

Computational Methods for High-Throughput Omics Data

Proteomics

„Proteomics is defined as the study of the ensemble of proteins at a given point in time, especially their expression pattern, ~~structure and function.~~”



Quantitation

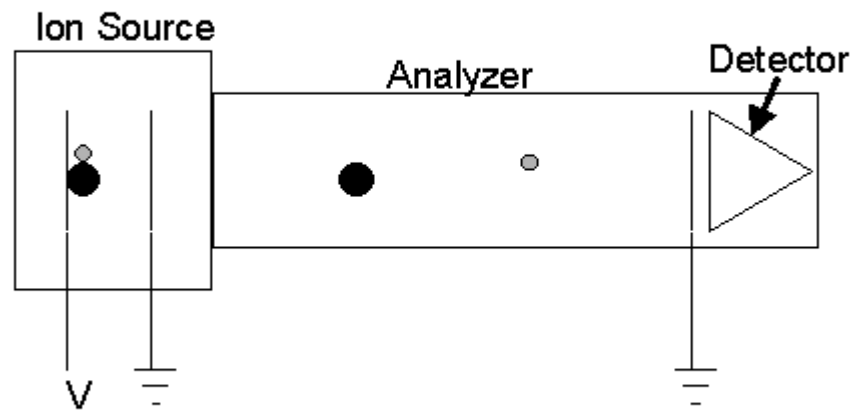
Identification

Basics

Protonation (H^+)



