

# Computational Methods for High-Throughput Omics Data

## General stuff ...

- 1 proteomics OR 1 genomics paper
- Questionnaire for audience
  - 5 to 10 questions for each talk
- Short summary [2-3 pages]
  
- Feedback sheet
  
- Camera [optional]

# Grading & Attendance

- $\geq 80\%$  (of 13)
- Final grade: average of talk & short summary

# 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

# Proteomics

Identification & Quantitation methods:

- 2-D electrophoresis (isoelectric point + mass(SDS-PAGE)) + WesternBlot | Staining
- ELISA
- Liquid Chromatography (LC) (i.e., HPLC | CE)
- **Mass Spectrometry (MS)**

# Basics

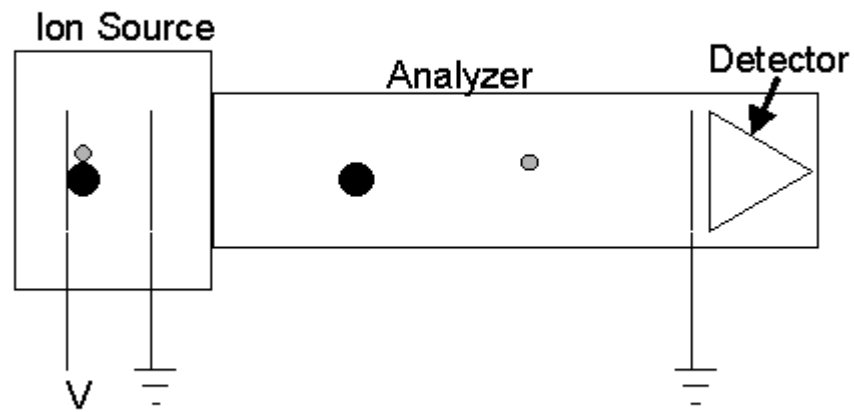


# Basics

Protonation ( $H^+$ )



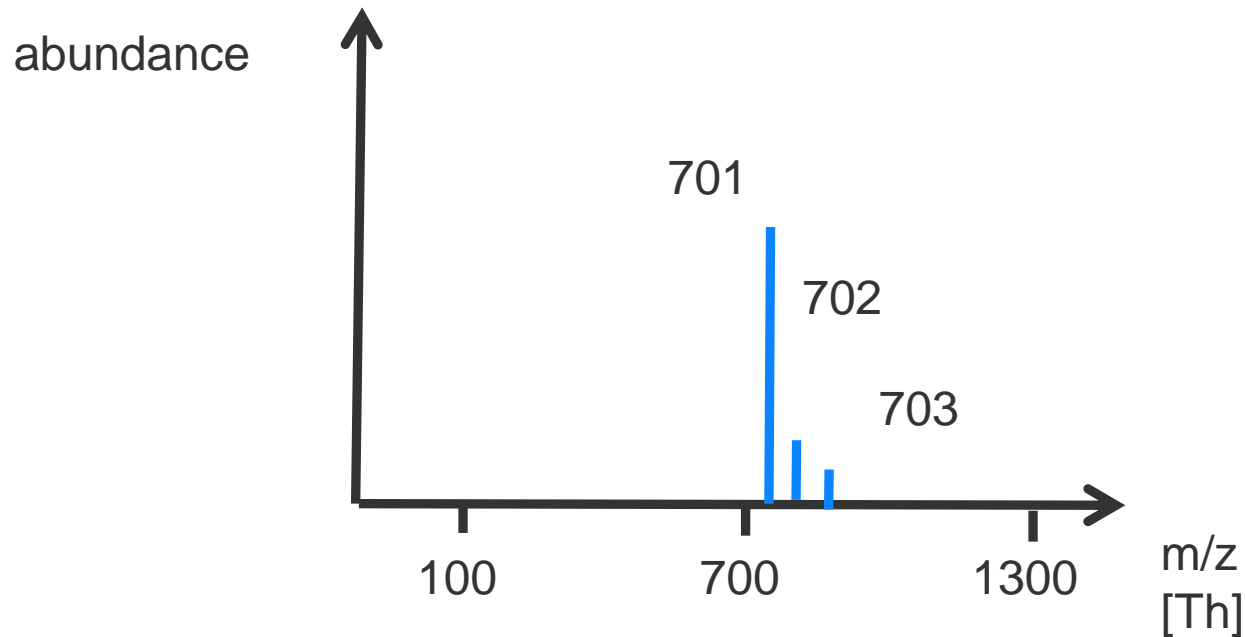
Linear TOF Mass Spectrometer



# Basics

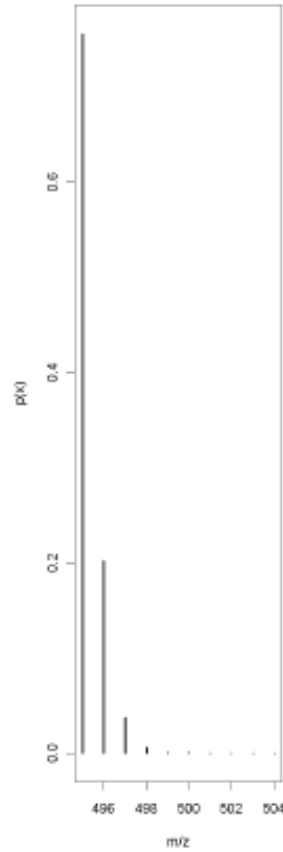
mass(„DPFINA“) = 700 Da = 700 u  
 charge = 1

