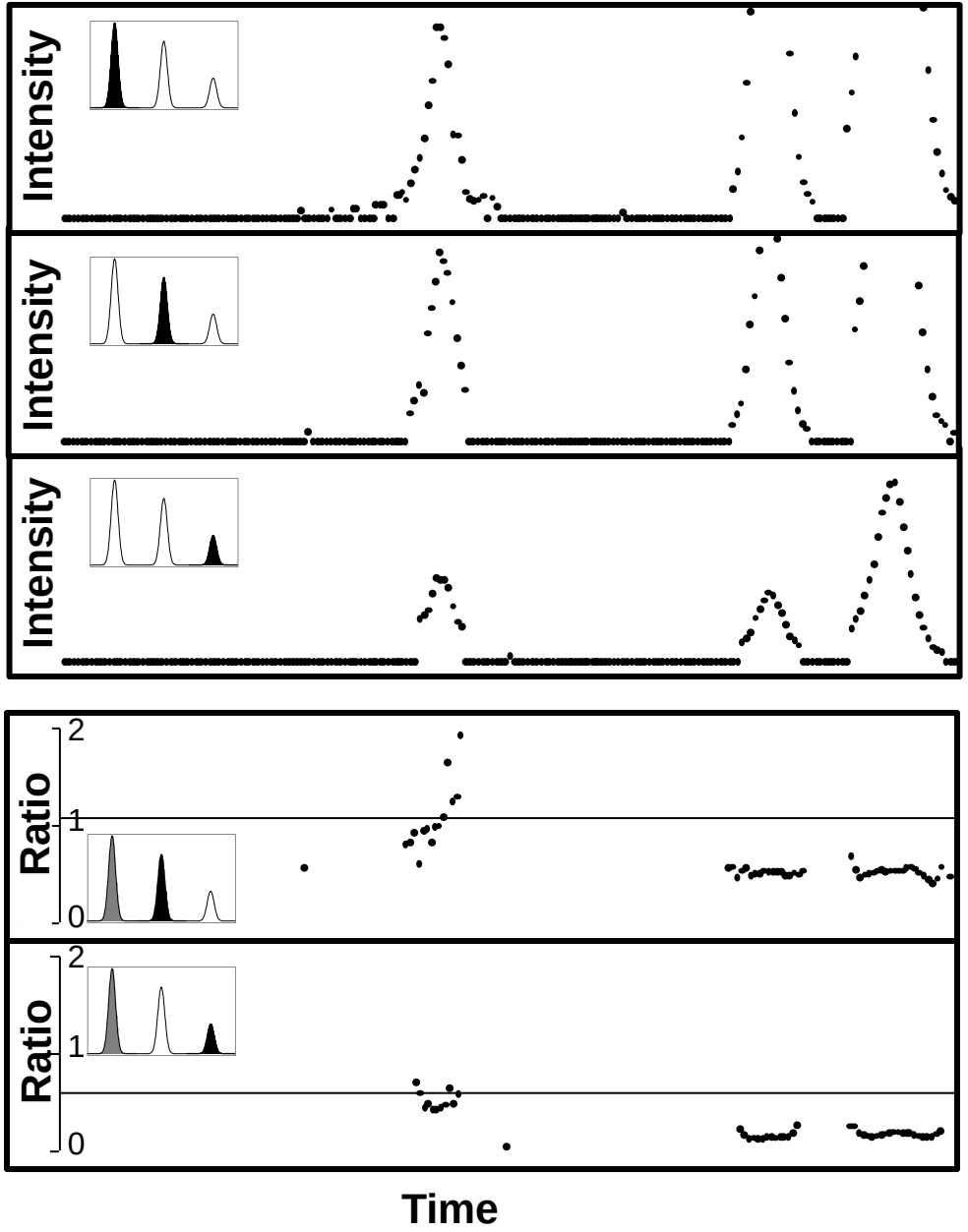
AADDTWEPFASGK



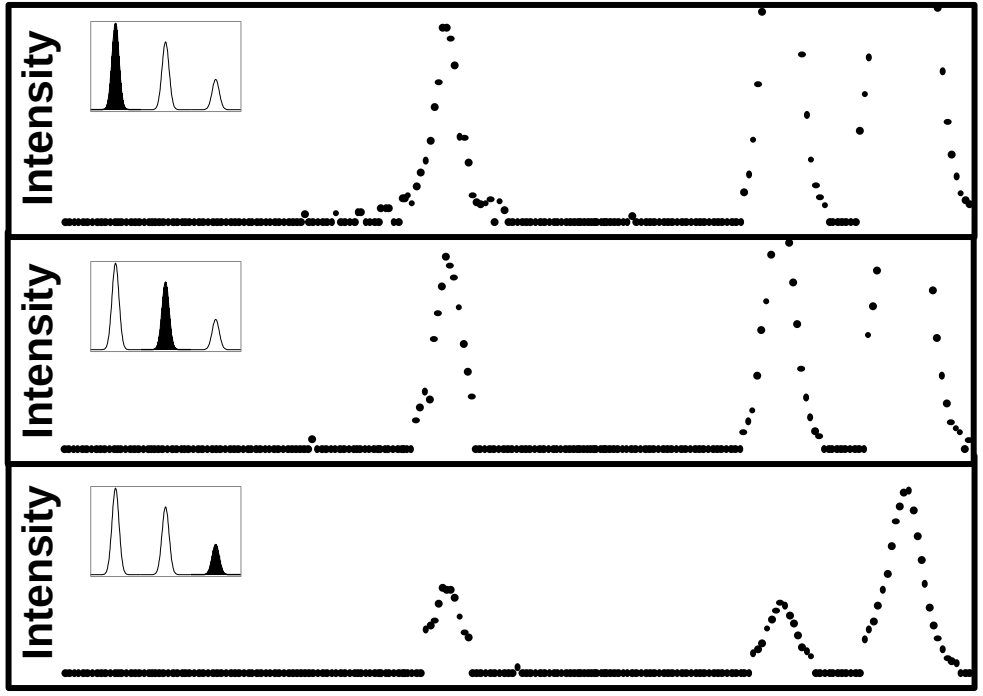
AADDTWEPFASGK

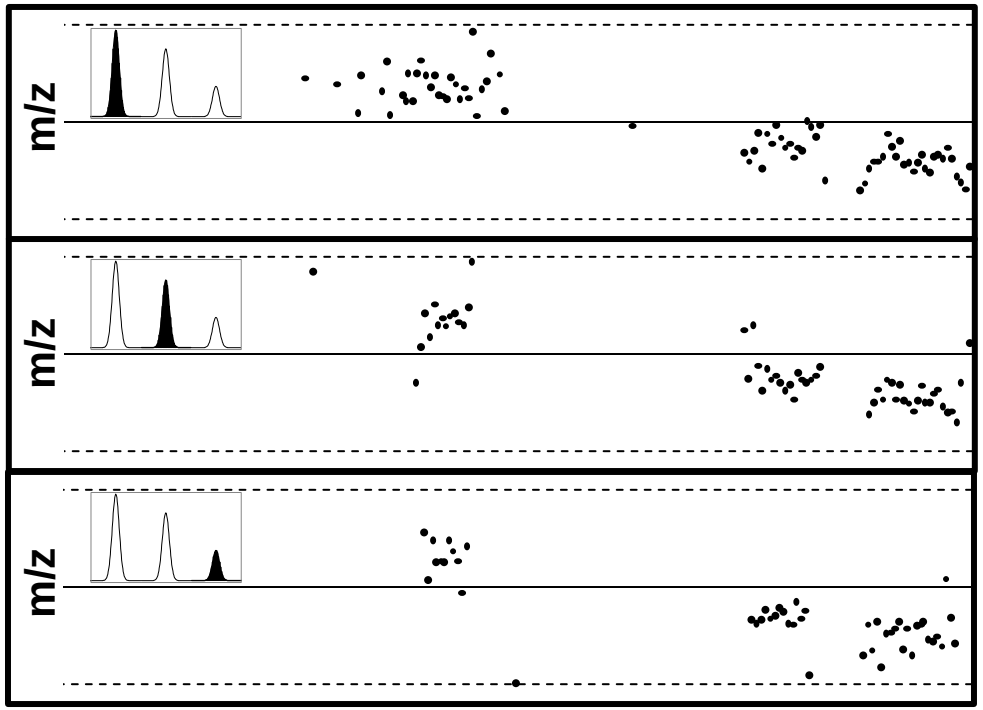


YVLTQPPSVSVAPGQTAR



YVLTQPPSVSVAPGQTAR





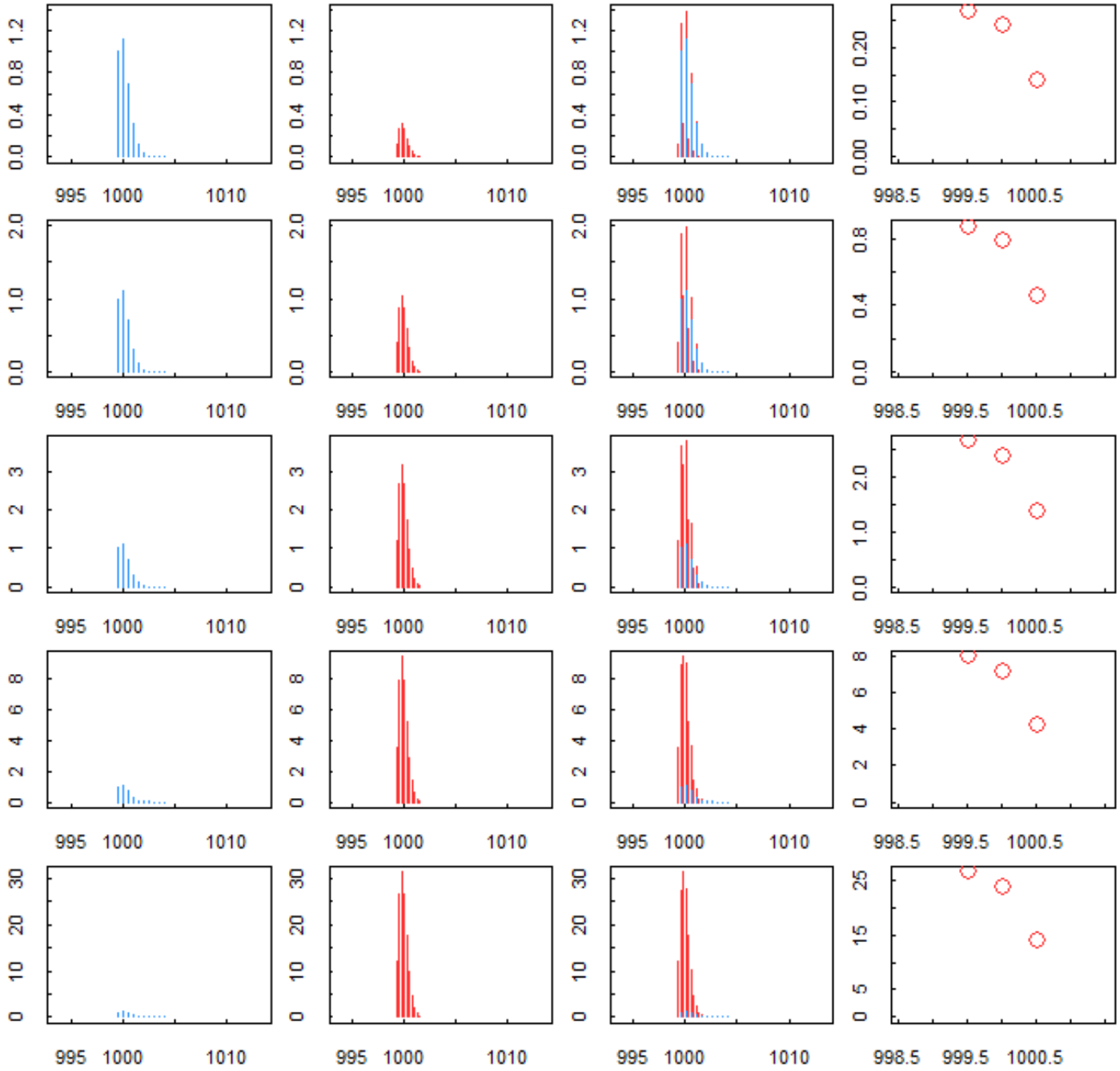