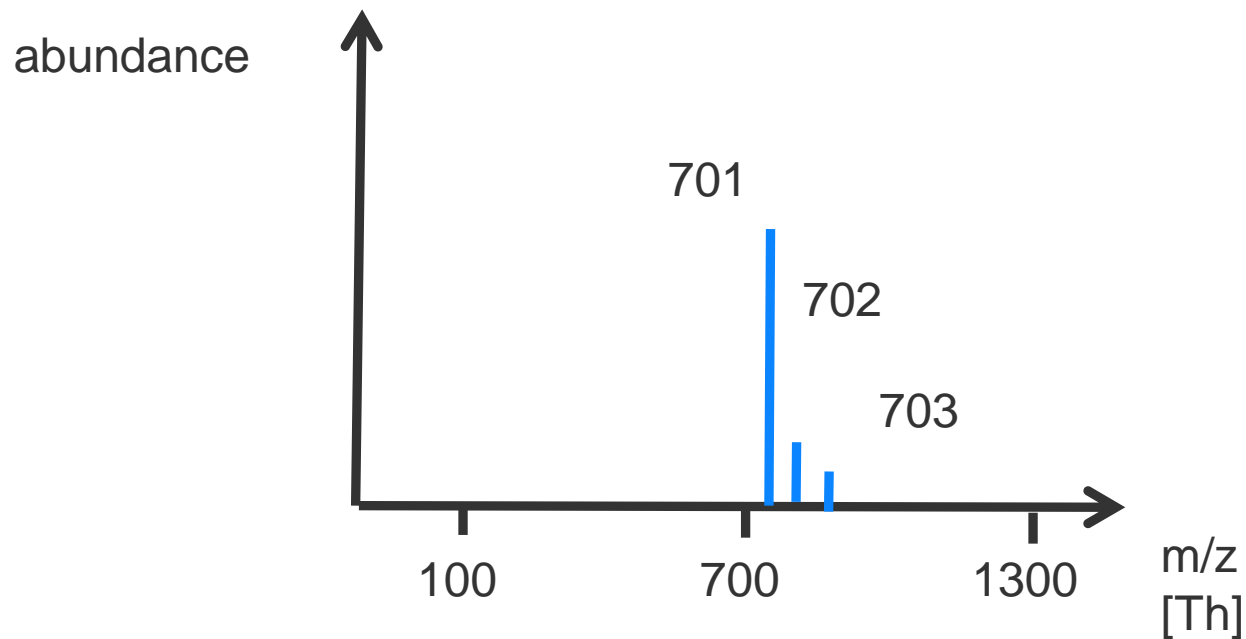
Linear TOF Mass Spectrometer



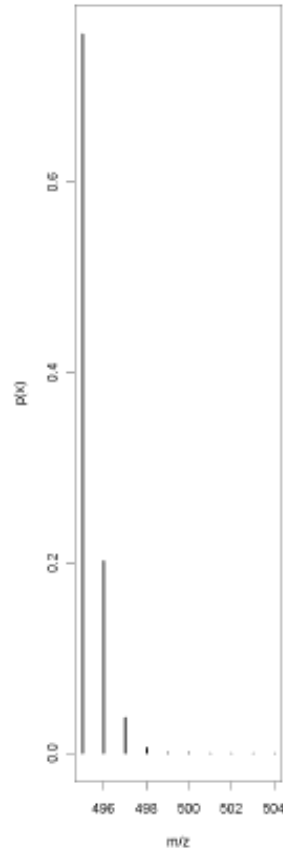
Basics

mass(„DPFINA“) = 700 Da = 700 u
charge = 1

SumFormula: $C_{31}H_{45}N_7O_{10}$
Isotopes: ^{12}C and ^{13}C

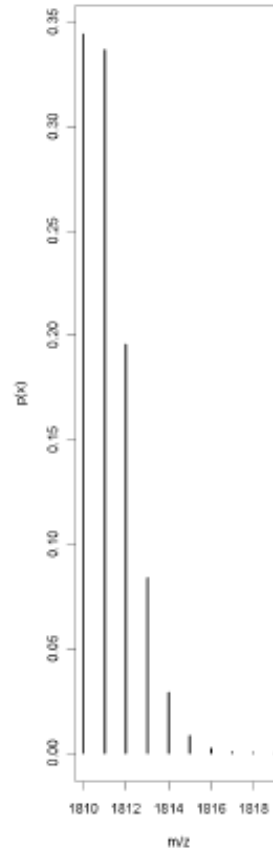


Basics

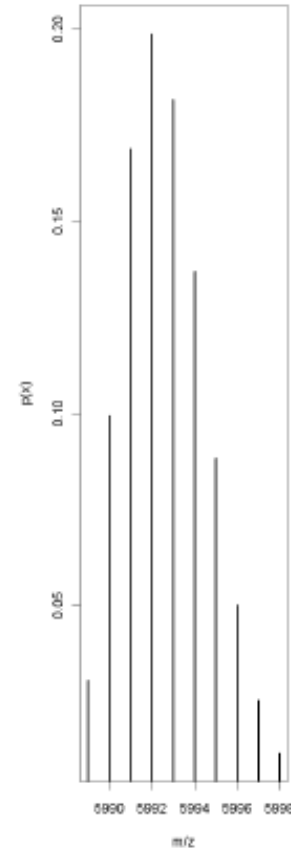


Mass (Da)