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



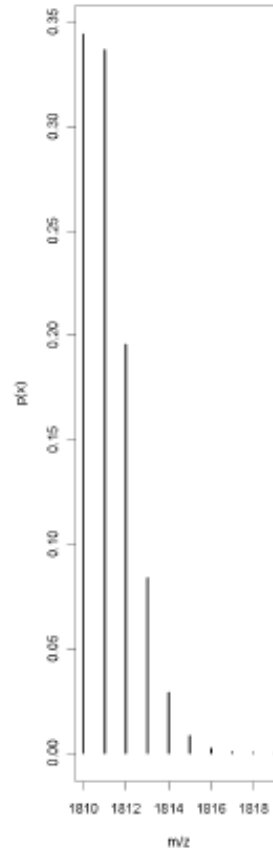


# Basics

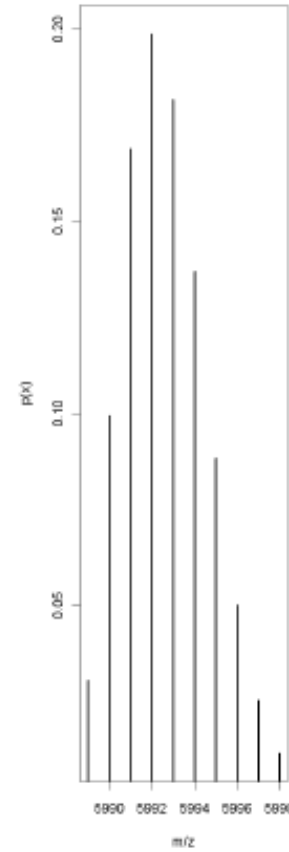


Mass (Da)