Interference

Analysis of low abundance proteins is **sensitive to interference** from other components of the sample.

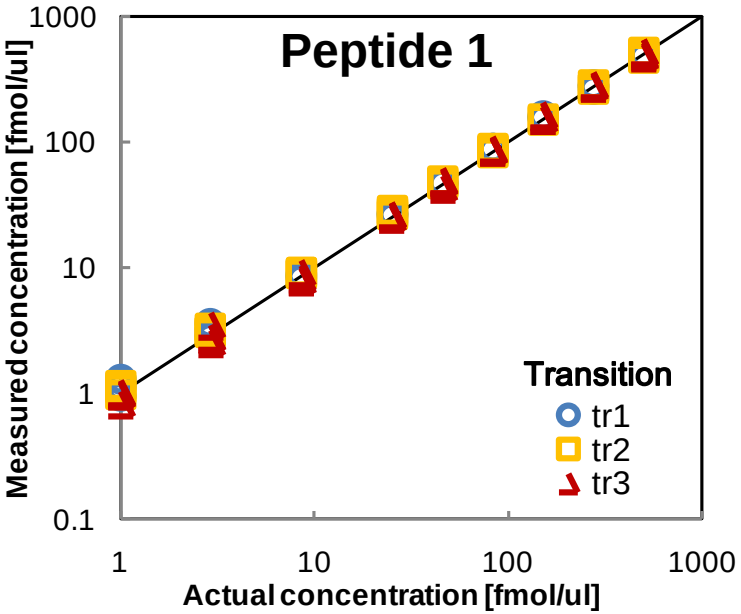
MS1 interference: other components of the sample that overlap with the isotope distribution.

