# Chap 7 Mass Spectrometry

Biochemistry and Molecular Biology

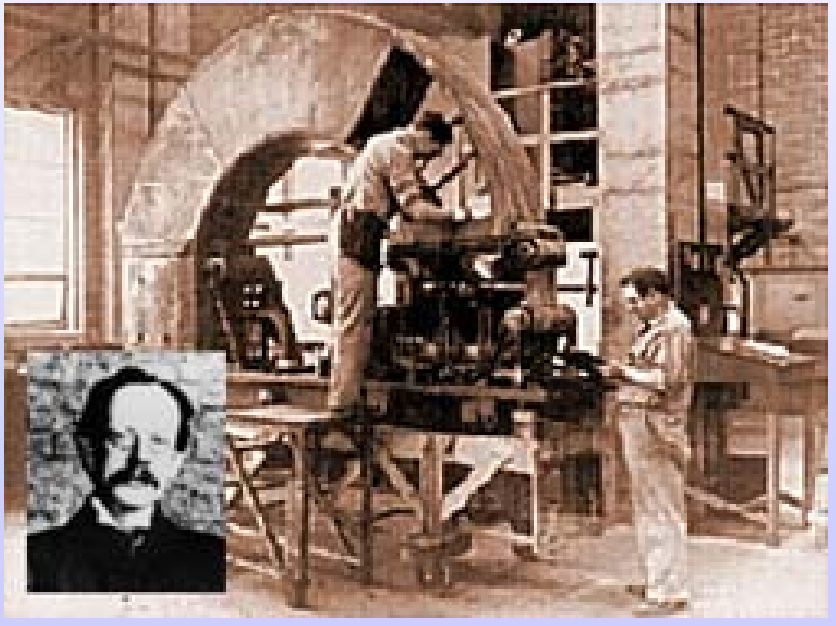
- 9.1 Introduction
- 9.2 Ionization
- 9.3 Mass Analysis
- 9,5 Structural Information by Tandem Mass Spectrometry
- 9.7 Computing and Database Analysis

# Biological Mass Spectrometry

A key tool for Protein. Peptides. Carbohydrate. DNA……and more

# Back to the history……

The first mass spectrometer - parabola spectrograph



1912 J. J. Thomson

# Back to the history……


J.J. Thomson

- J.J. Thomson observes a line at mass 22 in the spectrum of neon.
- J.J. Thomson delivers his Bakerian Lecture, “Rays of Positive Electricity” to the Royal Society of London.

- ***Francis Aston*** is awarded the Nobel Prize in chemistry for his discovery of isotopes of "inactive elements."

Francis Aston

| 1 | | | | | | | | | | | | | | | | | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 H 1.0079 | 2 | | | | | | | | | | | 13 | 14 | 15 | 16 | 17 | 2 He 4.0026 |
| 3 Li 6.941 | 4 Be 9.0122 | | | | | | | | | | | 5 B 10.811 | 6 C 12.011 | 7 N 14.007 | 8 O 15.999 | 9 F 18.998 | 10 Ne 20.180 |
| 11 Na 22.990 | 12 Mg 24.305 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 Al 26.982 | 14 Si 28.086 | 15 P 30.974 | 16 S 32.065 | 17 Cl 35.453 | 18 Ar 39.948 |
| 19 K 39.098 | 20 Ca 40.078 | 21 Sc 44.956 | 22 Ti 47.867 | 23 V 50.942 | 24 Cr 51.996 | 25 Mn 54.938 | 26 Fe 55.845 | 27 Co 58.933 | 28 Ni 58.693 | 29 Cu 63.546 | 30 Zn 65.409 | 31 Ga 69.723 | 32 Ge 72.64 | 33 As 74.922 | 34 Se 78.96 | 35 Br 79.904 | 36 Kr 83.798 |
| 37 Rb 85.468 | 38 Sr 87.62 | 39 Y 88.906 | 40 Zr 91.224 | 41 Nb 92.906 | 42 Mo 95.94 | 43 Tc (98) | 44 Ru 101.07 | 45 Rh 102.91 | 46 Pd 106.42 | 47 Ag 107.87 | 48 Cd 112.41 | 49 In 114.82 | 50 Sn 118.71 | 51 Sb 121.76 | 52 Te 127.60 | 53 I 126.90 | 54 Xe 131.29 |
| 55 Cs 132.91 | 56 Ba 137.33 | 57-71 * | 72 Hf 178.49 | 73 Ta 180.95 | 74 W 183.84 | 75 Re 186.21 | 76 Os 190.23 | 77 Ir 192.22 | 78 Pt 195.08 | 79 Au 196.97 | 80 Hg 200.59 | 81 Tl 204.38 | 82 Pb 207.2 | 83 Bi 208.98 | 84 Po (209) | 85 At (210) | 86 Rn (222) |
| 87 Fr (223) | 88 Ra (226) | 89-103 # | 104 Rf (261) | 105 Db (262) | 106 Sg (266) | 107 Bh (264) | 108 Hs (277) | 109 Mt (268) | 110 Ds (281) | 111 Rg (272) | 112 Uub (285) | 113 Uut (284) | 114 Uuq (289) | 115 Uup (288) | | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| * Lanthanide series | 57 La 138.91 | 58 Ce 140.12 | 59 Pr 140.91 | 60 Nd 144.24 | 61 Pm (145) | 62 Sm 150.36 | 63 Eu 151.96 | 64 Gd 157.25 | 65 Tb 158.93 | 66 Dy 162.50 | 67 Ho 164.93 | 68 Er 167.26 | 69 Tm 168.93 | 70 Yb 173.04 | 71 Lu 174.97 |
| # Actinide series | 89 Ac (227) | 90 Th 232.04 | 91 Pa 231.04 | 92 U 238.03 | 93 Np (237) | 94 Pu (244) | 95 Am (243) | 96 Cm (247) | 97 Bk (247) | 98 Cf (251) | 99 Es (252) | 100 Fm (257) | 101 Md (258) | 102 No (259) | 103 Lr (262) |

# Advances in Modern Mass Spectrometry

Limitations of traditional MS on biological applications

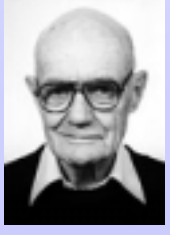
- High molecular weight >50,000
- Amount of Sample $< 10^{-12} - 10^{-15}$ mole

Intact Molecule

Non-covalent Complex

**E**lectro**S**pray **I**onization MS

**M**atrix **A**ssisted **L**aser **D**esorption **I**onization MS



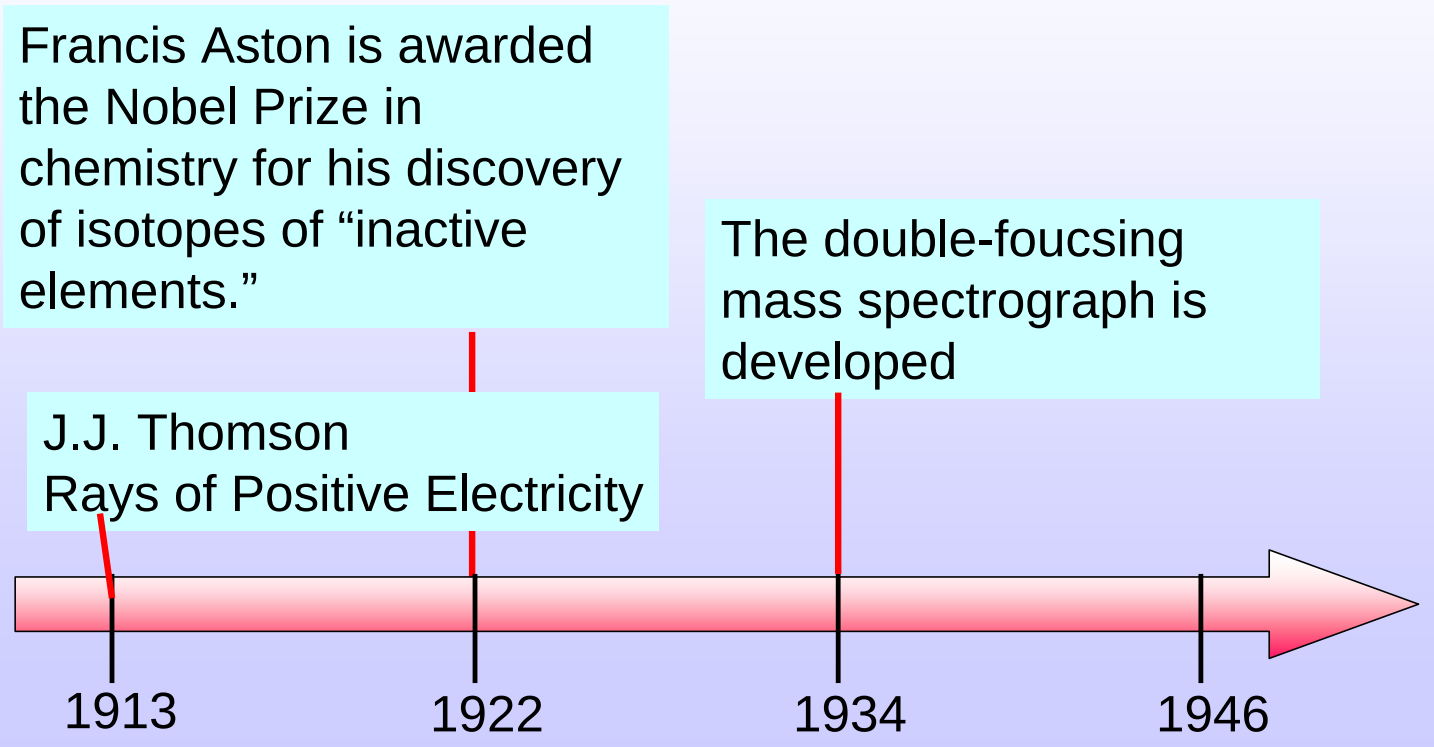
John B. Fenn

ESI

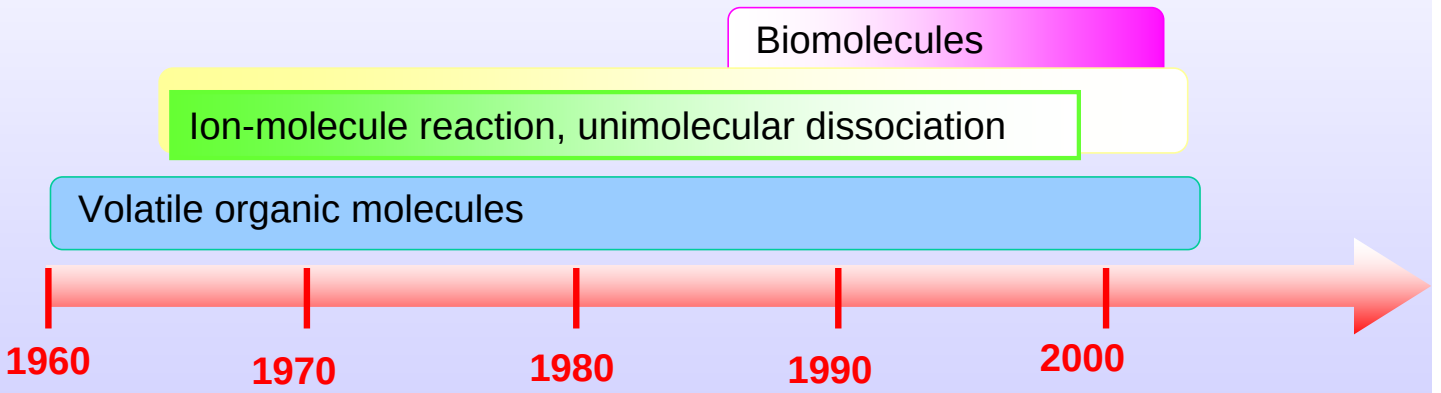
Koichi Tanaka

MALDI

# Evolution of Mass Spectrometry-1



# Evolution of Mass Spectrometry-2



# Human Genome Project

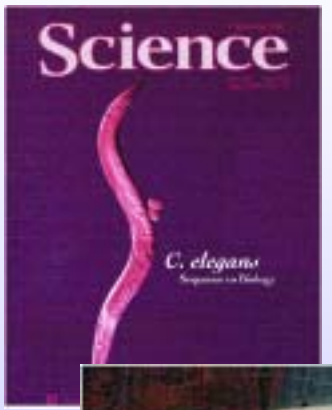

2003, 4, 14

35000-40000 Genes

# Genomics（基因體學）

1953: Watson and Crick: DNA double helix

*C. Elegans* (1998)