500



1800

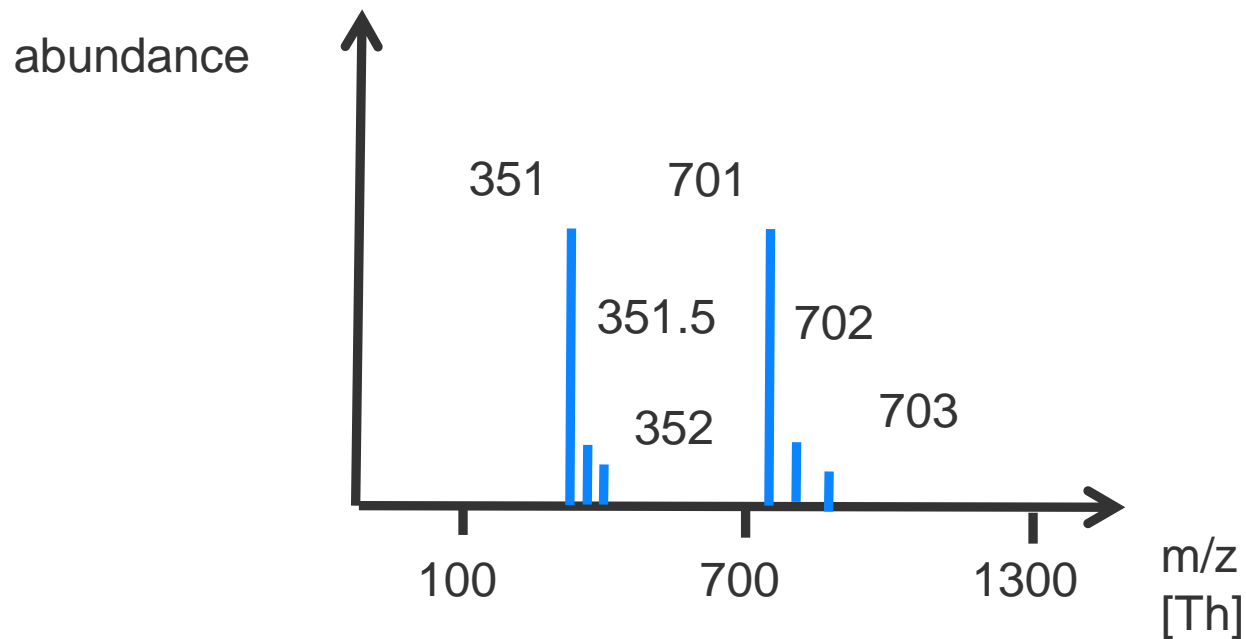


6000

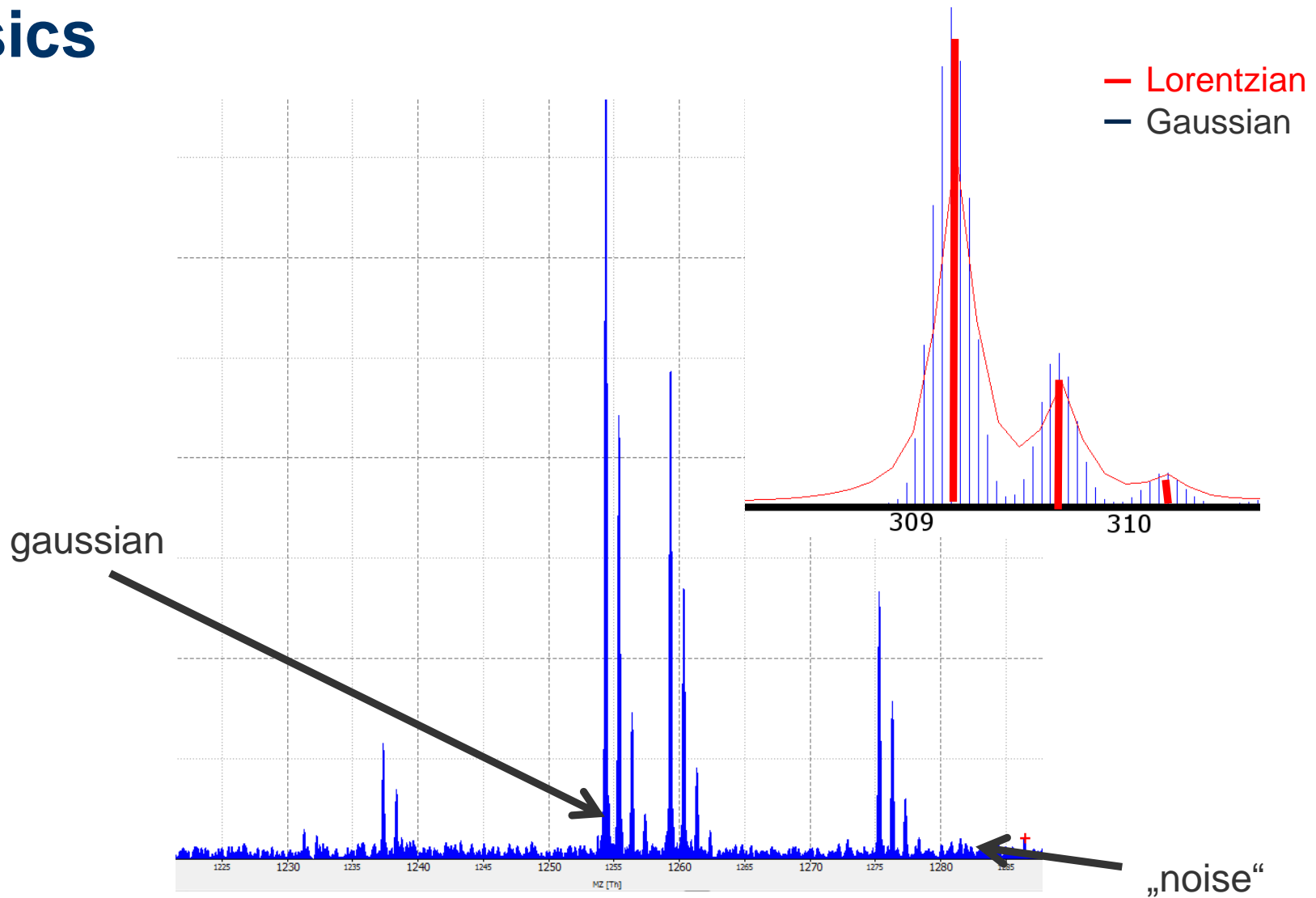
# Basics

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

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



# Basics



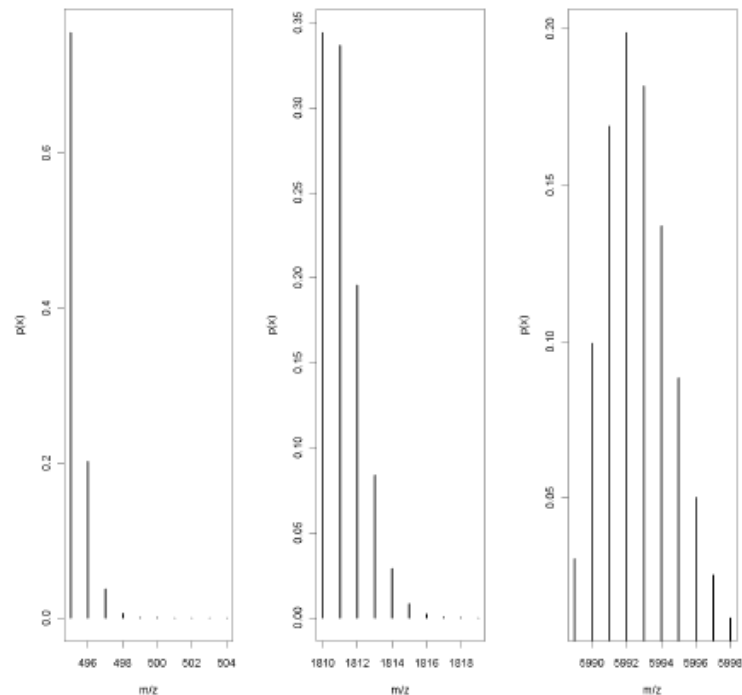
# Talk: Isotope Distribution

Calculation of isotope distributions in mass spectrometry.  
A trivial solution for a non-trivial problem

Hugo Kubinyi

*Hauptlaboratorium, BASF AG, D-6700 Ludwigshafen (Germany)*

(Received 26th September 1990)