500



1800

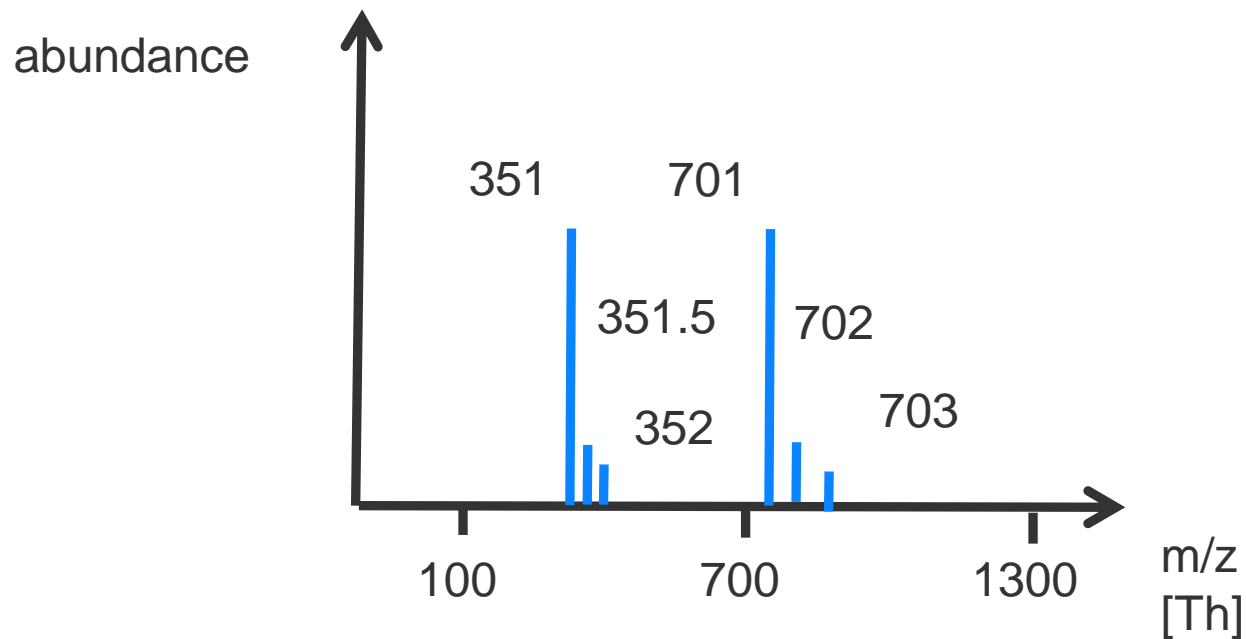


6000

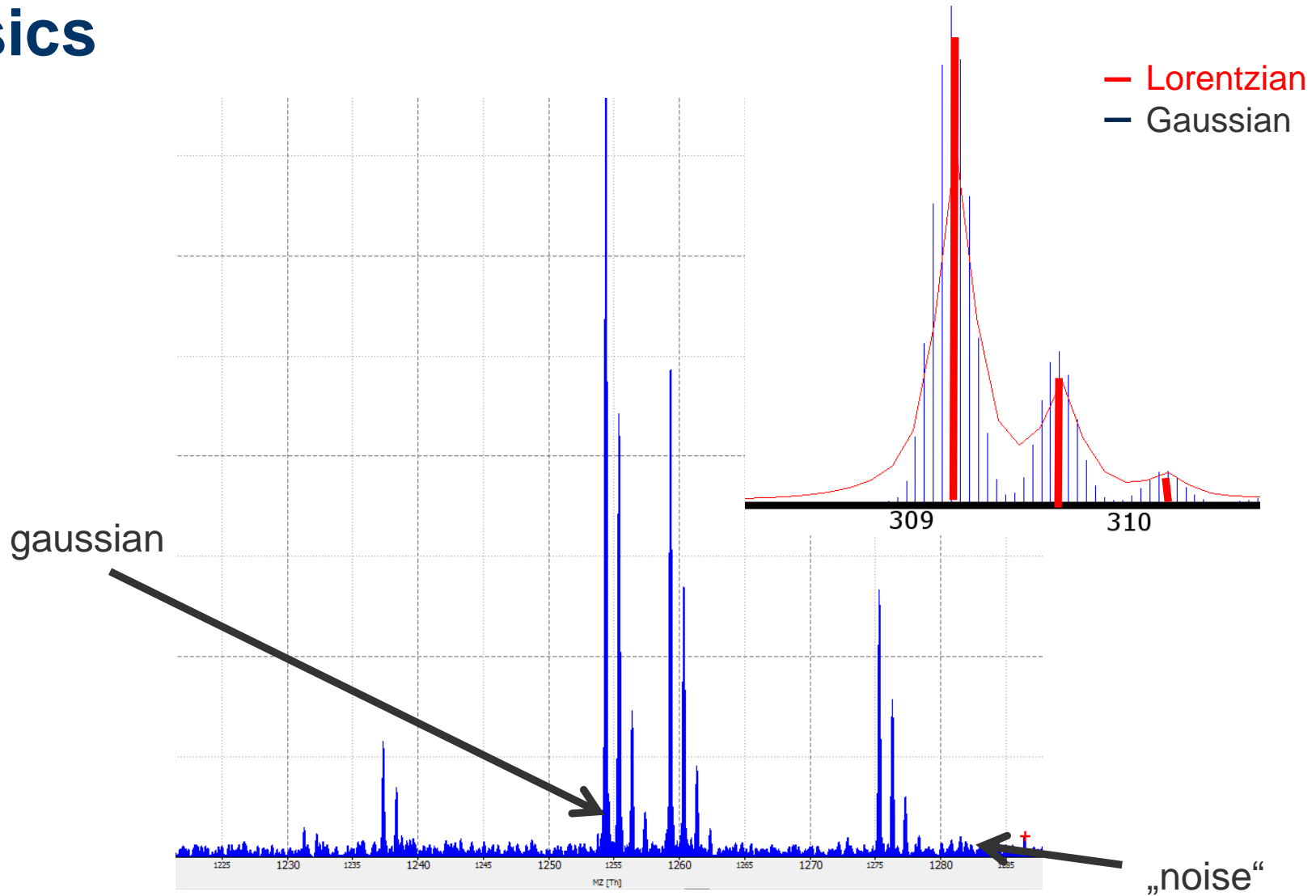
Basics

mass(„DPFINA“) = 700 Da = 700 u
charge = 2

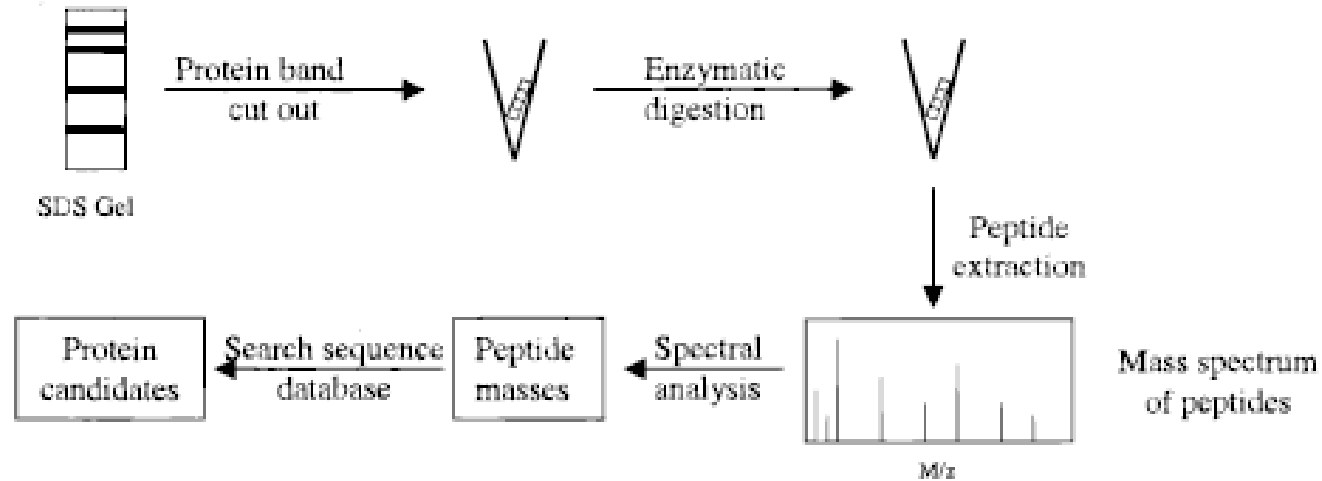
SumFormula: $C_{31}H_{45}N_7O_{10}$
Isotopes: ^{12}C and ^{13}C