*Rattus norvegicus*

*Danio rerio*

*Anopheles gambiae* (2002

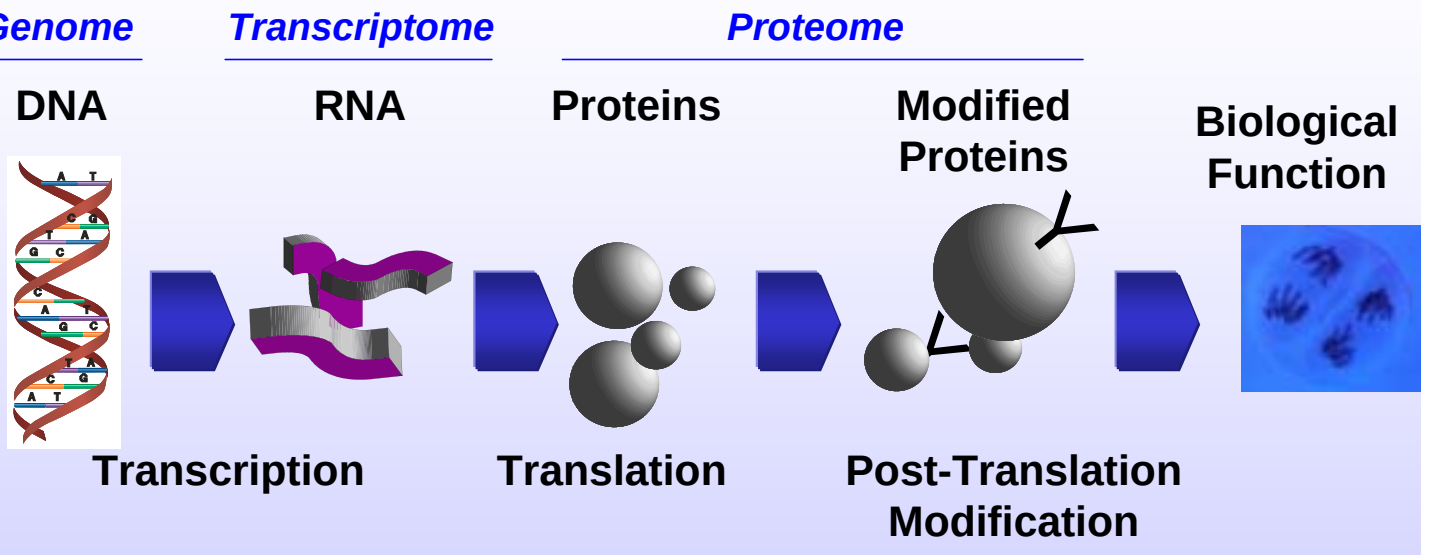
*Fugu rubripes* (2002)

*Homo sapiens* (2001)

*Drosophila melanogaster* (2000)

*Mus musculus* (2002)

# From Genomics to Proteomics




**PROTE**in complement to a gen**OME**

Entire protein complement in a given cell, tissue or organism

# PR  TE  MICS

## 蛋白質體學

Protein activity, modifications, localizations, amd interactions of proteins in complexes

***Proteomics*** can be defined as *the qualitative and quantitative comparison of proteomes under different conditions to further unravel biological processes*

# Protein chemistry and proteomics

## Protein Identification
## ( 蛋白質鑑定)

| Protein chemistry | Proteomics |
|---|---|
| Individual proteins | Complex mixtures |
| Complete sequence analysis | Partial sequence analysis |
| Emphasis on structure and function | Emphasis on identification by database matching |
| Structual biology | Systems biology |

# Disease Proteomics

## 一滴血可以診斷出疾病嗎?

**Use of proteomic patterns in serum to identify ovarian cancer**

*Emanuel F Petricoin III, Ali M Ardekani, Ben A Hitt, Peter J Levine, Vincent A Fusaro, Seth M Steinberg, Gordon B Mills, Charles Simone, David A Fishman, Elise C Kohn, Lance A Liotta*

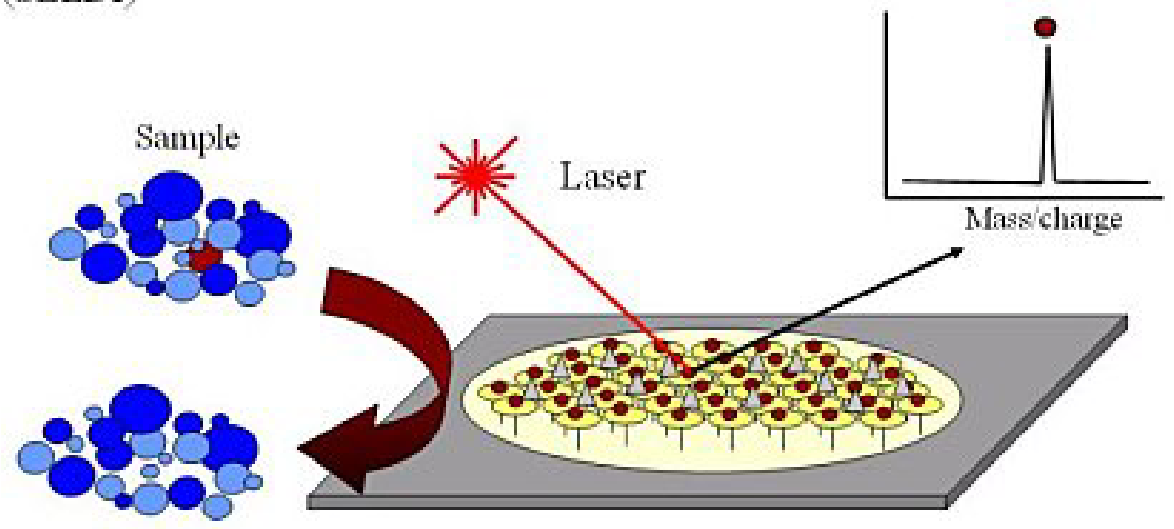
**Background** New technologies for the detection of early-stage ovarian cancer are urgently needed. Pathological changes within an organ might be reflected in proteomic patterns in serum. We developed a bioinformatics tool and used it to identify proteomic patterns in serum that distinguish neoplastic from non-neoplastic disease within the ovary.

572 THE LANCET · Vol 359 · February 16, 2002 · www.thelancet.com

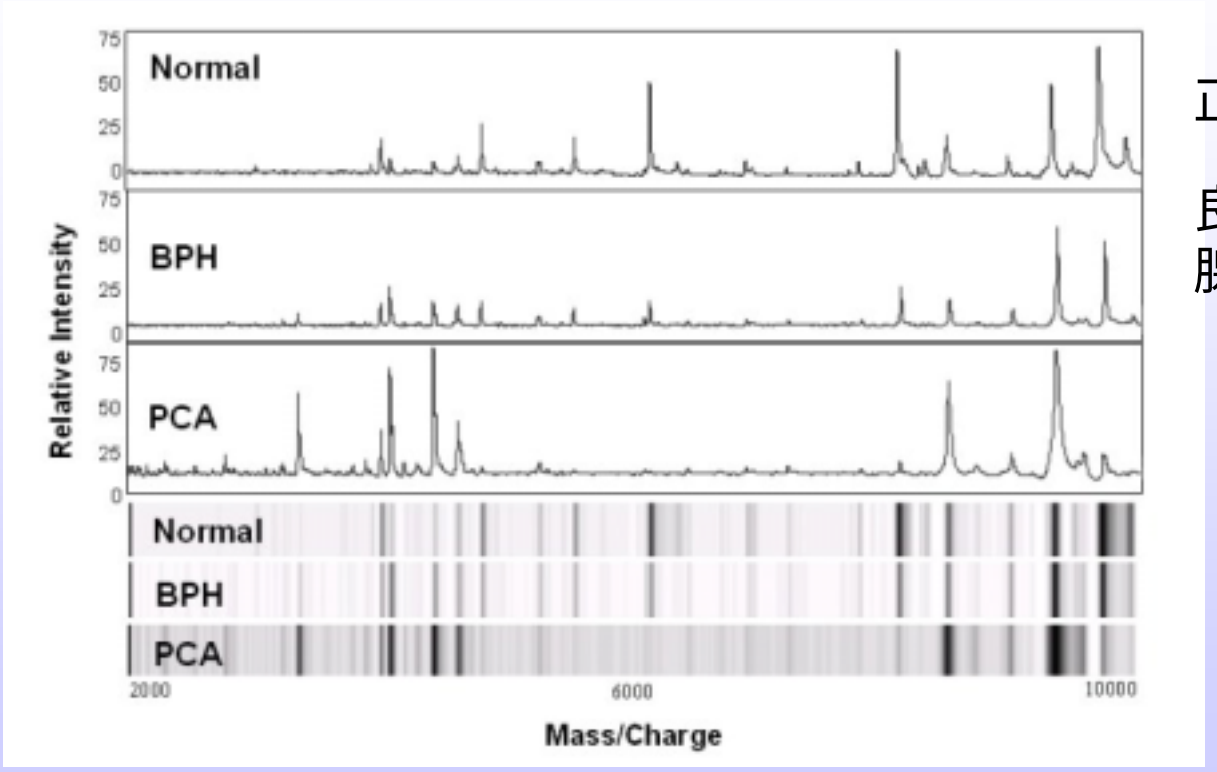
## *Protein Chips* 蛋白質晶片

### SELDI ProteinChip™ Technology Process

- Sample goes *directly* onto the ProteinChip™ Array
- Proteins● are captured and *retained* on the chip (affinity capture )
- EAM△ is added to the chip
- Retentate map is "read" by Surface-Enhanced Laser Desorption/Ionization (SELDI)