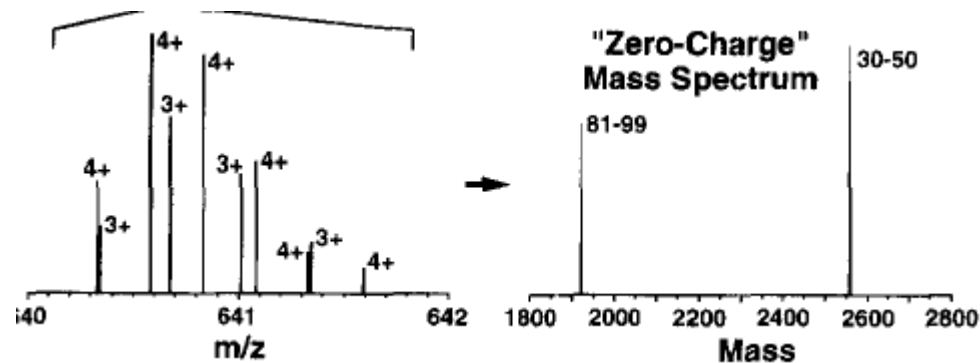


# Talk: Decharging

## A Universal Algorithm for Fast and Automated Charge State Deconvolution of Electrospray Mass-to-Charge Ratio Spectra

Zhongqi Zhang and Alan G. Marshall\*

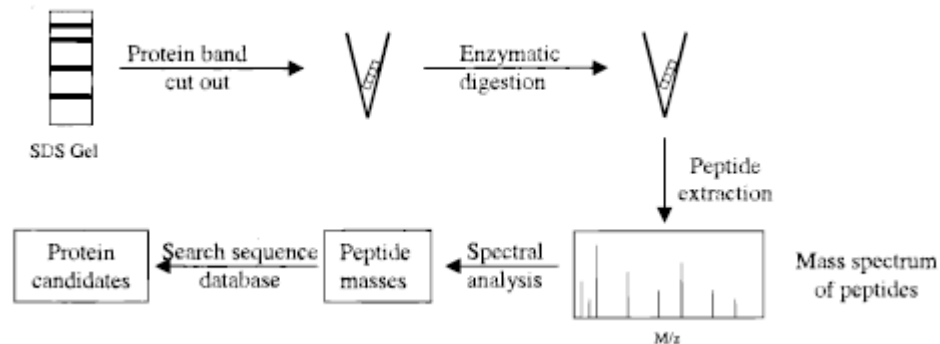
Center for Interdisciplinary Magnetic Resonance, National High Magnetic Field Laboratory, Florida State University, Tallahassee, Florida, USA