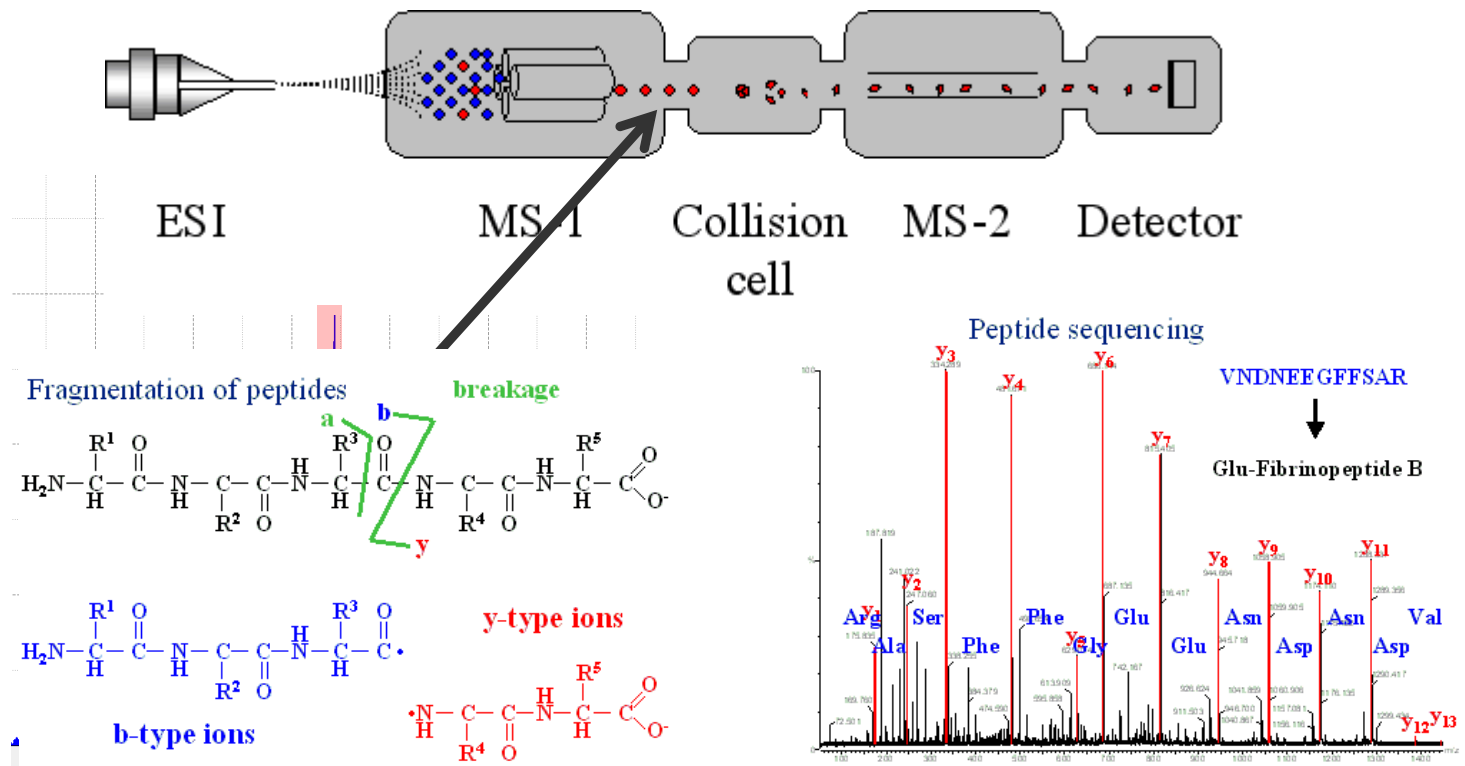
Basics



Talk: Identification [PMF]

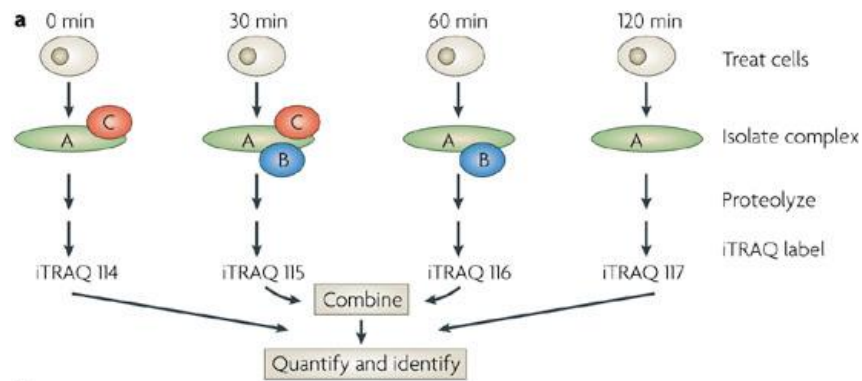


Talk: Identification

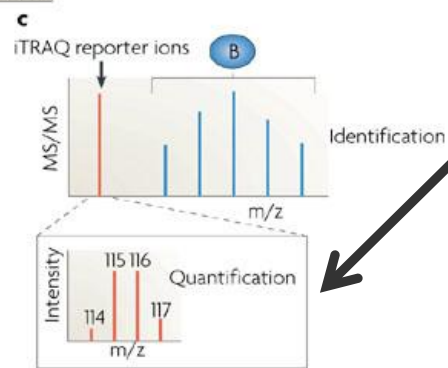
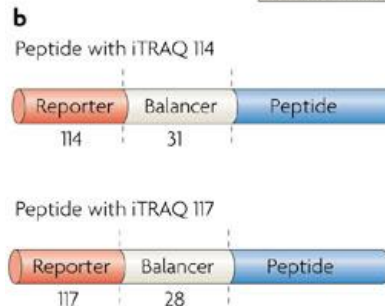