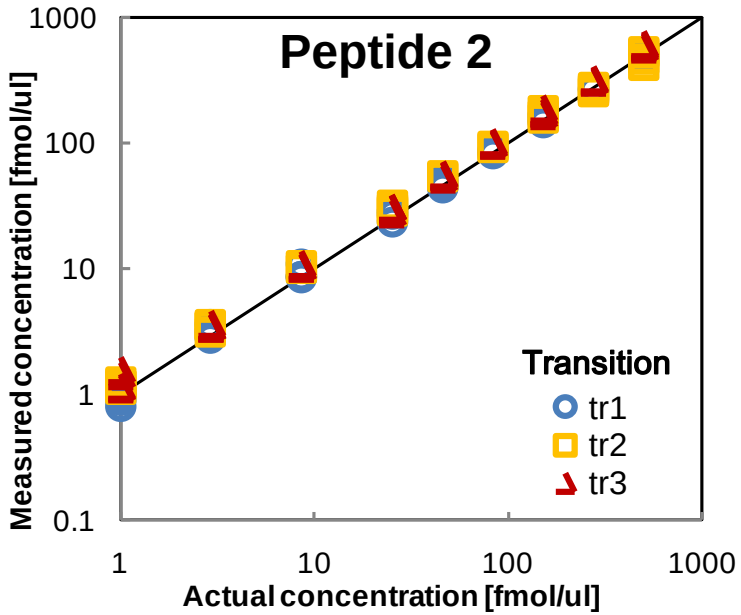
MS/MS interference: other components of the sample with same precursor and fragment masses as the transitions that are monitored.

MS1 interference



Quantitation using MRM





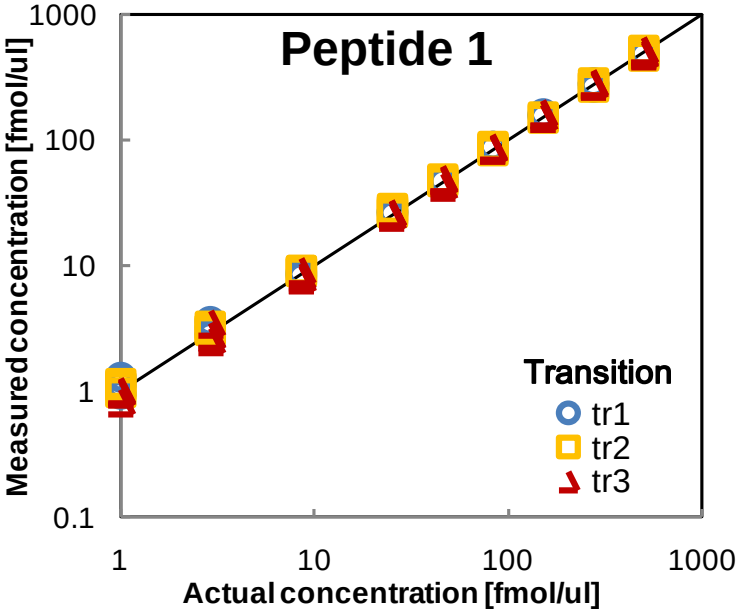
Data taken from CPTAC Verification Work Group Study 7.

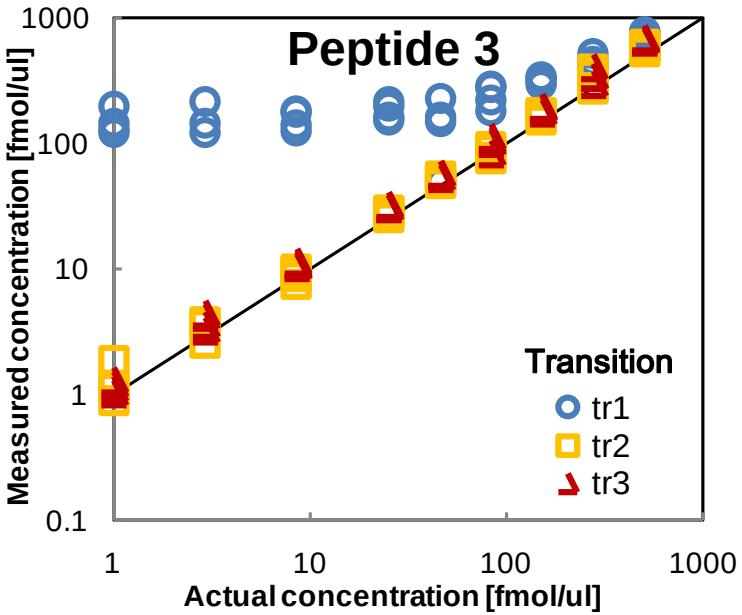
10 peptides
3 transitions per peptide
Concentrations 1-500 fmol/µl
Human plasma background
8 laboratories
4 repeat analysis per lab

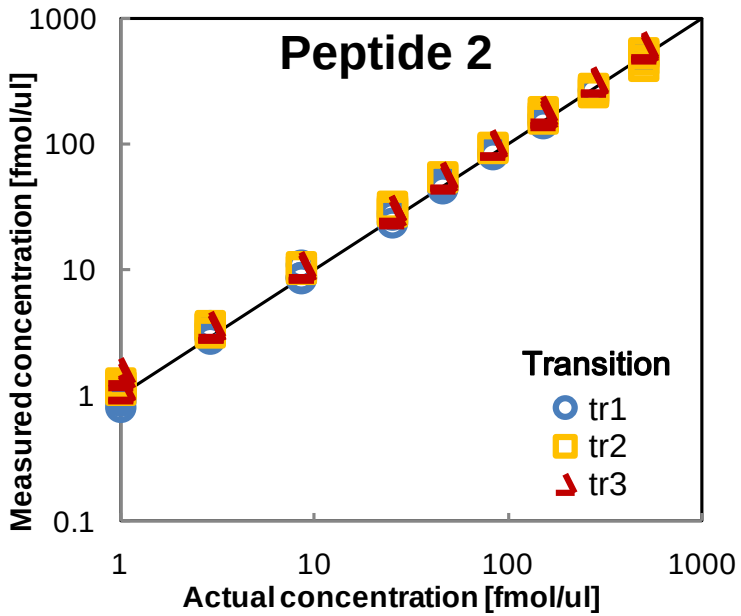
Addona et al., Nature Biotechnol. 27 (2009) 633-641