# Talk: Identification [PMF]

## ProFound: An Expert System for Protein Identification Using Mass Spectrometric Peptide Mapping Information

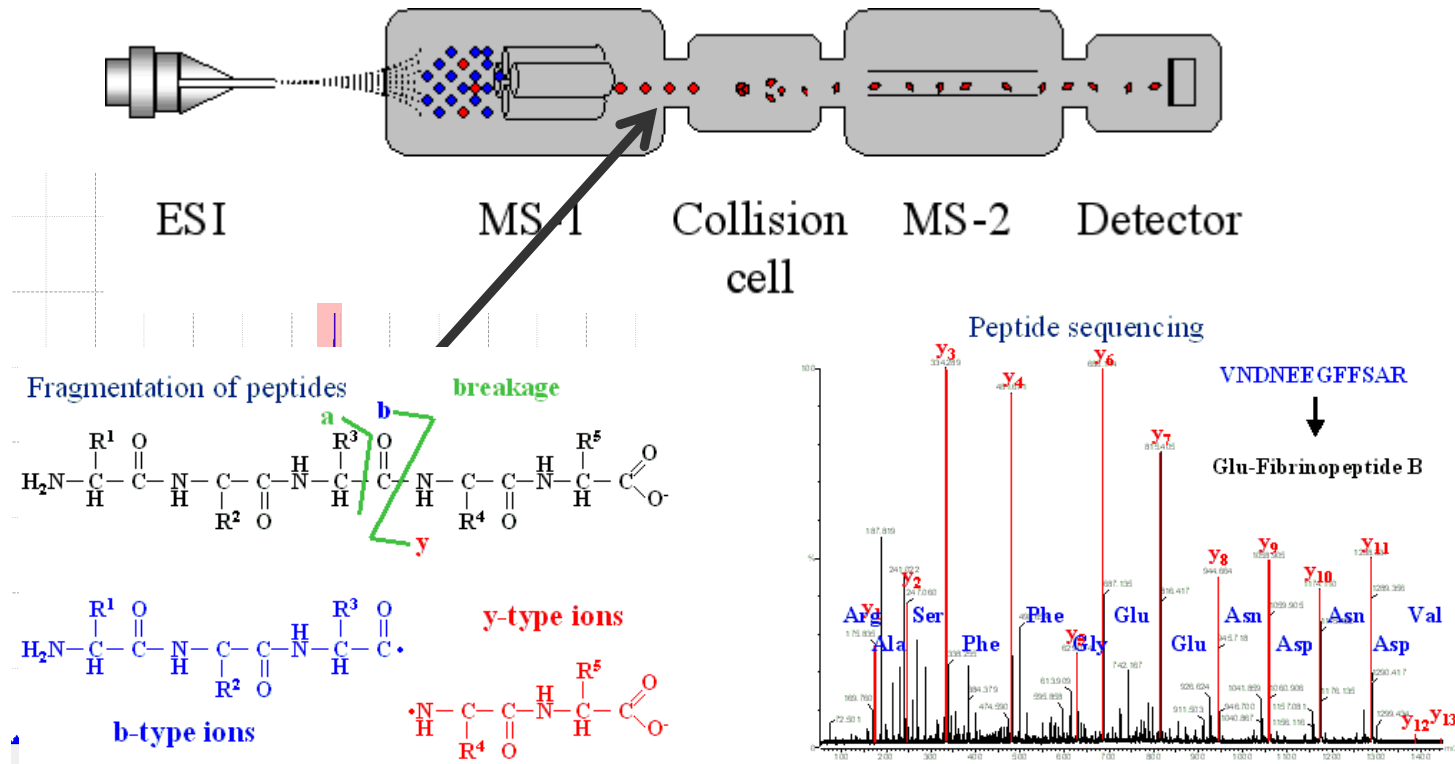
Wenzhu Zhang\* and Brian T. Chait\*



# Talk: Identification

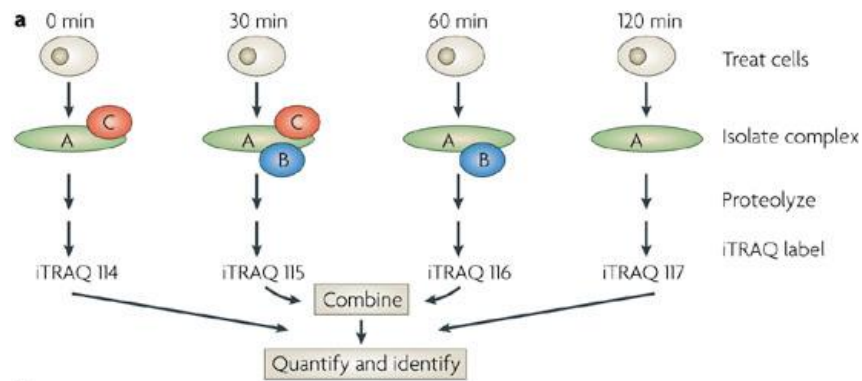
## InsPecT: Identification of Posttranslationally Modified Peptides from Tandem Mass Spectra

Stephen Tanner,<sup>\*,†</sup> Hongjun Shu,<sup>‡</sup> Ari Frank,<sup>§</sup> Ling-Chi Wang,<sup>||</sup> Ebrahim Zandi,<sup>||</sup> Marc Mumby,<sup>‡</sup> Pavel A. Pevzner,<sup>§</sup> and Vineet Bafna<sup>§</sup>

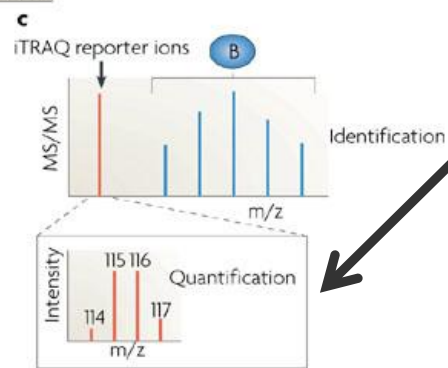
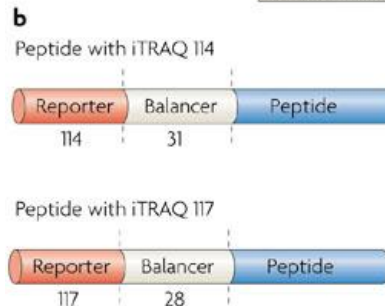


# Talk: Multiplexing using iTRAQ

Isobaric tag for relative and absolute quantitation



← Peptides!



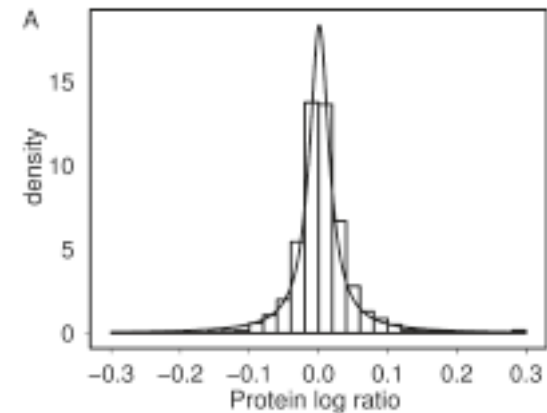
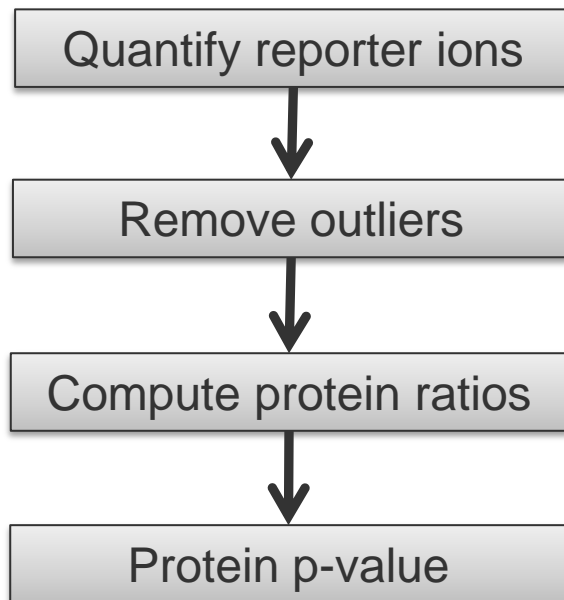
Nature Reviews | Molecular Cell Biology