# 找尋腫瘤標記蛋白質




正常人

良性攝護腺腫瘤

攝護腺癌

# Imaging Mass Spectrometry

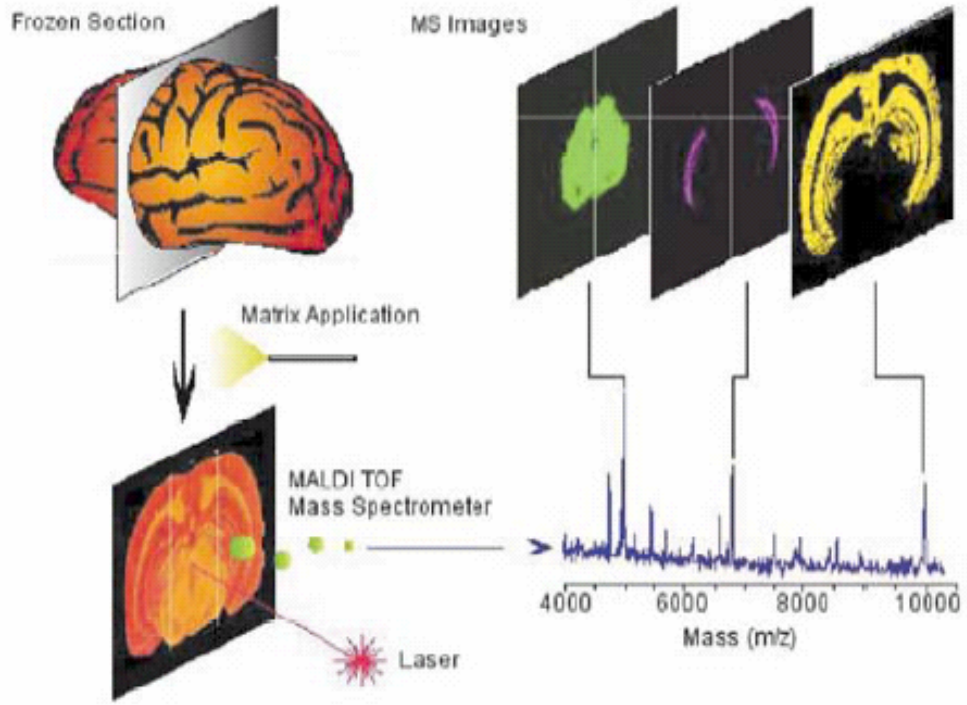
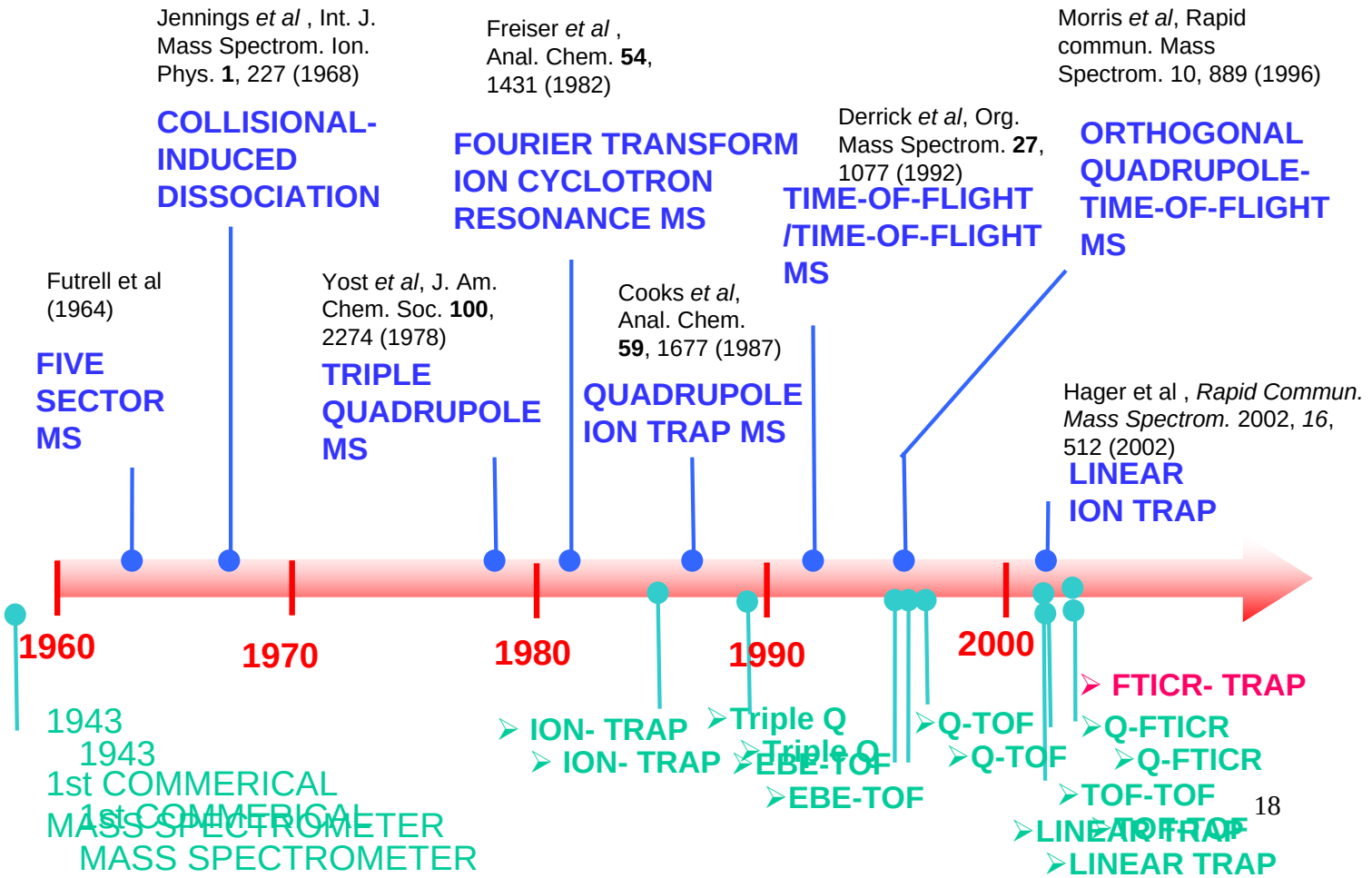


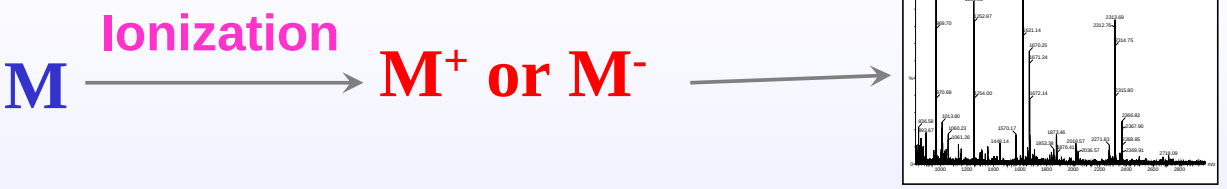

**Fig. 1** Methodology developed for the spatial analysis of tissue by MALDI mass spectrometry. Frozen sections are mounted on a metal plate, coated with an UV-absorbing matrix and placed in the mass spectrometer. A pulsed UV laser desorbs and ionizes analytes from the tissue and their m/z values are determined using a time-of-flight analyzer. From a raster over the tissue and measurement of the peak intensities over thousands of spots, mass spectrometric images are generated at specific molecular weight values.

Nat. Med. **2001**, *7*, 493-496

# Wide Applications of Mass Spectrometry



# What is mass spectrometry?



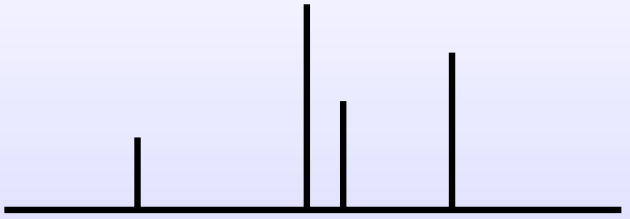

Mass-to-charge Ratio ( ***m/z*** )

MS is an analytical tool that measure the molecular weight of molecules based on the motion of charged particle in an electrical or magnetic field.

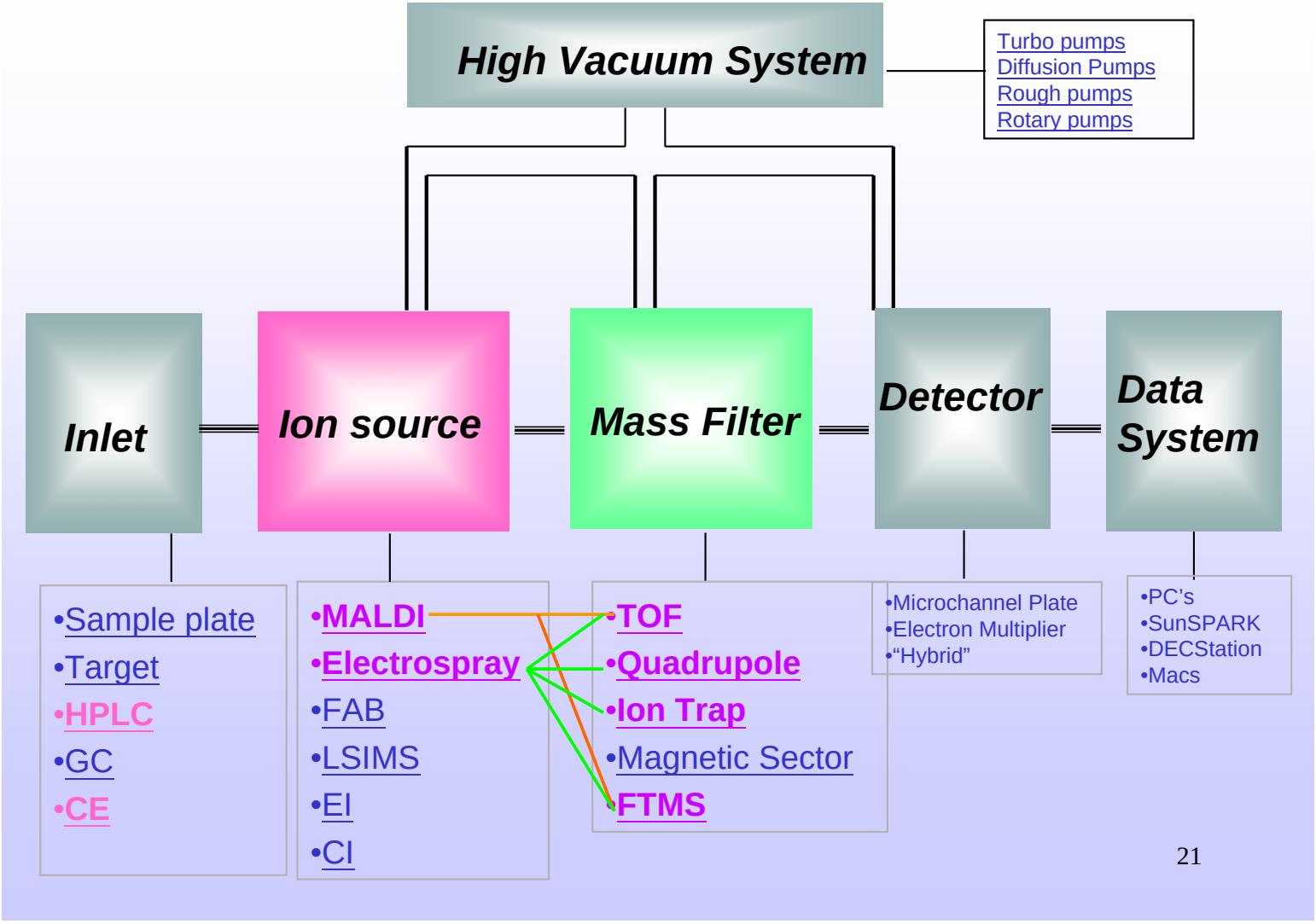
# Two different ways of measurement

- Composition of analyte (MS)

  e.g. Peptide mass fingerprinting



- Structure of analyte (MS/MS, tandem MS)


High Vacuum System
Turbo pumps
Diffusion Pumps
Rough pumps
Rotary pumps
Inlet
Ion source
Mass Filter
Detector
Data System
•Sample plate
•Target
•HPLC
•GC
•CE
•MALDI
•Electrospray
•FAB
•LSIMS
•EI
•CI
•TOF
•Quadrupole
•Ion Trap
•Magnetic Sector
•FTMS
•Microchannel Plate
•Electron Multiplier
•"Hybrid"
•PC's
•SunSPARK
•DECStation
•Macs