Talk: Multiplexing using iTRAQ

Isobaric tag for relative and absolute quantitation

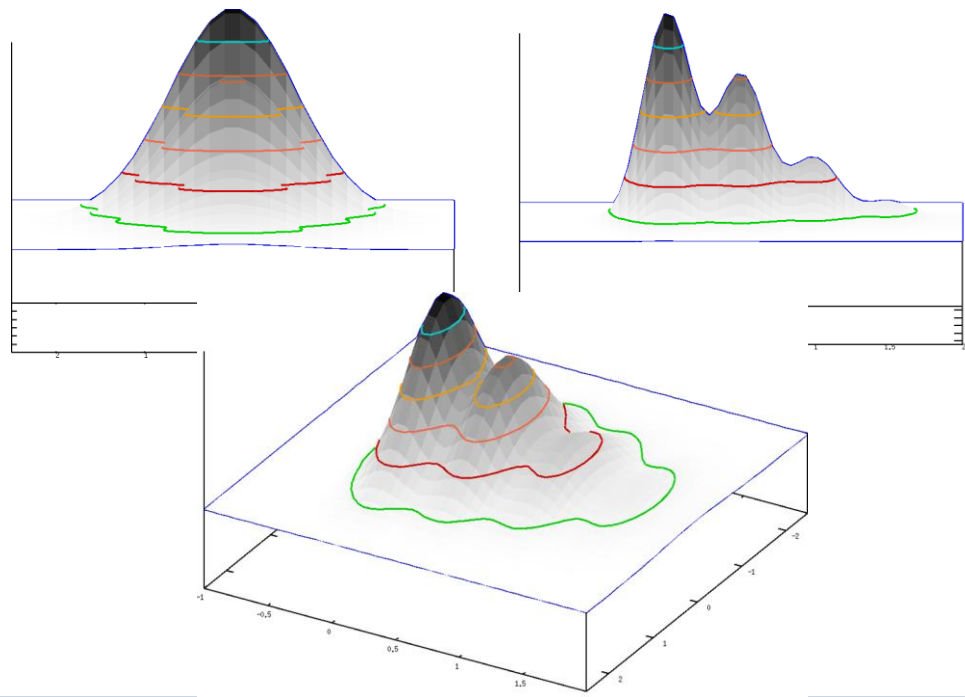
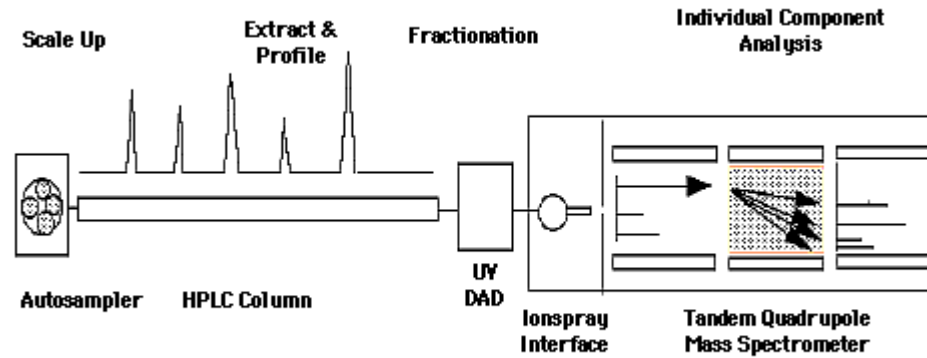


← Peptides!