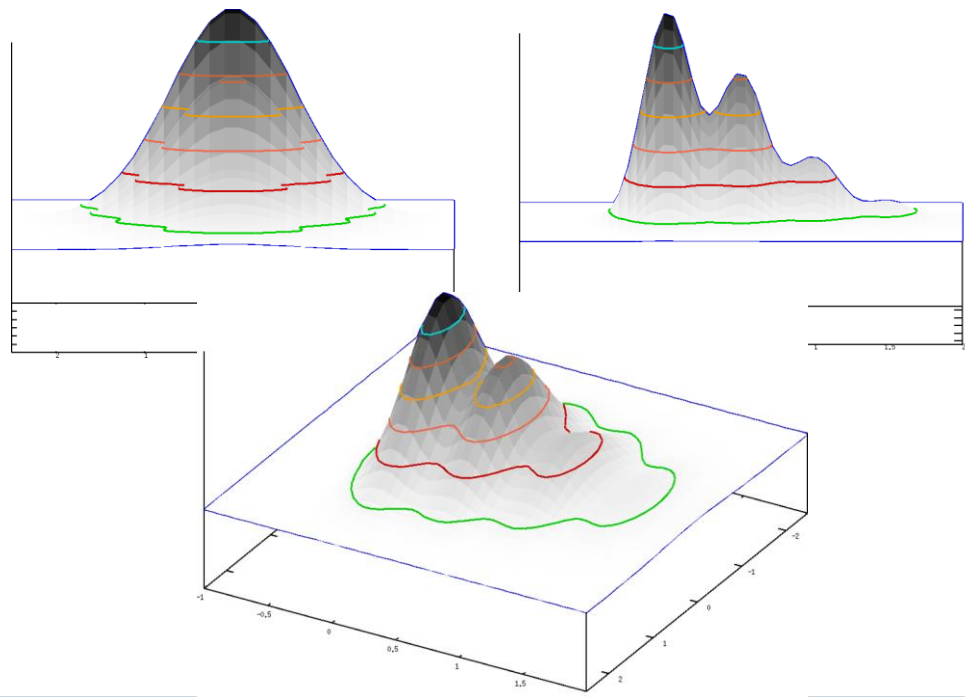
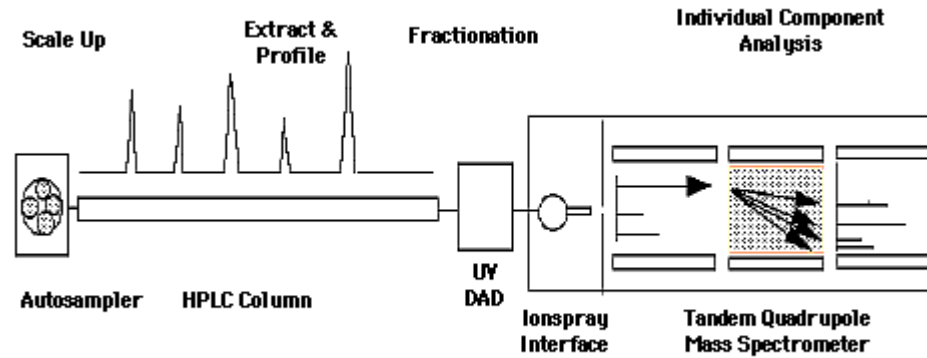
# Talk: Multiplexing using iTRAQ