# Ionization Methods

- Electron Impact (EI)
- Fast Atom Bombardment (FAB)
- Electrospray Ionization (ESI)
- Matrix-Assisted Laser Desorption Ionization (MALDI)

# Advances in Modern Mass Spectrometry

Limitations of traditional MS on biological applications

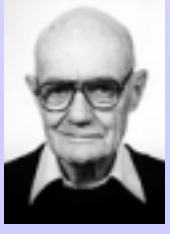
- High molecular weight >50,000
- Amount of Sample $< 10^{-12} - 10^{-15}$ mole

Intact Molecule

Non-covalent Complex

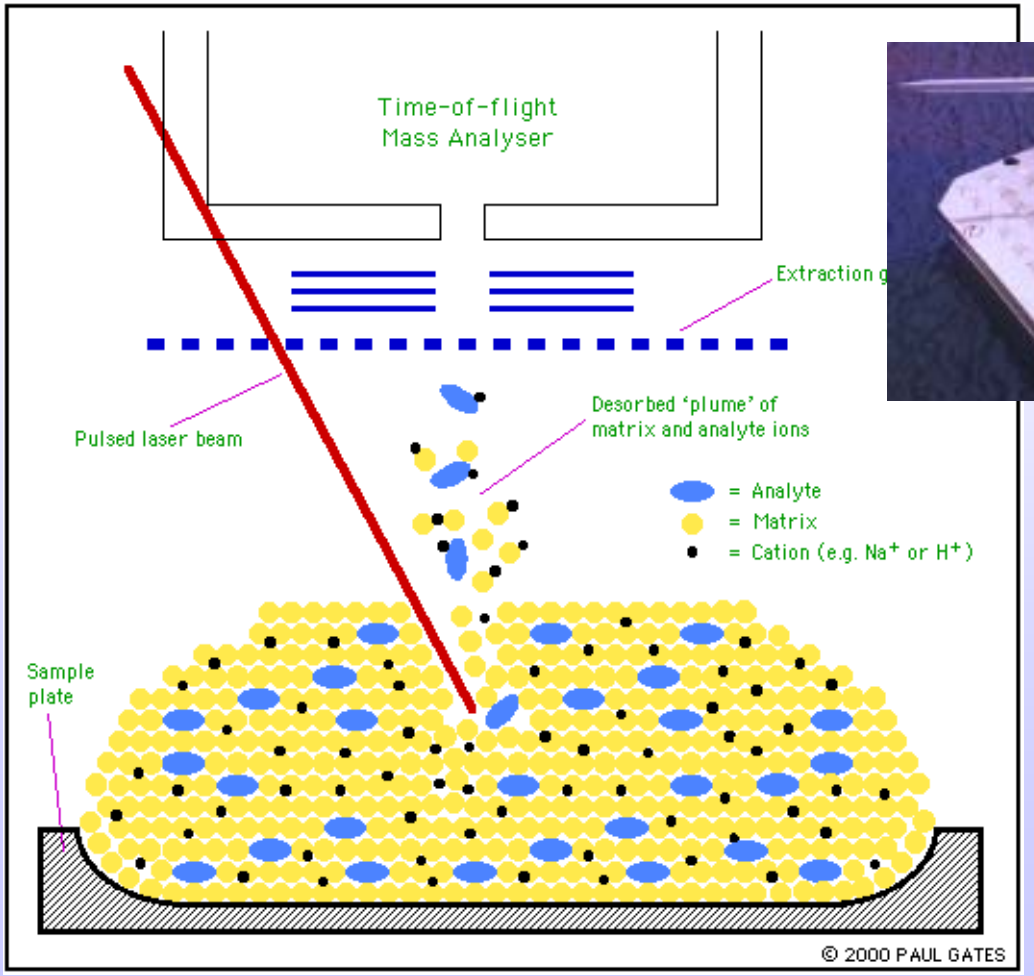
**E**lectro**S**pray **I**onization MS

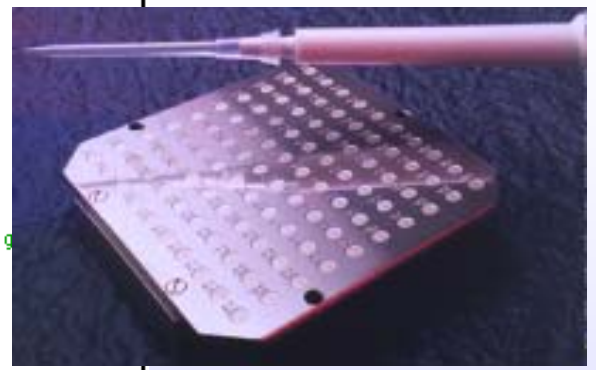
**M**atrix **A**ssisted **L**aser **D**esorption **I**onization MS



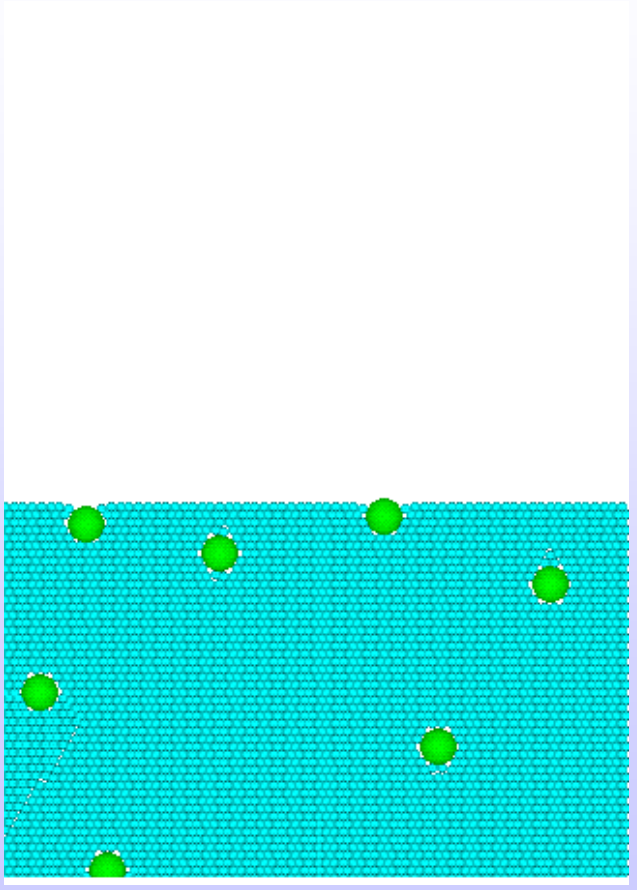
2002 Nobel Prize in Chemistry

# *M*atrix-*A*ssisted *L*aser *D*esorption *I*onization (*MALDI*)



# *M*atrix-*A*ssisted *L*aser *D*esorption *I*onization (*MALDI*)



Analyte

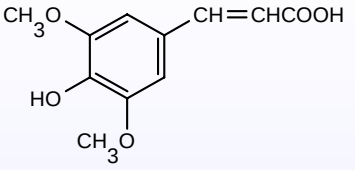
Metal Ion

Matrix

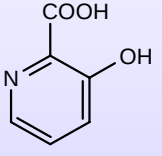
*http*://www.ionsource.com

# Matrix selection

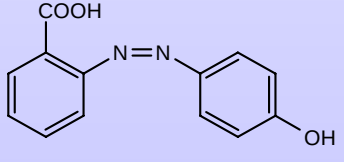
| | |
|---|---|
| α-Cyano-4-hydroxy-cinnamic acid (CHCA) | Peptides<10kDa |
| Sinapinic Acid | Proteins >10kDa |
| 2,5-Dihydroxybenzoic acid (DHB) | Neutral Carbohydrates, Synthetic Polymers |
| "Super DHB" | Proteins, Glycosylated proteins |
| 3-Hydroxypicolinic acid | Oligonucleotides |
| HABA | Proteins, Oligosaccharides |



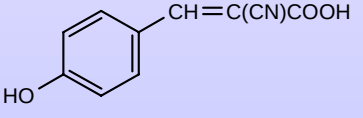

Sinapinic acid (3,5-Dimethoxy-4-hydroxy cinnamic acid)




3-hydroxypicolinic acid (3-HPA)

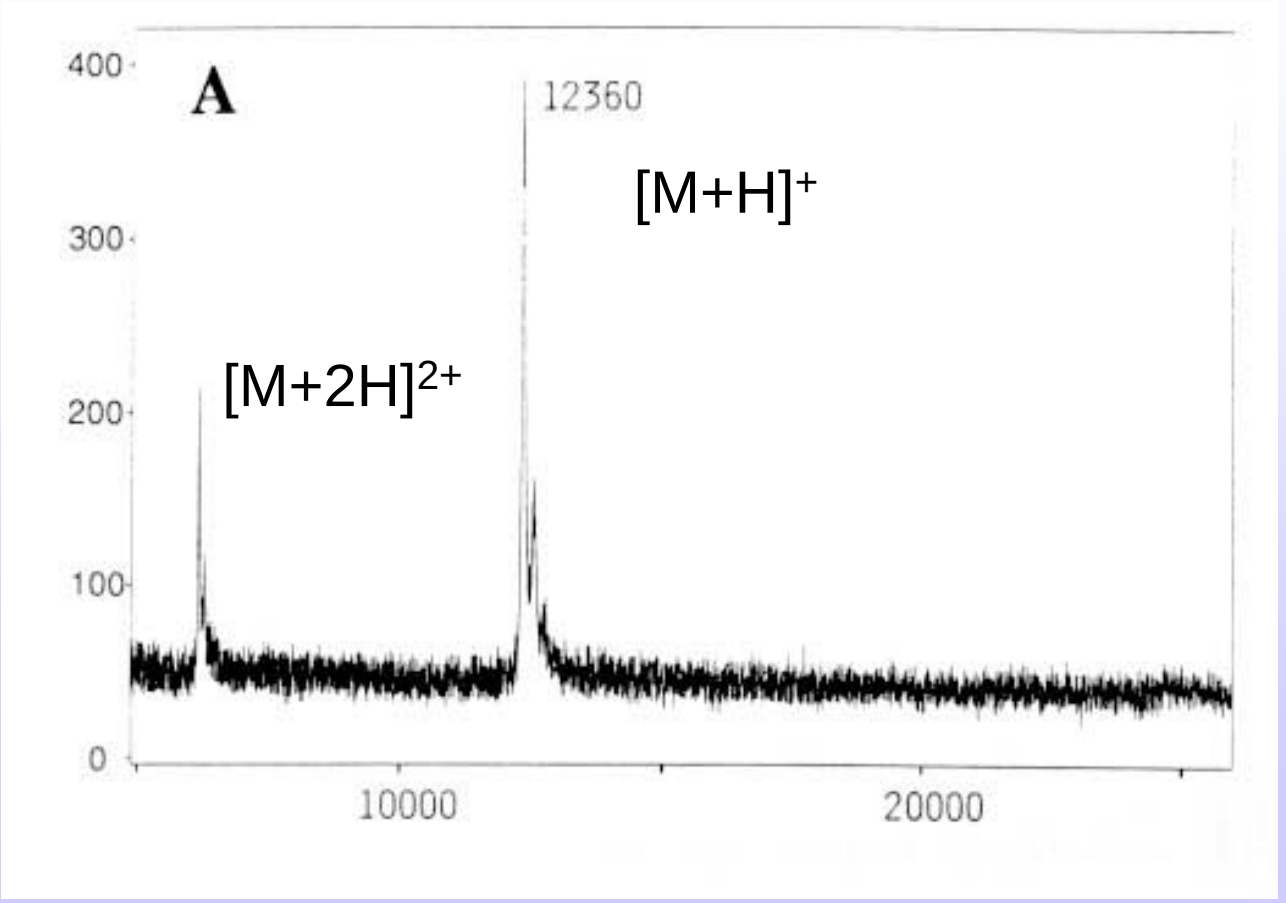



2-(4-hydroxyphenylazo)-benzoic acid (HABA)