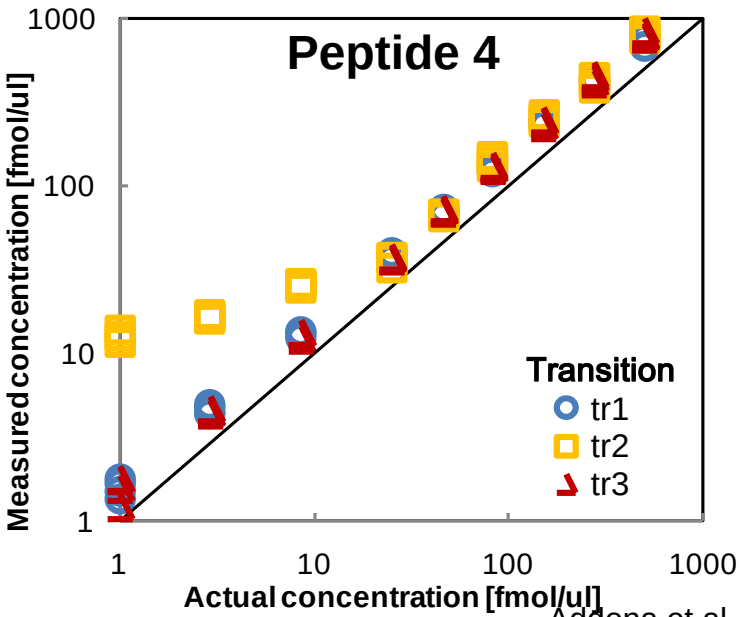
Addona et al., NBT 2009

Quantitation using MRM



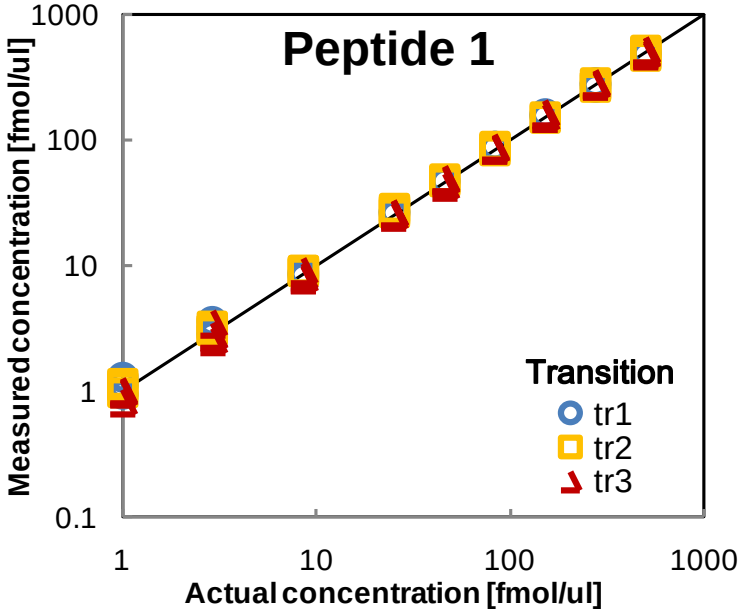


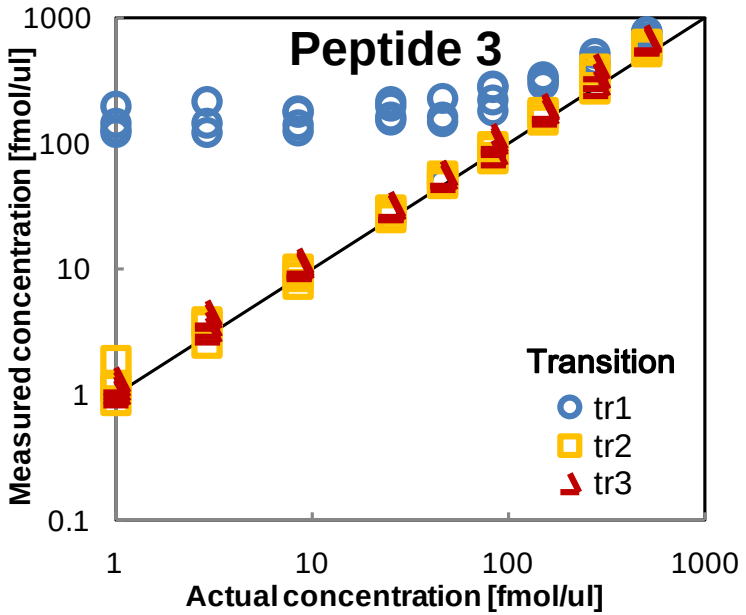


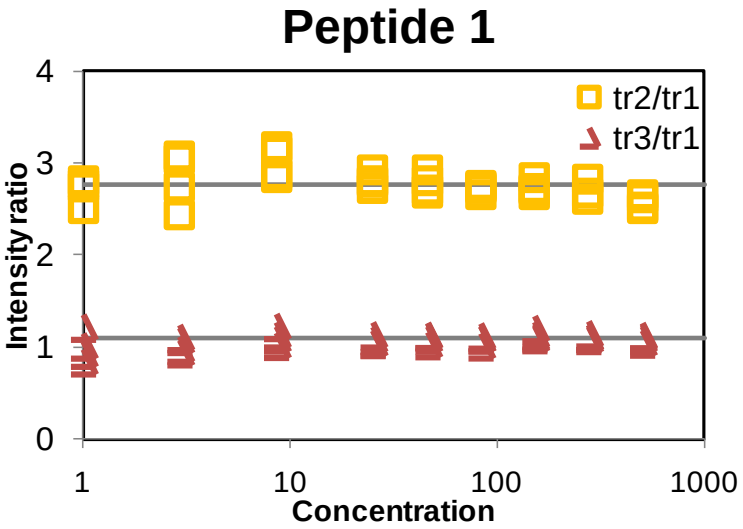


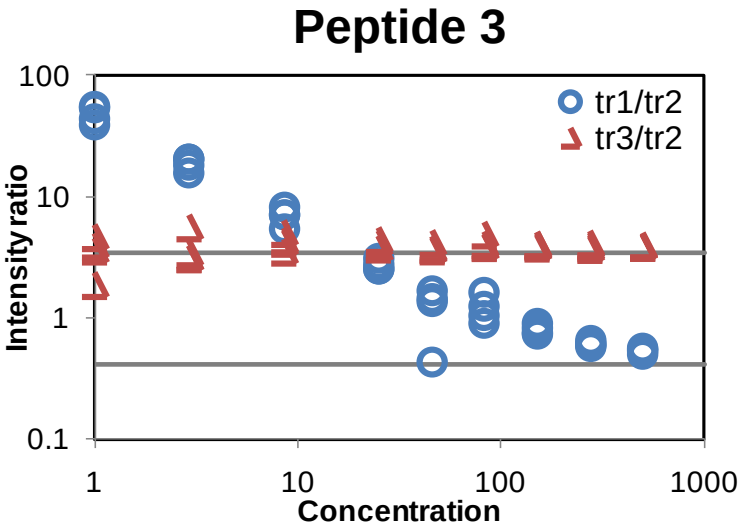
Addona et al., NBT 2009

Ratios of intensities of transitions





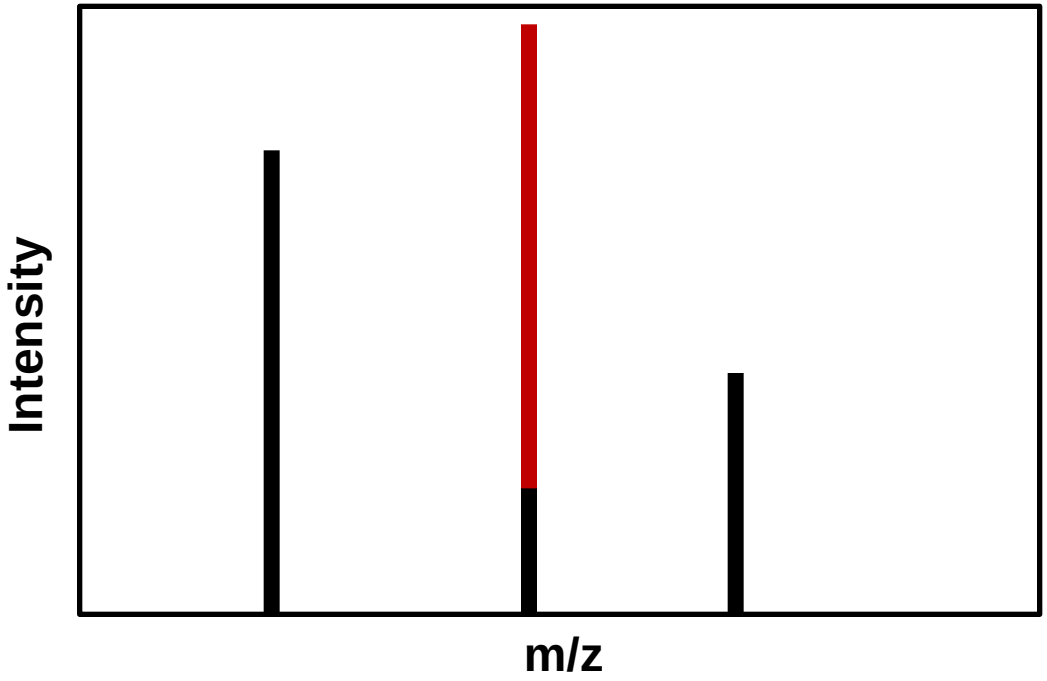




Addona et al., NBT 2009

Model: Noise and Interference

Can the knowledge of the relative intensity of the transitions be used to correct for interference?