## General Statistical Modeling of Data from Protein Relative Expression Isobaric Tags

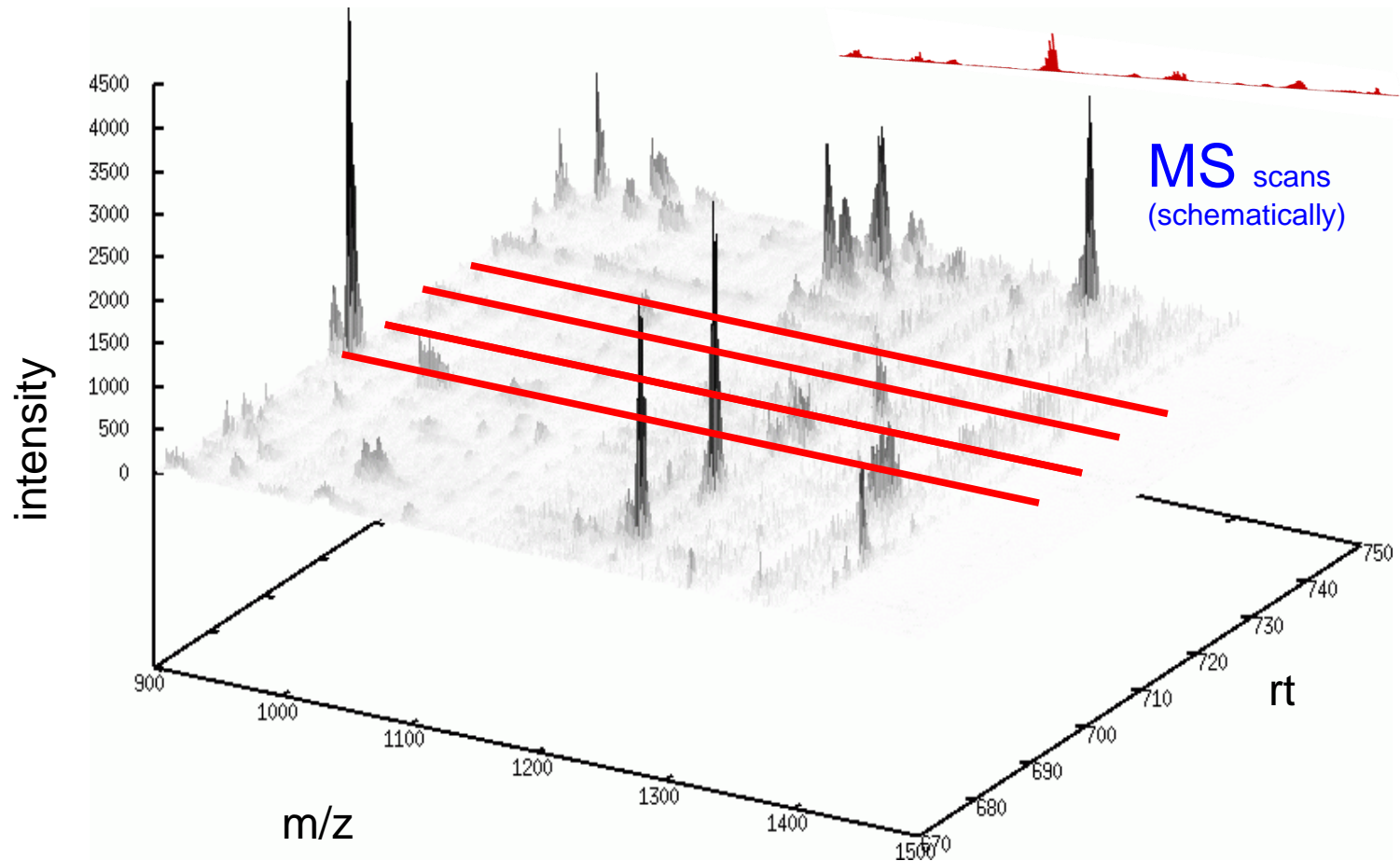
Florian P. Breitwieser,<sup>†</sup> André Müller,<sup>†</sup> Loïc Dayon,<sup>‡</sup> Thomas Köcher,<sup>¶</sup> Alexandre Hainard,<sup>‡</sup> Peter Pichler,<sup>§</sup> Ursula Schmidt-Erfurth,<sup>||</sup> Giulio Superti-Furga,<sup>†</sup> Jean-Charles Sanchez,<sup>‡</sup> Karl Mechtler,<sup>¶</sup> Keiryn L. Bennett,<sup>†</sup> and Jacques Colinge<sup>\*†</sup>