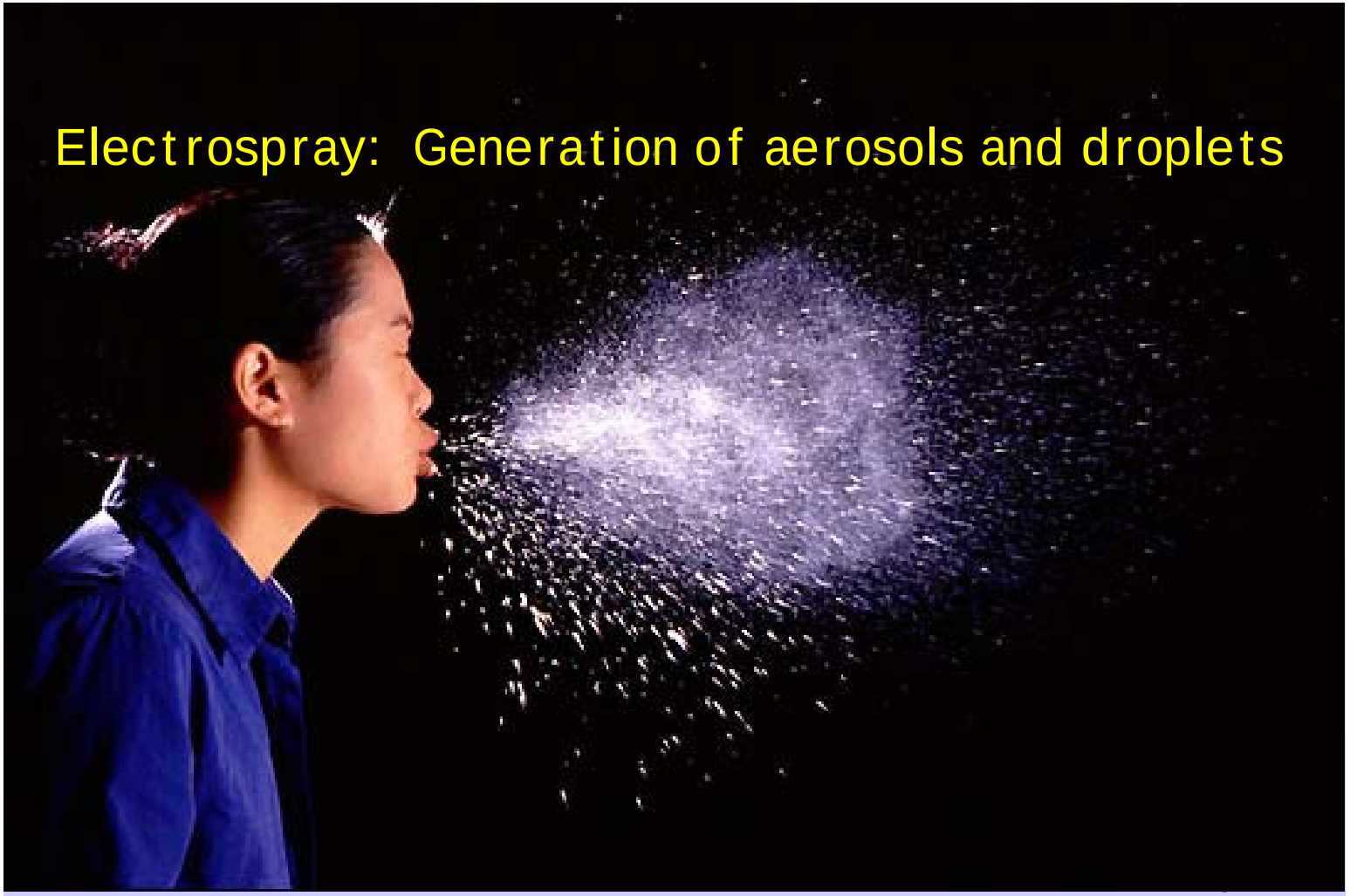



α-cyano-4-hydroxycinnamic acid

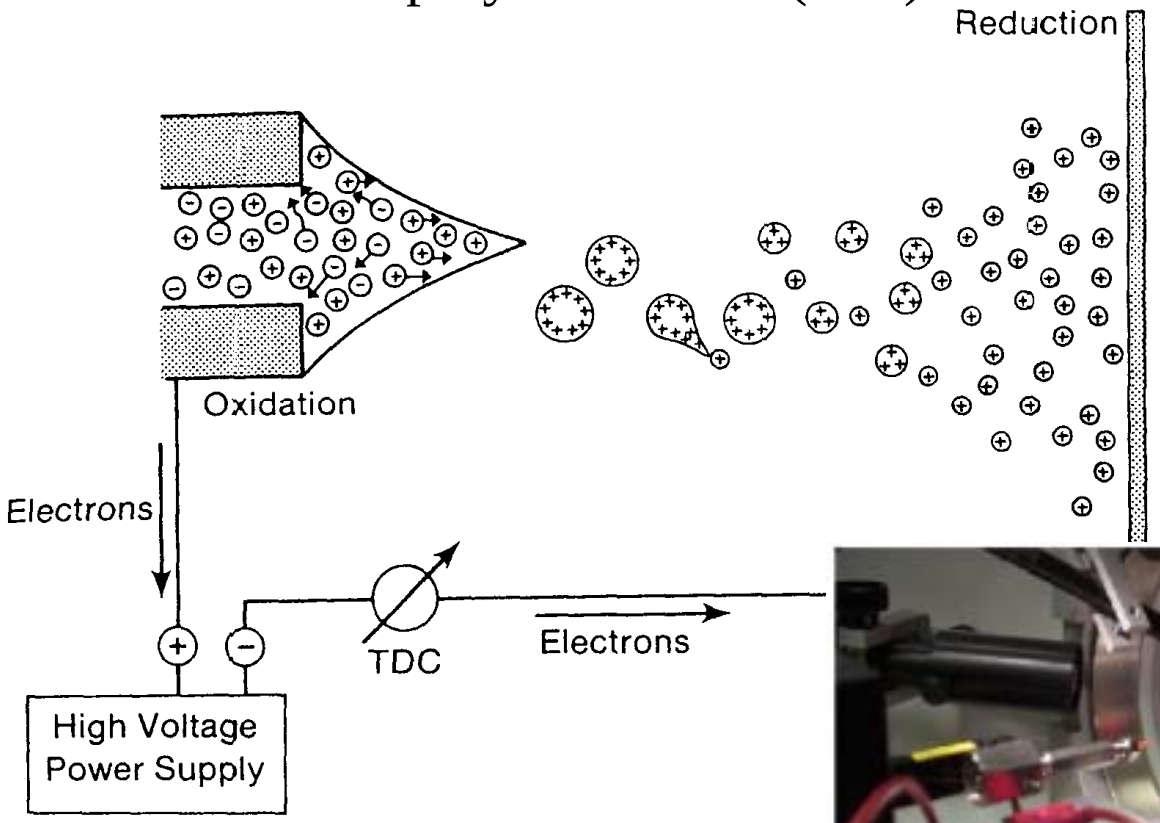
# A typical mass spectra

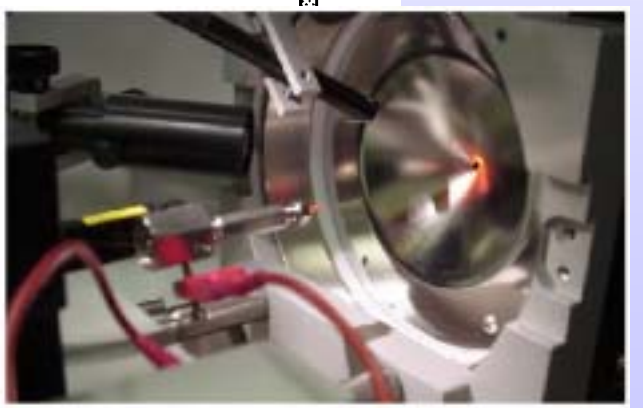


# Electrospray: Generation of aerosols and droplets

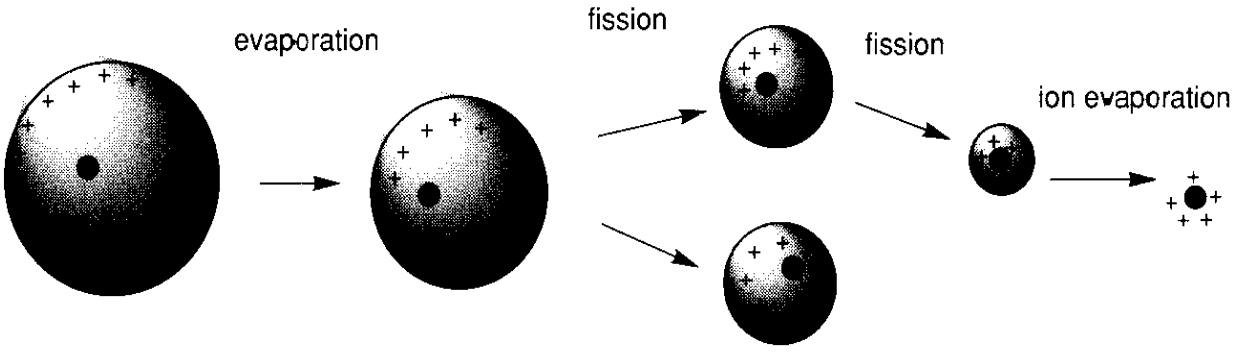


# *E*lectro*S*pray *I*onization (**ESI**)



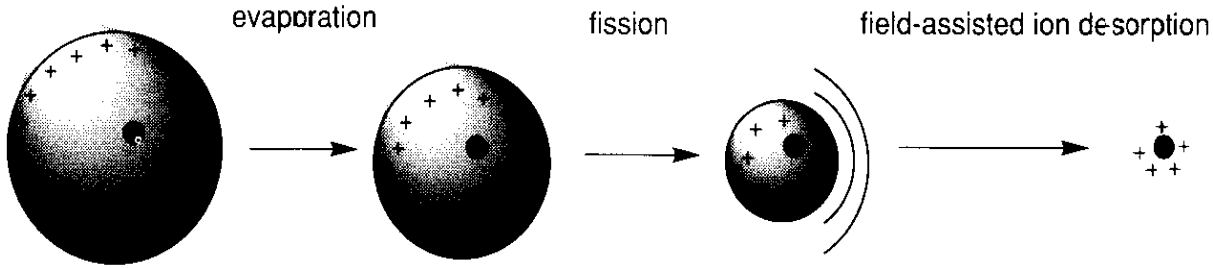



## *Charge Residue Model*




(a)