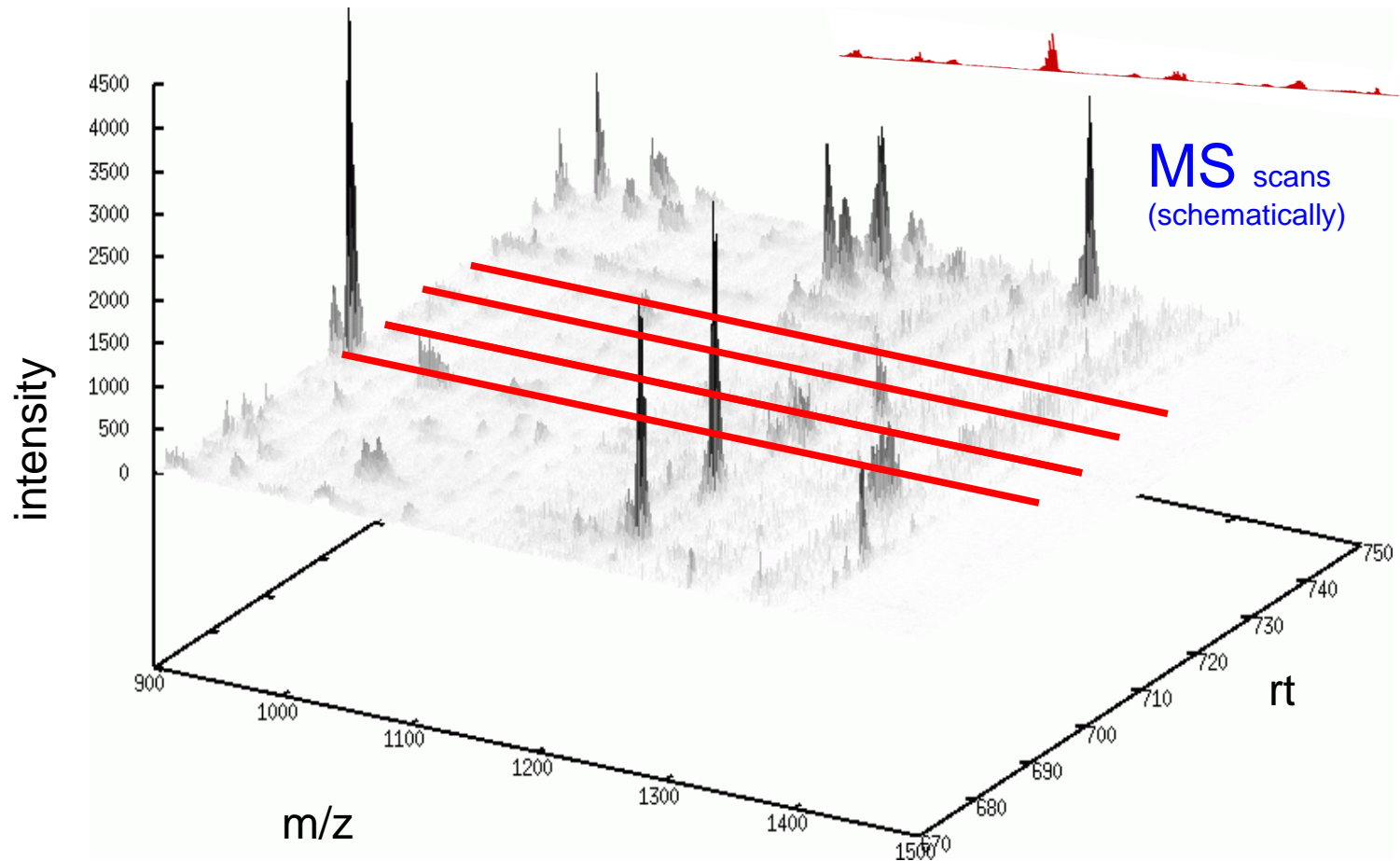


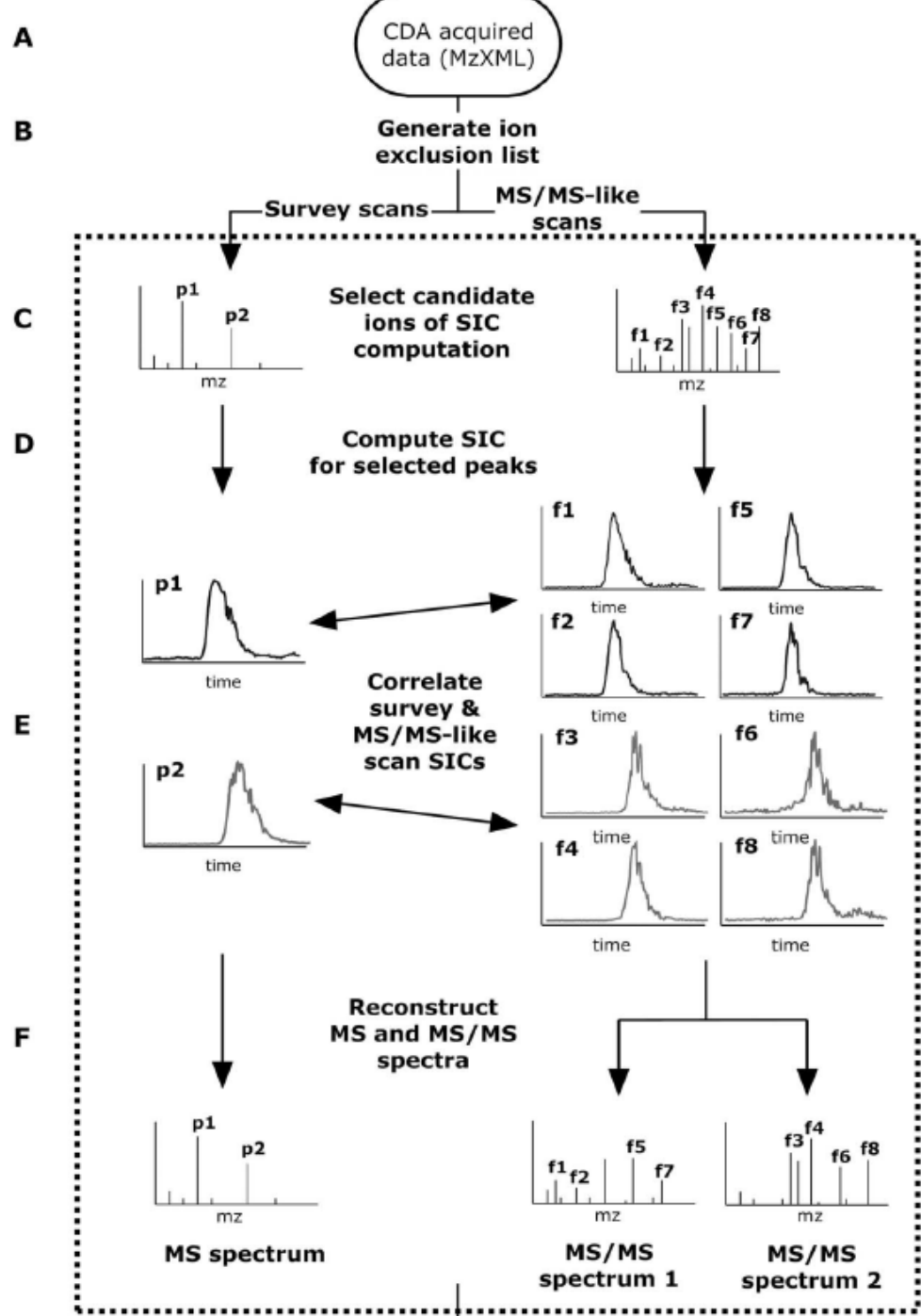
Nature Reviews | Molecular Cell Biology

Identification: ETISEQ

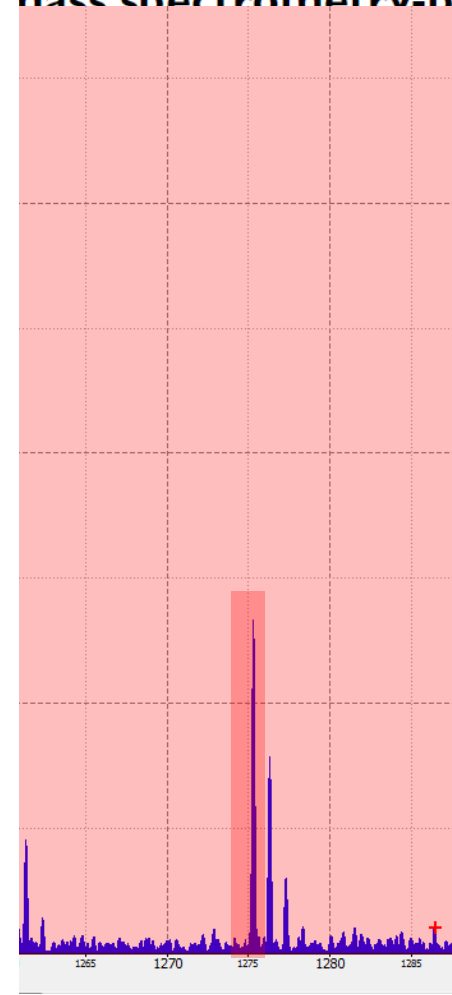


Identification: ETISEQ





Retention time ion sequencing mass spectrometry-based



Isotope Distribution

Watch & Learn

<http://www.youtube.com/watch?v=dq7xKVvXWew>

Isotope Distribution Computation

The probability that a molecule composed out of one atom of element E and one atom of element E' has a total of n additional neutrons is

$$\pi_{EE'}[n] = \sum_{i=0}^n \pi_E[i] \pi_{E'}[n - i].$$

Note that $\pi_{EE'}[\ell] = 0$ for $\ell > k_E + k_{E'}$.

This type of composition is very common in mathematics and known as a *convolution* operation, denoted by the operator $*$.

Using the convolution operator, we can rewrite the above equation as

$$\pi_{EE'} = \pi_E * \pi_{E'}.$$

For example, a hypothetical molecule composed out of one carbon and one nitrogen would have $\pi_{CN} = \pi_C * \pi_N$,

$$\pi_{CN}[0] = \pi_C[0] \pi_N[0],$$

$$\pi_{CN}[1] = \pi_C[0] \pi_N[1] + \pi_C[1] \pi_N[0],$$

$$\pi_{CN}[2] = \pi_C[0] \pi_N[2] + \pi_C[1] \pi_N[1] + \pi_C[2] \pi_N[0] \quad (= \pi_C[1] \pi_N[1]).$$

Isotope Distribution

Largest n that needs computation?

$$\pi_{EE'}[\ell] = 0 \text{ for } \ell > k_E + k'_E.$$

$$\sum_{i=0}^n \pi_E[i] = ?$$

Speed up computation?! (Kubinyi)