- Noise is a normally distributed increase or decrease in the intensity.

- Interference is an increase in the intensity of one or more transitions.

Detection of interference

Interference is detected by comparing the ratio of the intensity of pairs of transitions with the expected ratio and finding outliers.

Transition i has interference if $Z_{threshold} < Z_i$

where $Z_{threshold}$ is the interference detection threshold;

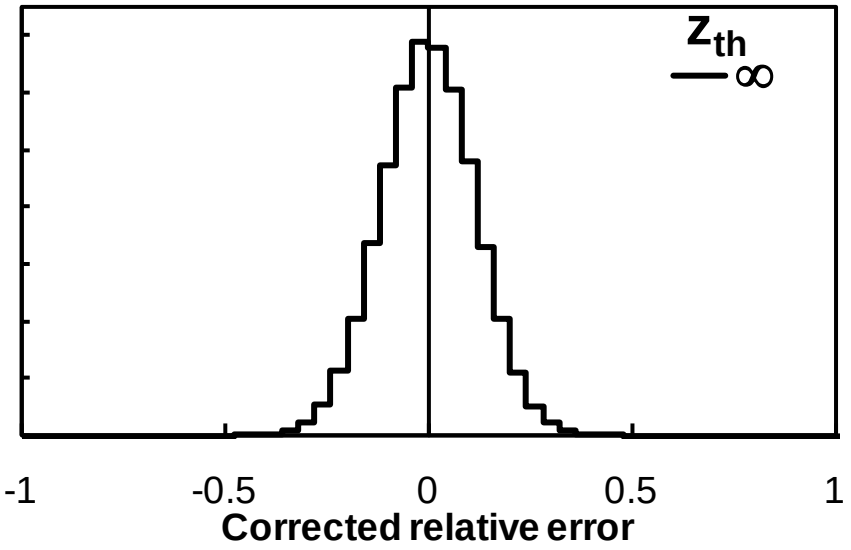
$$Z_i = \max_{j \neq i} Z_{ji} = \max_{j \neq i} \frac{r_{ji} - I_j / I_i}{\sigma_{ji}};$$

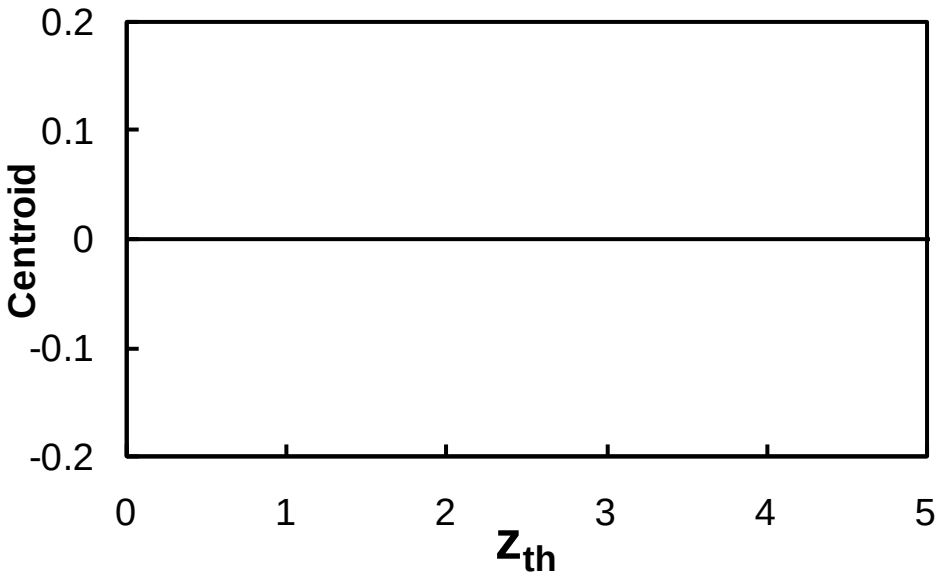
z_{ji} is the number of standard deviations that the ratio between the intensities of transitions j and i deviate from the noise;

I_i and I_j are the log intensities of transitions i and j;

r_{ji} is the median of the log intensity of transitions j and i;
σ_{ji} is the noise in the ratio.

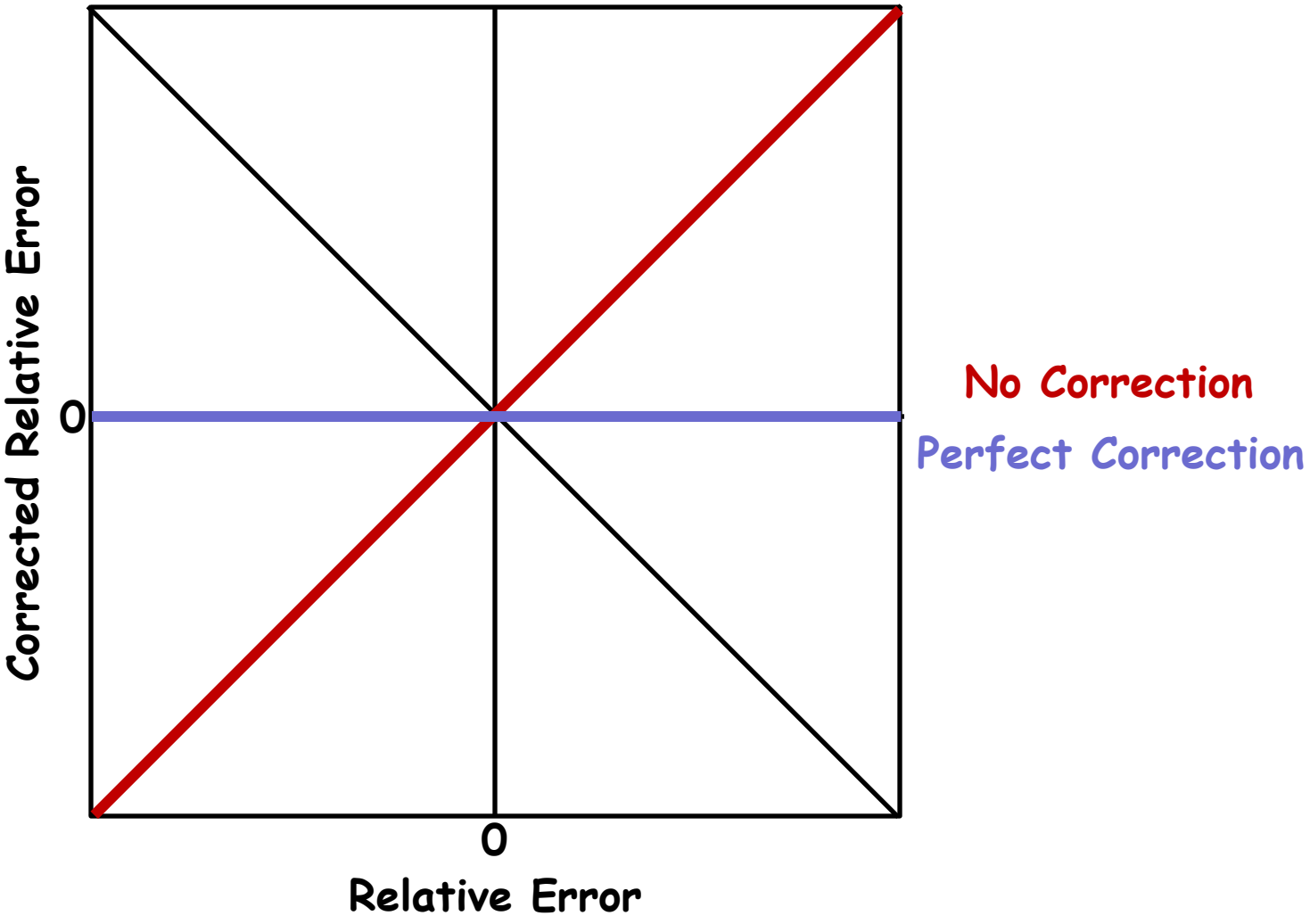
Error in quantitation after correction in presence of noise but no interference