Coulomb repulsion : Charge repulsion > surface tension

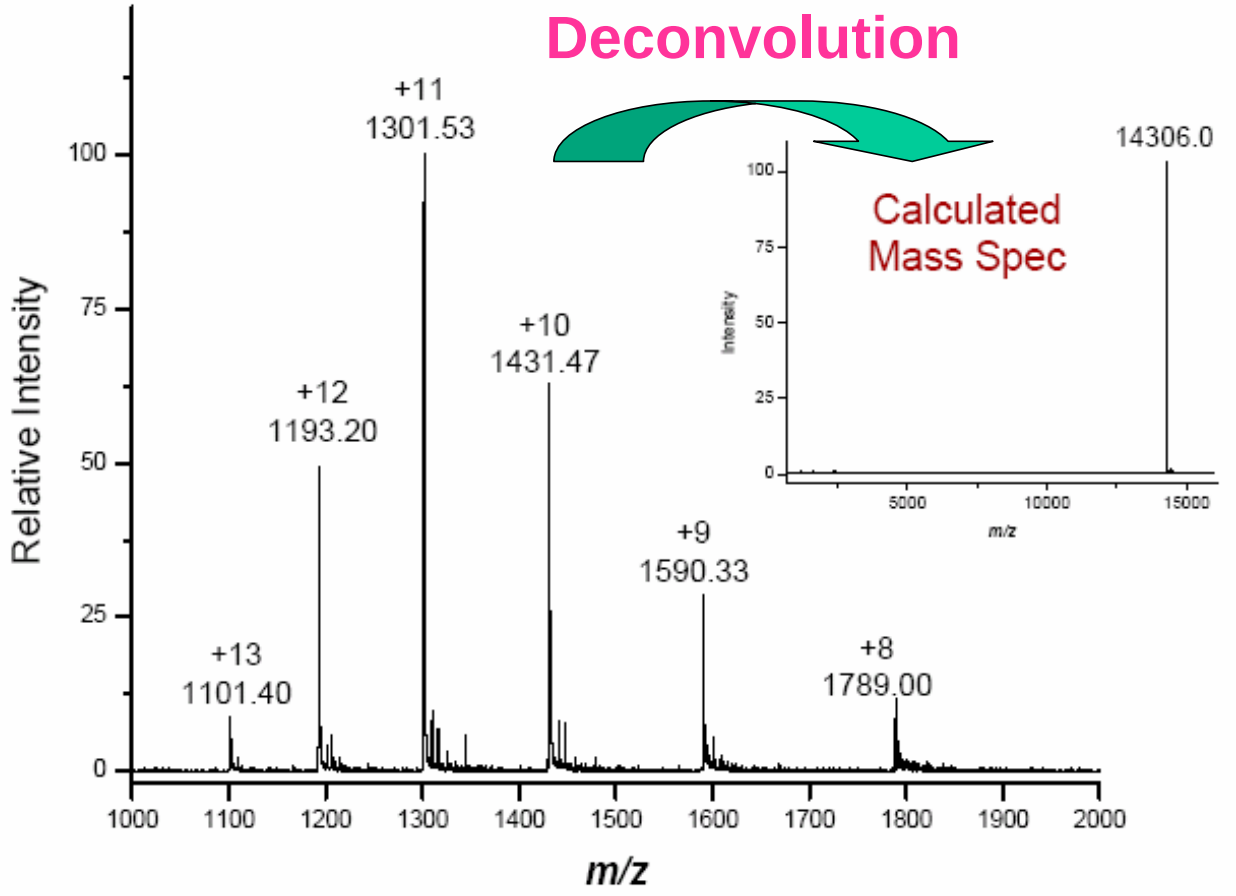


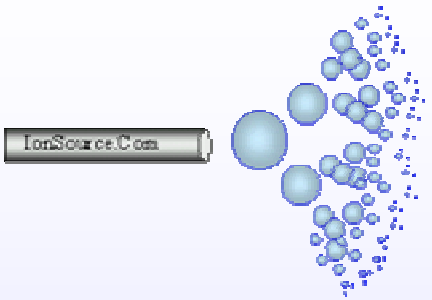

(b)

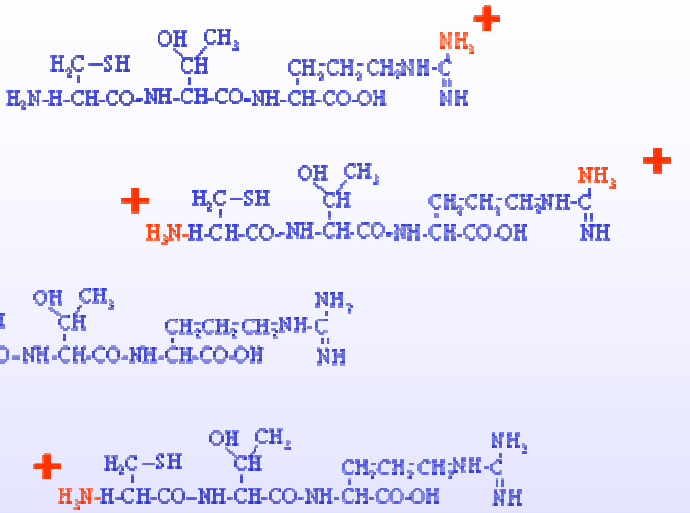
## *Ion Desorption Model*

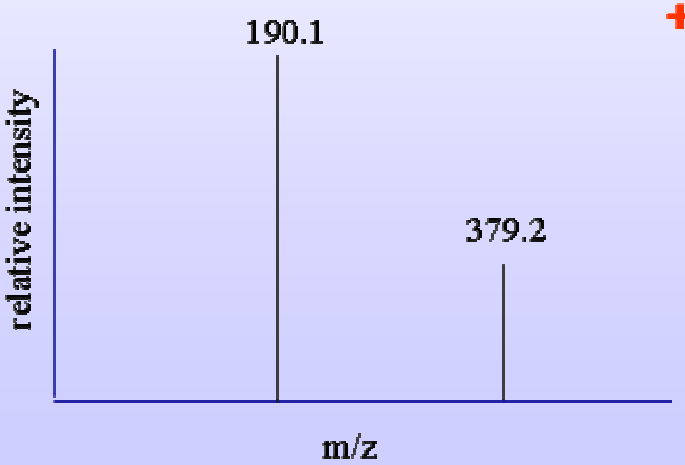
Ion desorption from the droplet surface

# Multiple-charged ESI spectra

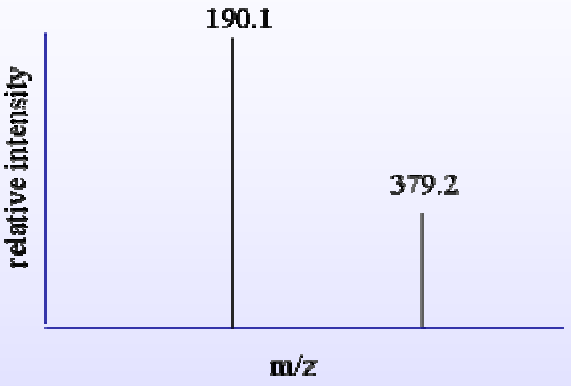



IonSource.Com




190.1
379.2
relative intensity
m/z

# Deconvolution




Step 1: assume 190.1 = mass/charge

379.2 = mass/charge

Step 2: assume the two peaks are related

190.1 = m/(z+1)

379.2=m/z

Step 3: solve m and z

m=378.2, z=1

| Charge state | Calculation | Unprotonated mass |
|---|---|---|
| +1 | (379.2-1)*1 | 378.2 |
| +2 | (190.1-1)*2 | 378.2 |
| | average | 378.2 |

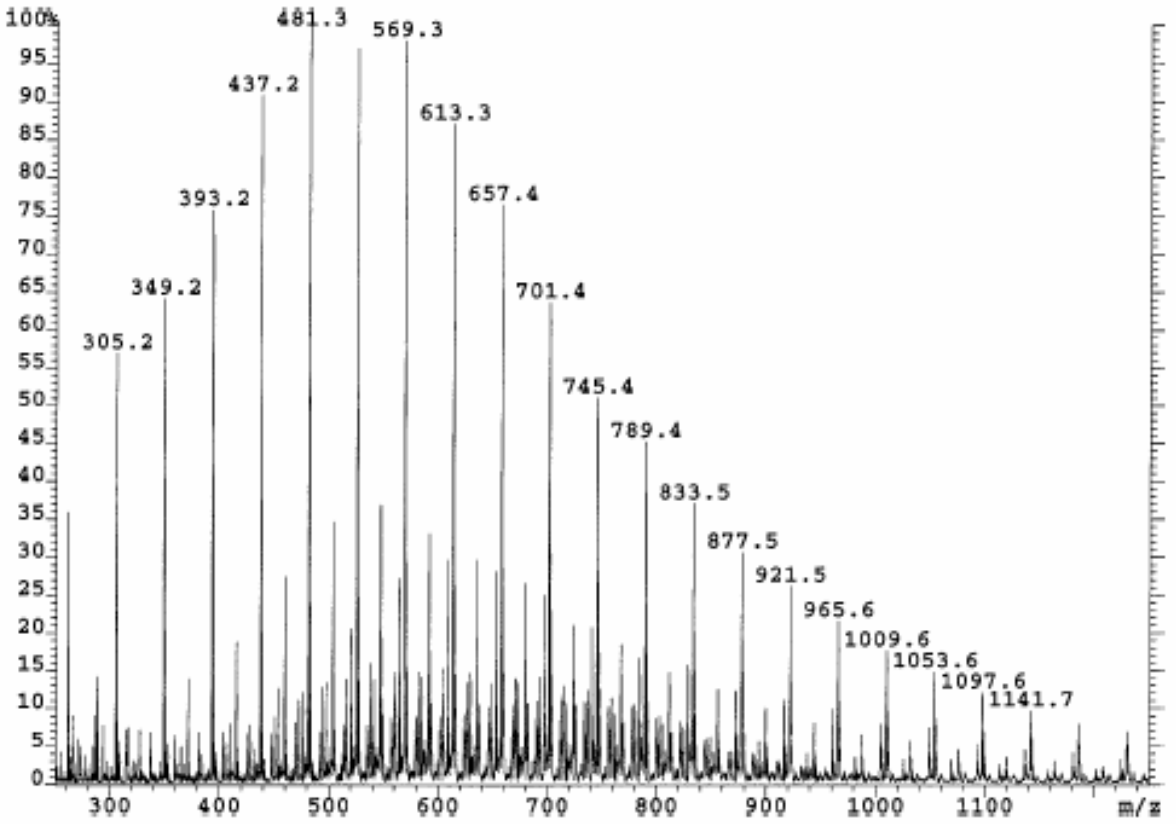
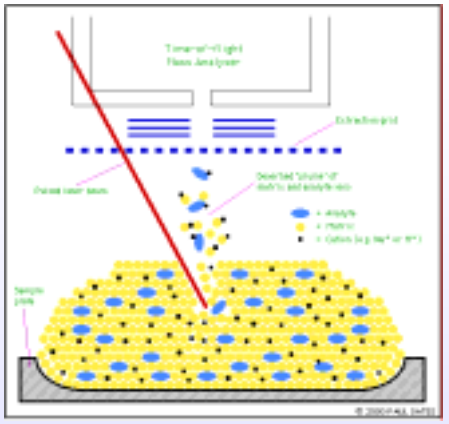
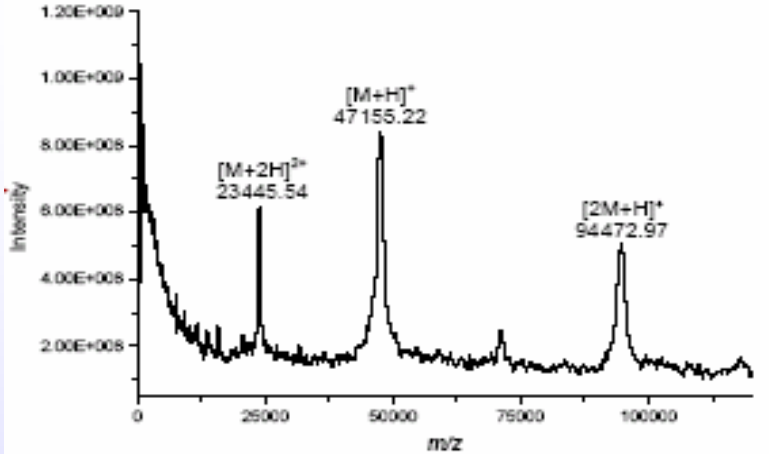