Π_{8C}

$$\Pi_{1C} * \Pi_{1C} = \Pi_{2C}$$

$$\Pi_{1C} * \Pi_{2C} = \Pi_{3C}$$

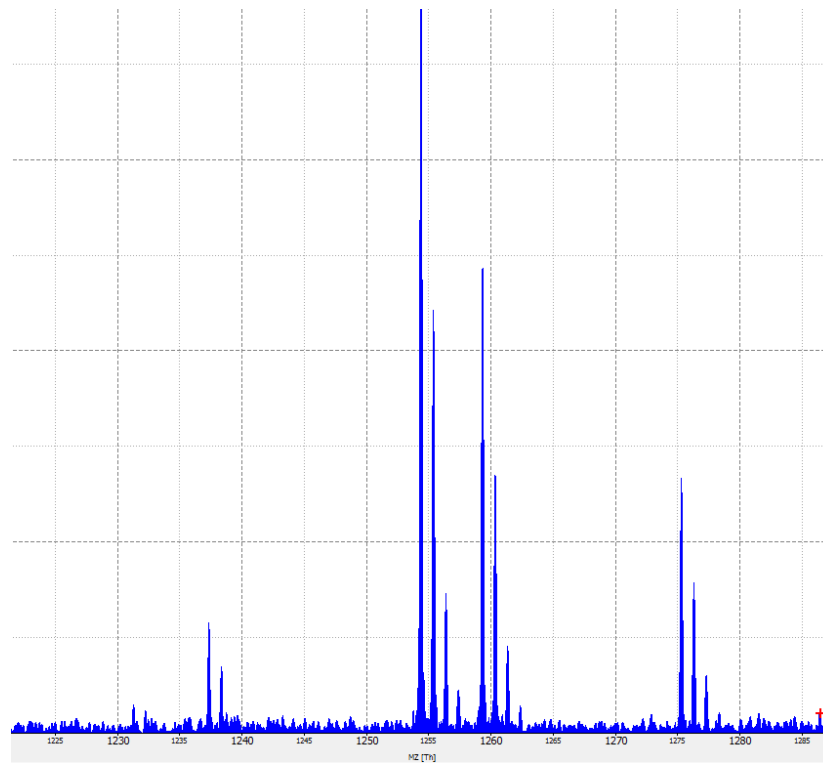
...

$$\Pi_{2C} * \Pi_{2C} = \Pi_{4C}$$

Nomenclature

BASE PEAK

highest peak in a spectrum



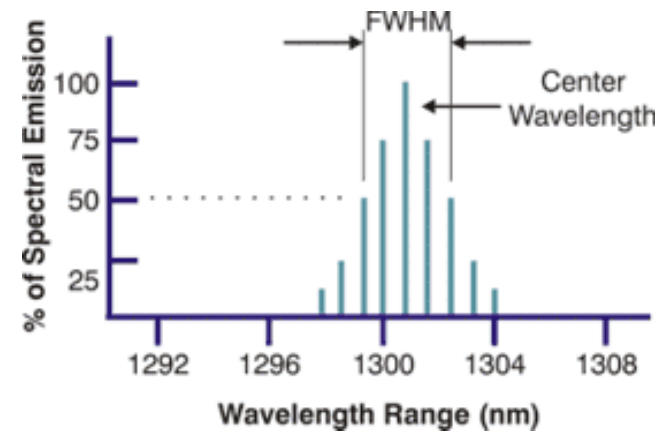
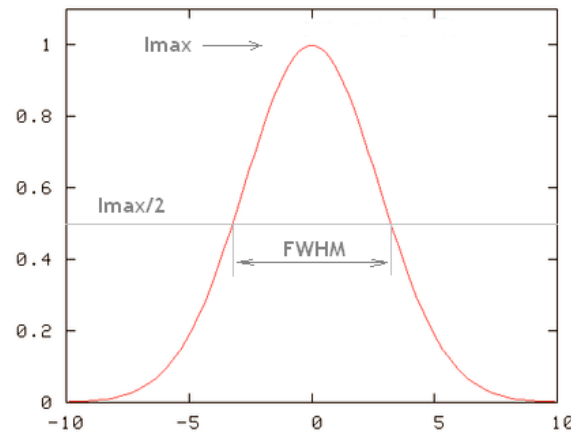
Nomenclature

BASE PEAK

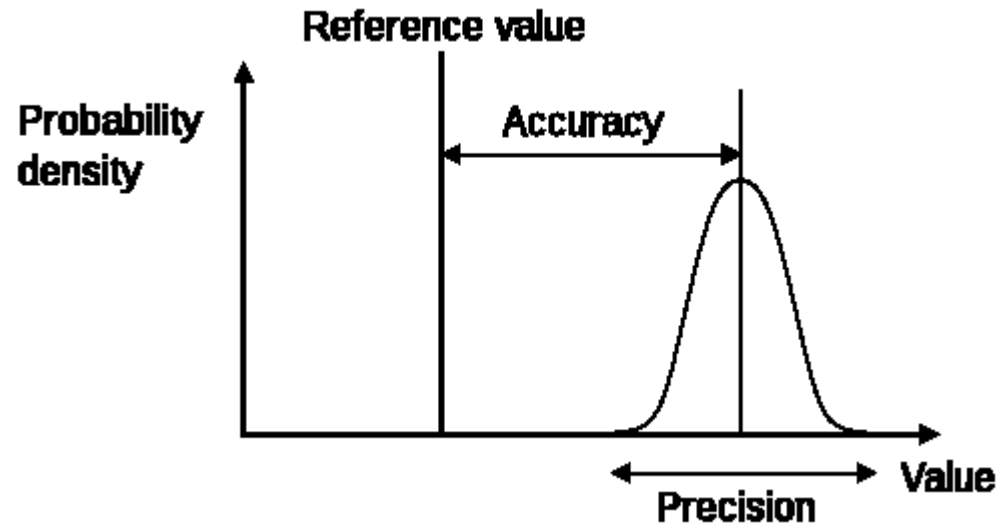
highest peak in a spectrum

RESOLUTION

$R = m / \Delta m$ @ 400 Th



Accuracy & Precision



Talks

- How to:
 - slides
 - timing
- Preliminary discussion
- Preliminary slides & questions