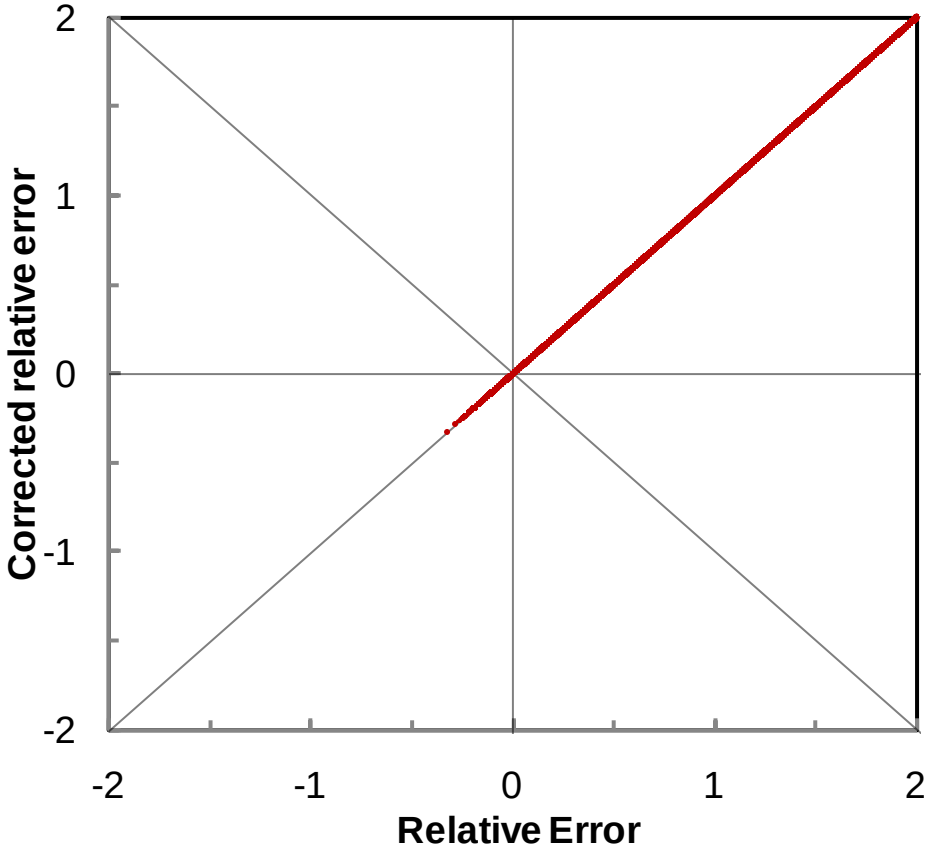
Relative noise = 0.2
No interference
Relative intensity of transitions: 1:1:1