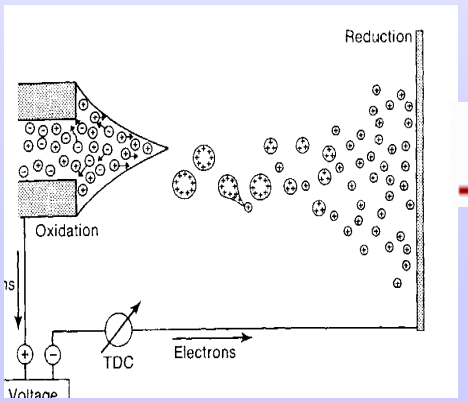

Figure 2 Positive-ion ES mass spectrum of a PEG mixture

# Molecular Weight of Proteins

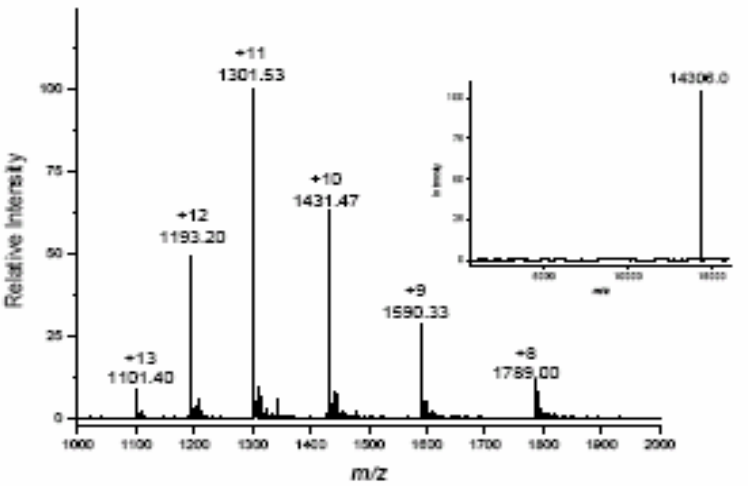


MALDI



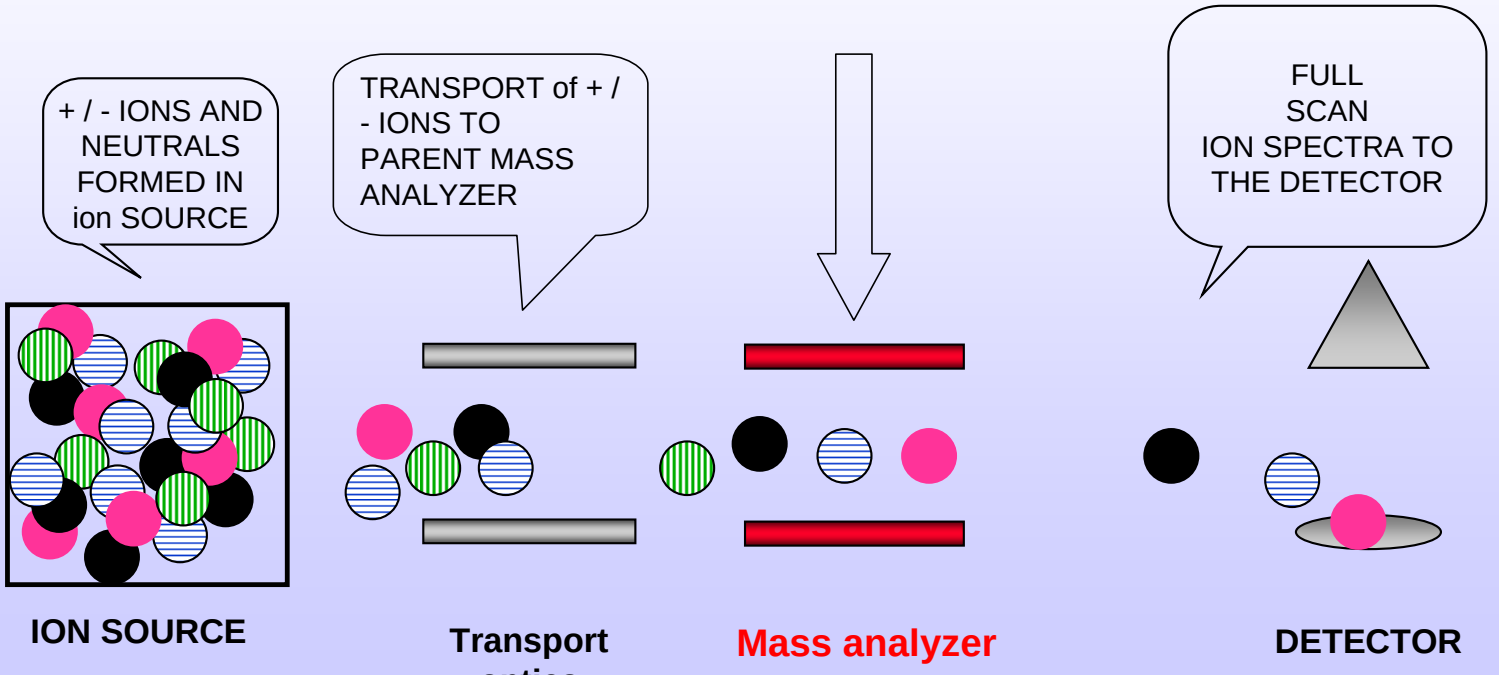



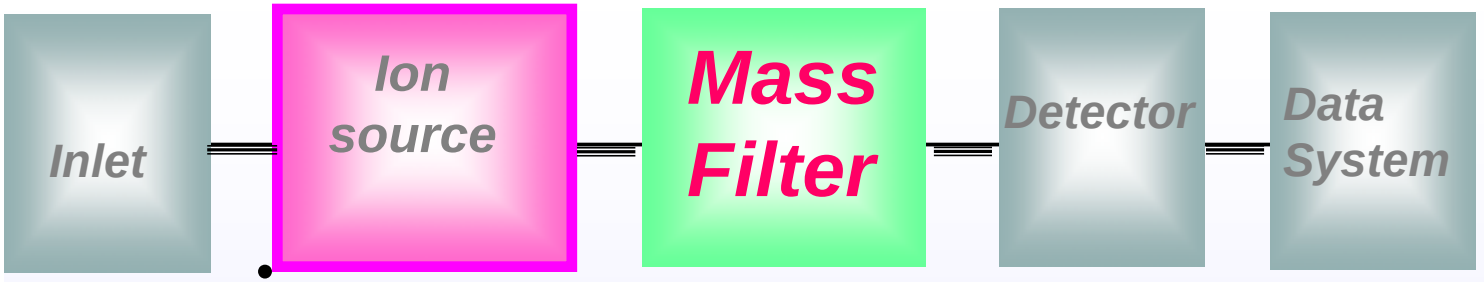

ESI



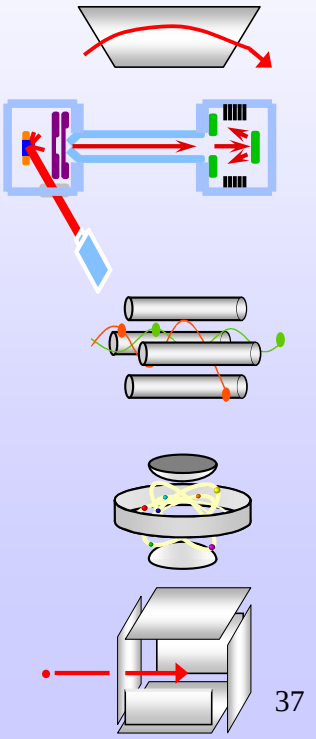
# Mass Analysis

Ions are separated according to their mss-to-charge (*m/z*)

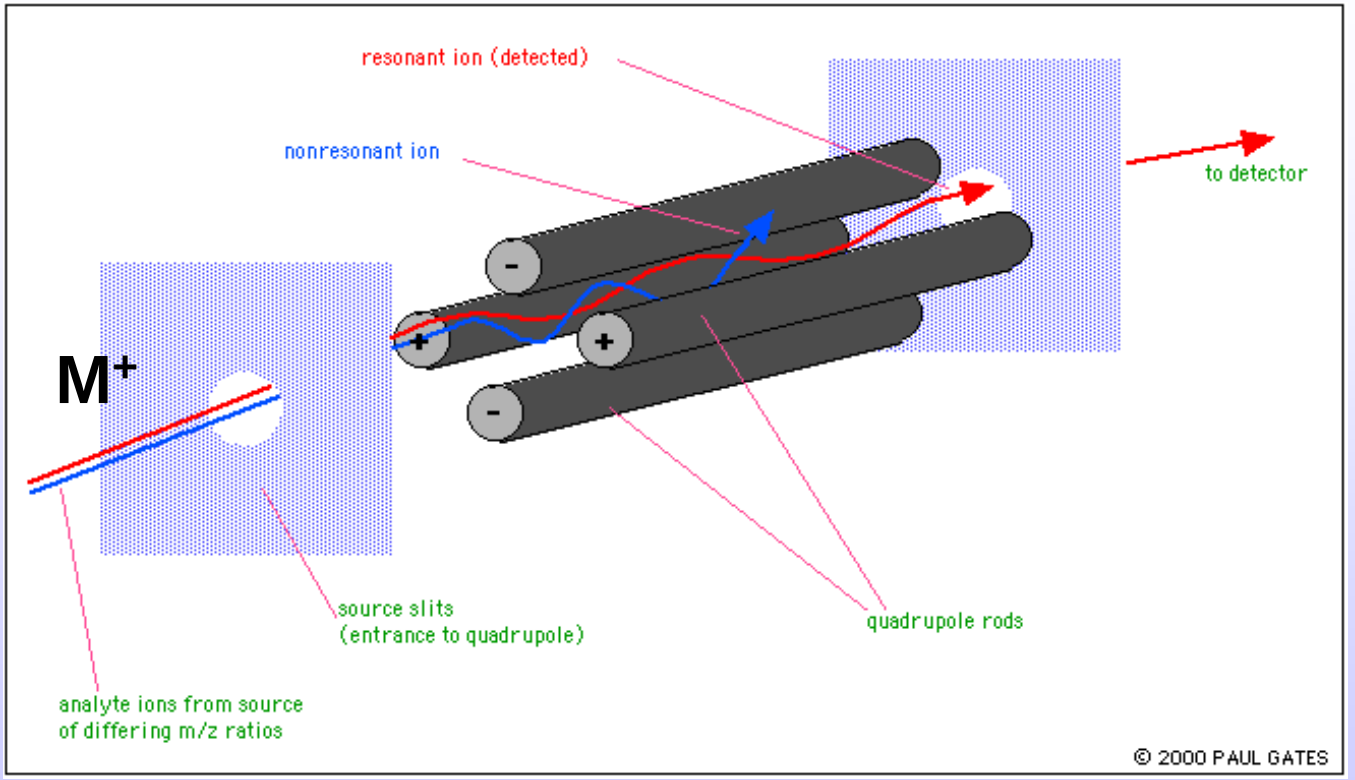





- Sector Instruments
- Time-of-flight Analyzer
- Quadrupole Mass Filter
- Ion-Trap Instrument
- FT-ICR