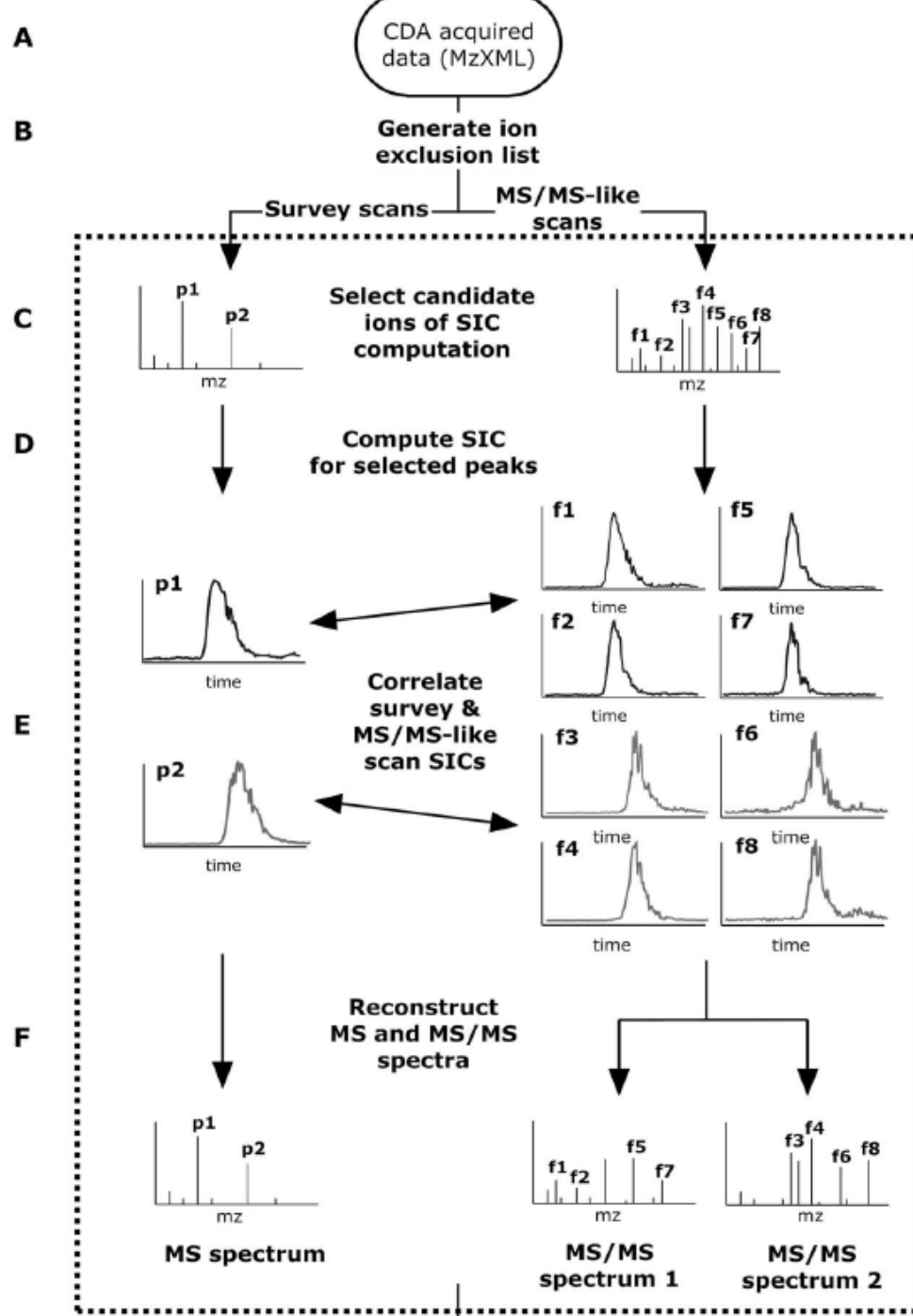


# Identification: ETISEQ

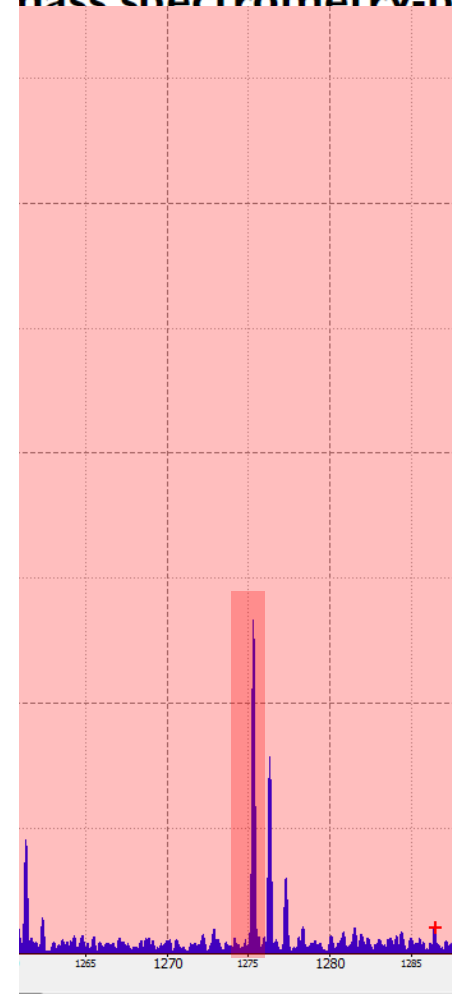


# Identification: ETISEQ





## Retention time ion sequencing mass spectrometry-based



# Talk: Experimental Design

## Statistical Design of Quantitative Mass Spectrometry-Based Proteomic Experiments

Ann L. Oberg<sup>†</sup> and Olga Vittek<sup>\*†</sup>

(a) Balanced Incomplete Block

Disease group	Replicate set 1										...
	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	
$D_1$	$X_{L1}$	$X_{L2}$	$X_{L1}$	$X_{L2}$							...
$D_2$	$X_{L2}$				$X_{L1}$	$X_{L2}$	$X_{L1}$				...
$D_3$		$X_{L1}$			$X_{L2}$			$X_{L1}$	$X_{L2}$		...
$D_4$			$X_{L2}$			$X_{L1}$		$X_{L2}$		$X_{L1}$	...
$D_5$				$X_{L1}$			$X_{L2}$		$X_{L1}$	$X_{L2}$	...

(b) Reference

Disease group	Replicate set 1					...
	Block 1	Block 2	Block 3	Block 4	Block 5	
$R$	$R_{L1}$	$R_{L1}$	$R_{L1}$	$R_{L1}$	$R_{L1}$	...
$D_1$	$X_{L2}$					...
$D_2$		$X_{L2}$				...
$D_3$			$X_{L2}$			...
$D_4$				$X_{L2}$		...
$D_5$					$X_{L2}$	...

(c) Loop

Disease group	Replicate set 1					...
	Block 1	Block 2	Block 3	Block 4	Block 5	
$D_1$	$X_{L1}$				$X_{L2}$	...
$D_2$	$X_{L2}$	$X_{L1}$				...
$D_3$		$X_{L2}$	$X_{L1}$			...
$D_4$			$X_{L2}$	$X_{L1}$		...
$D_5$				$X_{L2}$	$X_{L1}$	...