Corrections for interference

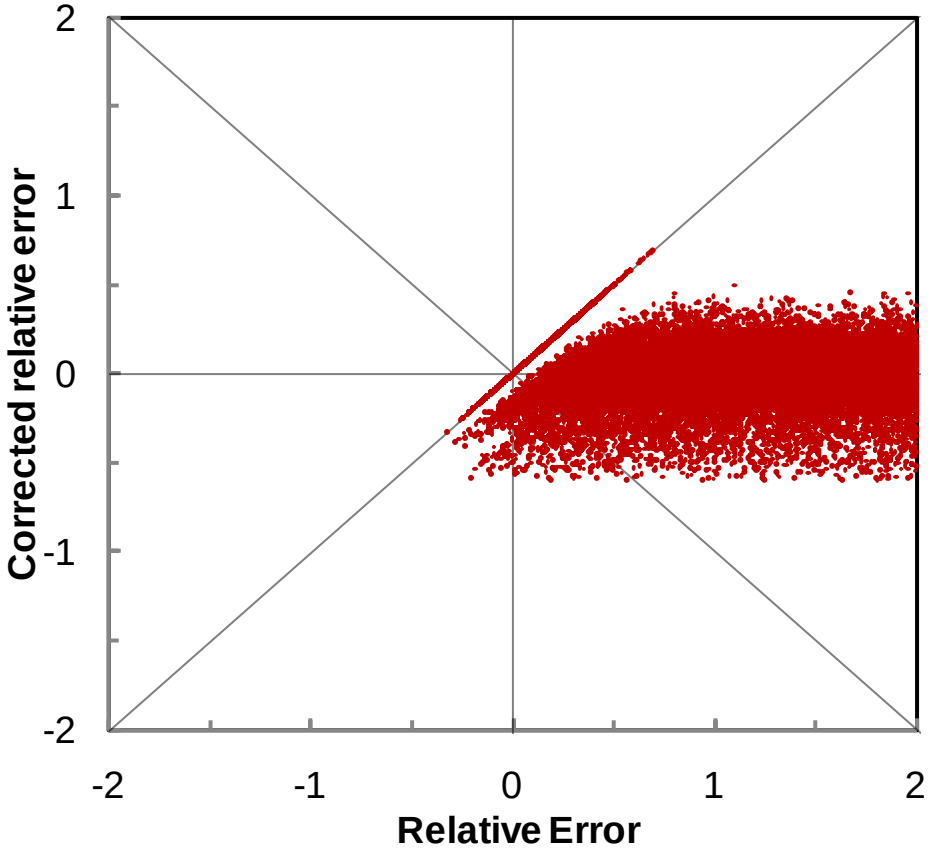


Error in quantitation after correction in presence of interference and noise

No correction

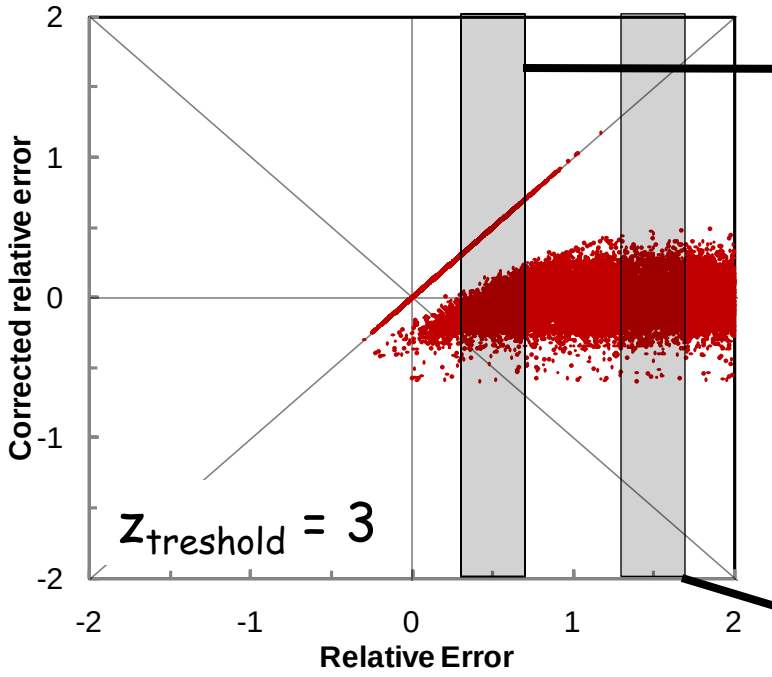


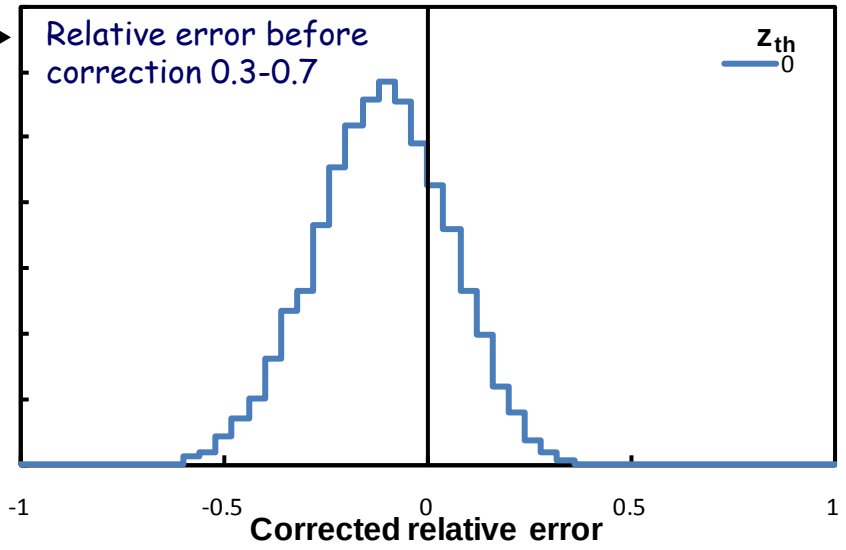
Correction (z_{th}=2)

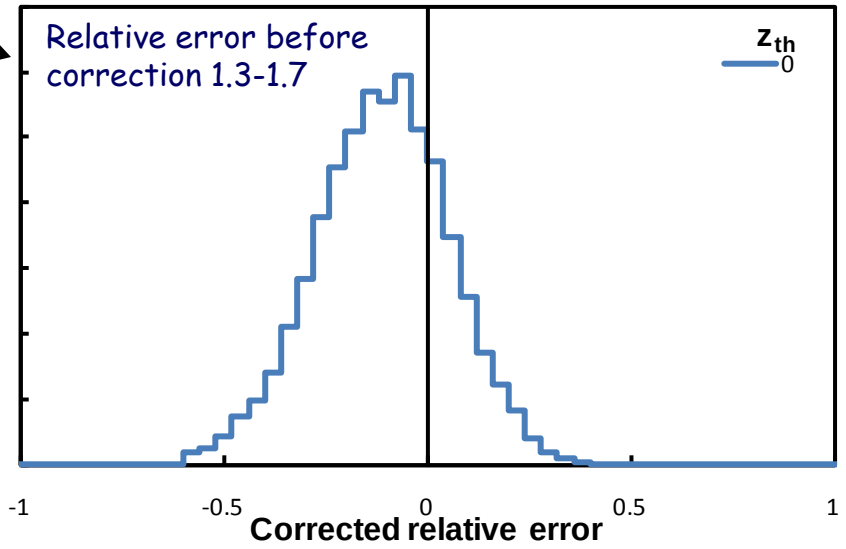


Relative noise = 0.2
Interference in 1 out of 3 transitions
Relative intensity of transitions: 1:1:1

Error in quantitation after correction in presence of interference and noise

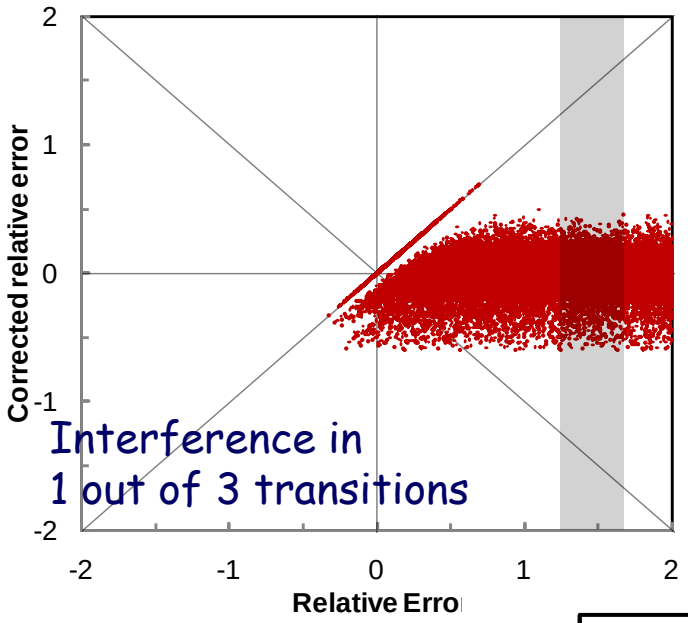


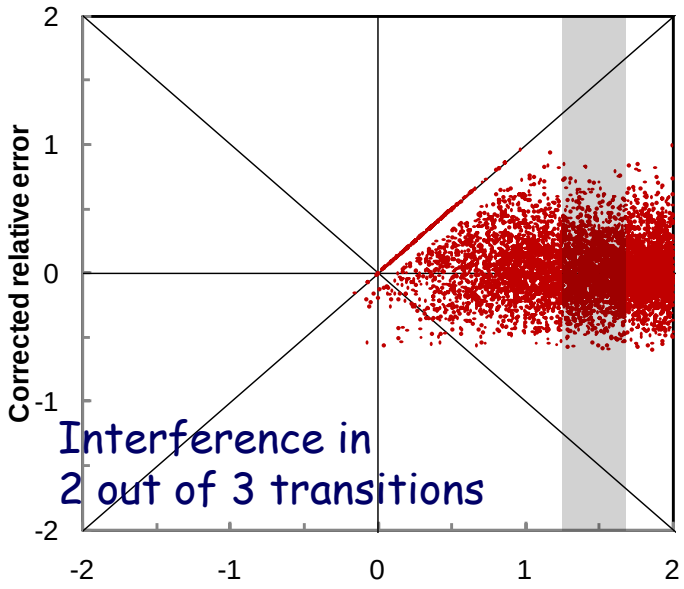


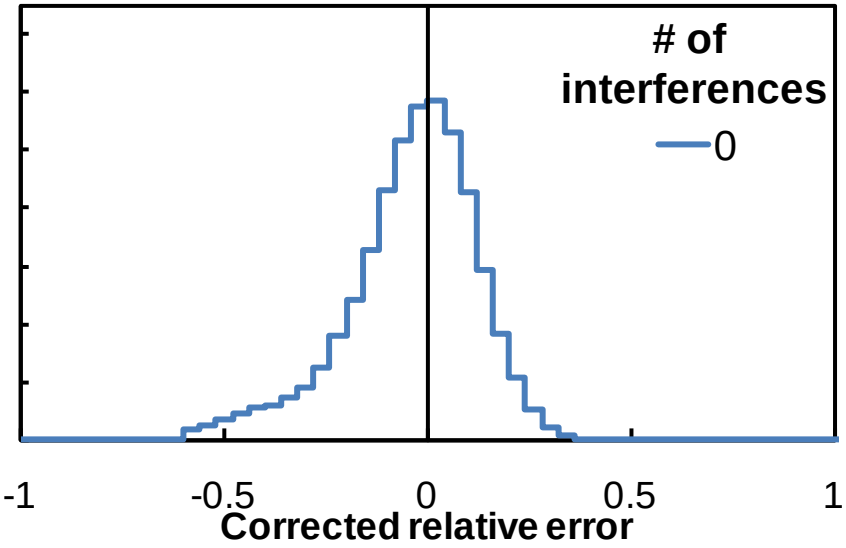


Relative noise = 0.2
Interference in 1 out of 3 transitions
Relative intensity of transitions: 1:1:1

Error in quantitation after correction in presence of interference and noise

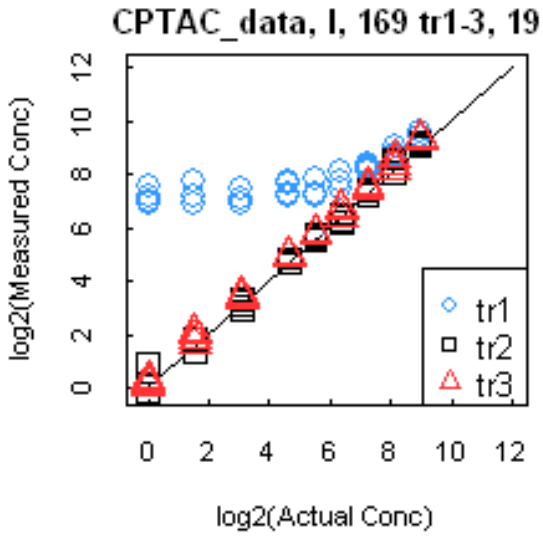


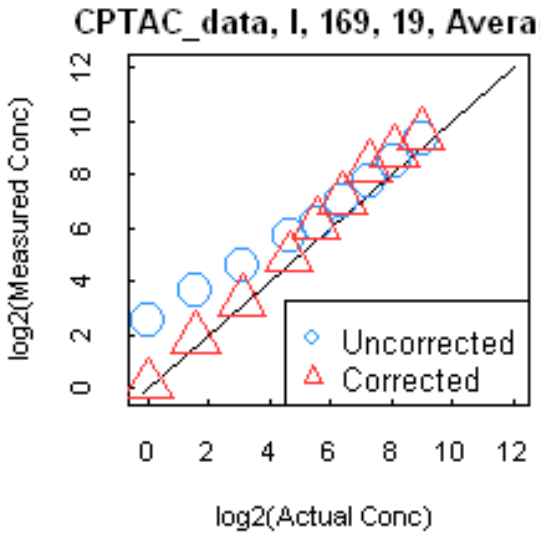


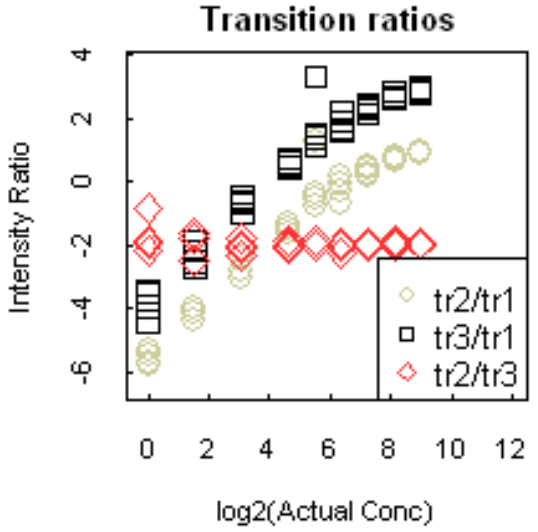


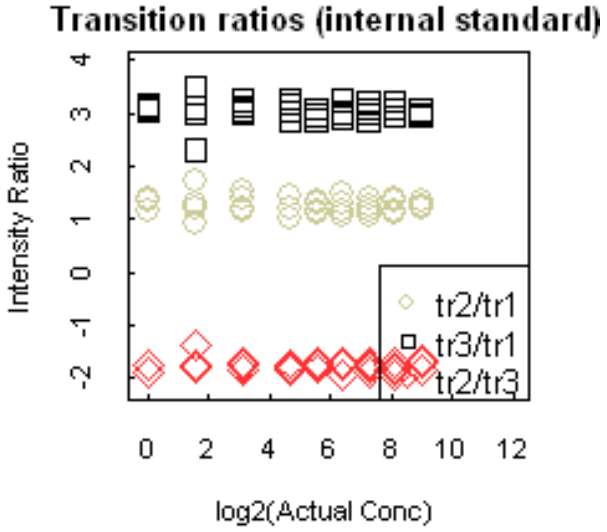
z_{th} = 2
Relative noise = 0.2
Relative intensity of transitions:

Correction for MS2 interference







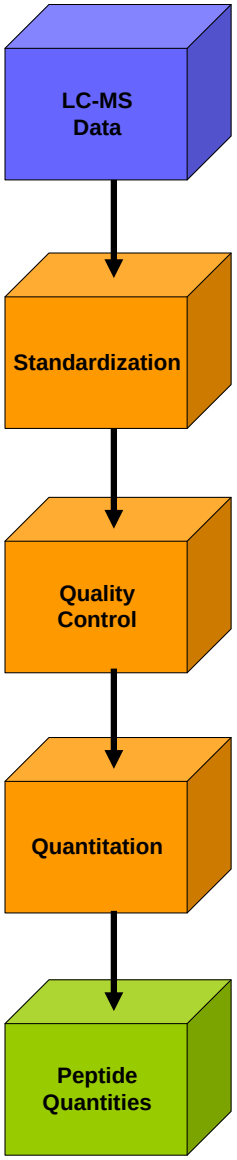


Workflow for quantitation with LC-MS

Standardization
Retention time alignment
Mass calibration
Intensity normalization

Quality Control
Detection of problems with samples and analysis

Quantitation
Peak detection
Background subtraction
Limits for integration in time and mass
Exclusion of interfering peaks



Takeaway Message

- There are many different ways to quantitate proteins – choose the one that is appropriate for your application.
- In general the earlier you can introduce isotopic labels the better the accuracy.
- Always monitor for interference.

Protein quantitation I: Overview (Week 8)