# Quadrupole Mass Filter (m/z -4000)




$a_u = a_x = -a_y = 4zU / m\omega^2 r_o^2$

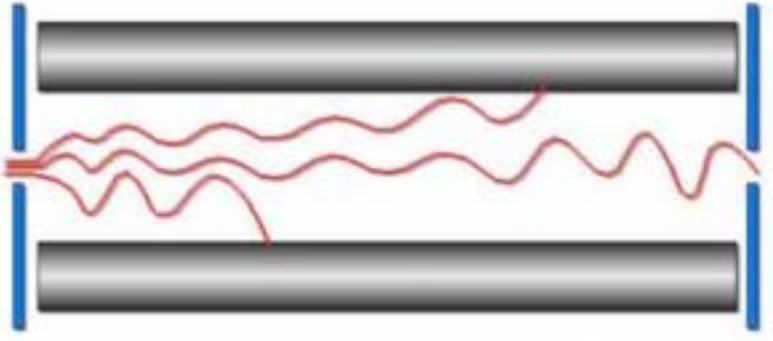
$q_u = q_x = -q_y = 2zV / m\omega^2 r_o^2$

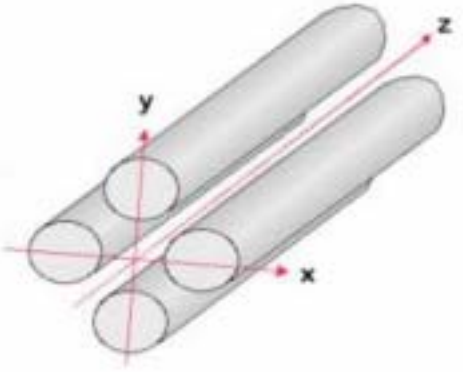
$a/q = 2U/V$, others fixed

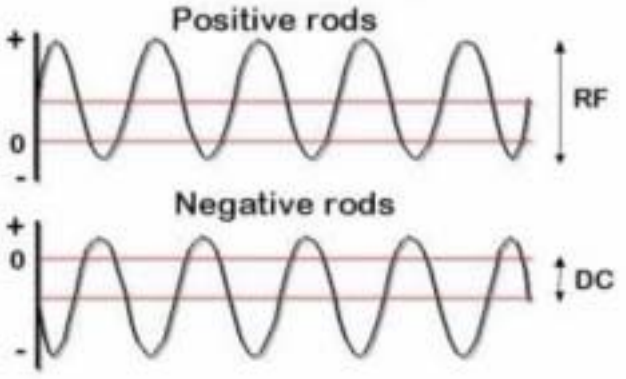


Length, diameter, kinetic energy -→m/e range, resolution

# Elements of a quadrupole analyzer



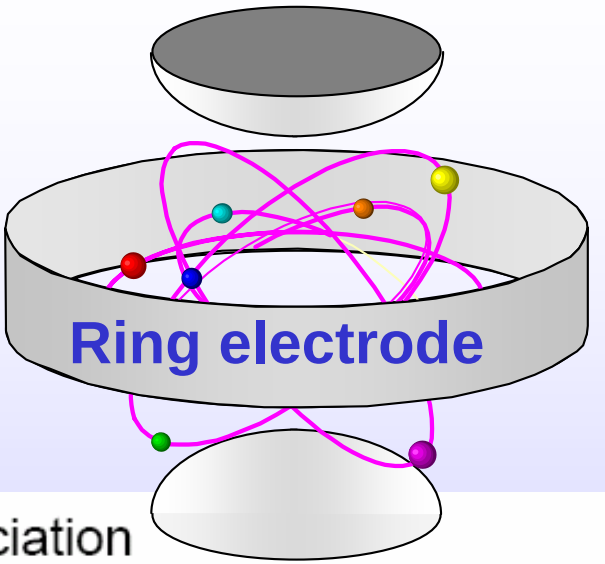



S Barnes-UAB 1/27/04

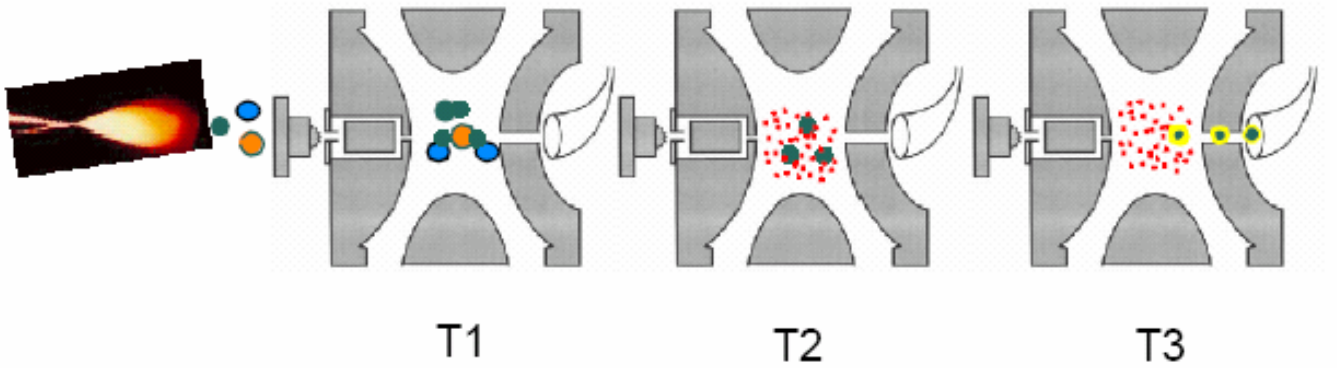
# Ion-Trap Instrument

- Principle very similar to quadrupole
- Ions stored by RF & DC fields
- Scanning field can eject ions of specific m/z

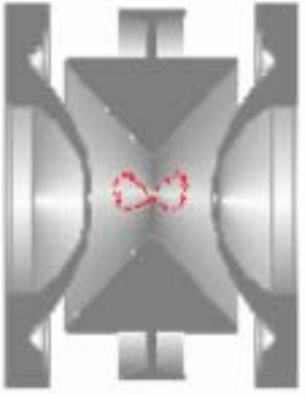



$MS^n$

Collisions with gas msec dissociation

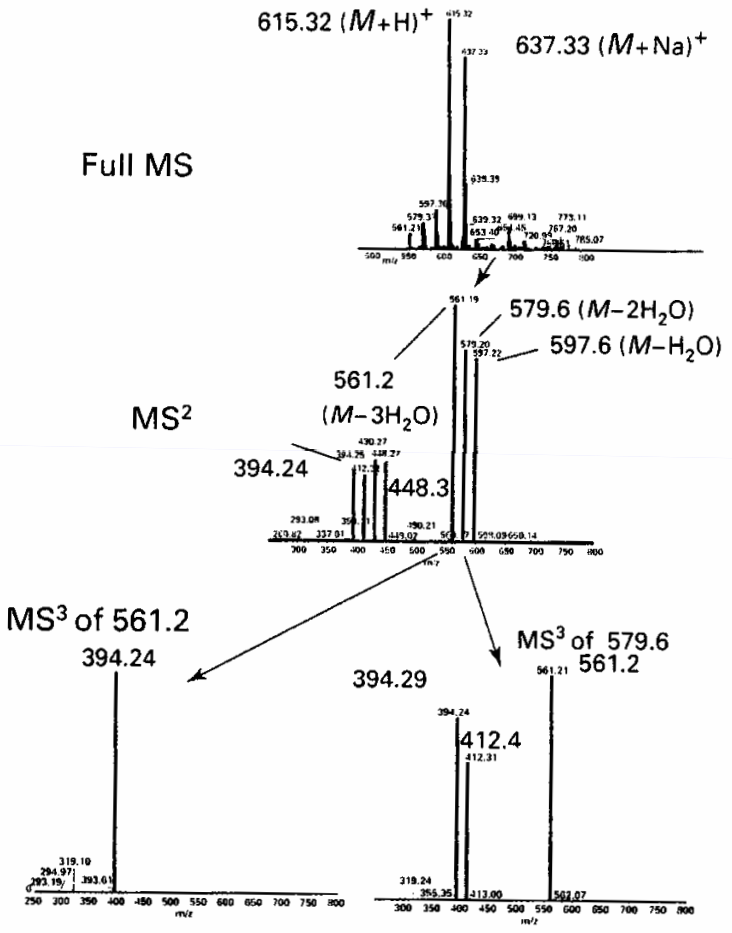


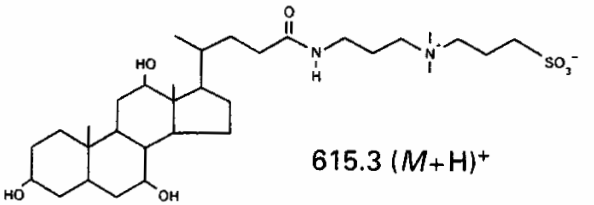
http://www.chem.wm.edu/dept/faculty/jcpout/faculty.html

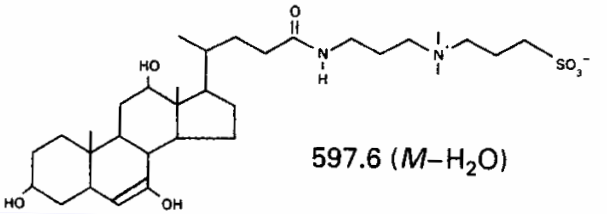


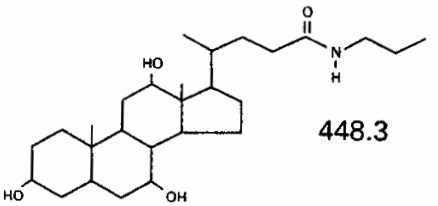
Ion path in a trap

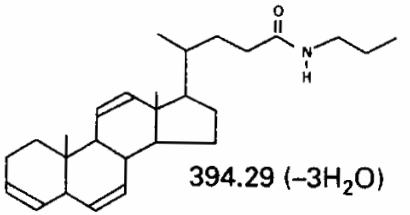
# Multiple MS/MS ($MS^n$) for Structural determination



# Time-of-flight Mass Spectrometry

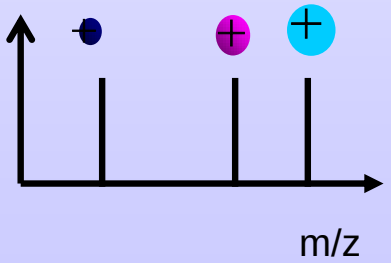
- Ions are generated in the source zone of the instrument.
- A potential (V) is applied across the source to extract and accelerate the ions from the source into the field-free 'drift' zone of the instrument.
- Ions travel with velocity **v = d/t** ; d:tube distance, t:time
- All ions produced will leave the source at the same time with the same kinetic energy (**KE = ½ *mv*² = *zV***), due to their having been accelerated through the same potential difference (ideally).
- The time-of-flight of the ions produced will only be dependent on the mass and the charge of the produced ion.

$$m/z = [2\,t^2\,V] / d^2$$

- The larger the ion, the slower its velocity and thus the longer it takes to traverse the field-free drift zone.


Mass spectrum

$$t = \left( \frac{m}{2zeEs} \right)^{\frac{1}{2}} D$$




t= 100 µ second for mass 50 ( small moleculae)

1000 µ second for mass 5,000 (peptide, polymer…)

3000 µ second for mass 50,000 (protein, DNA….)

# Calibration of Mass Spectrum

***Flight time***

$$t = \left( \frac{m}{2zeEs} \right)^{\frac{1}{2}} D$$

D: drift length

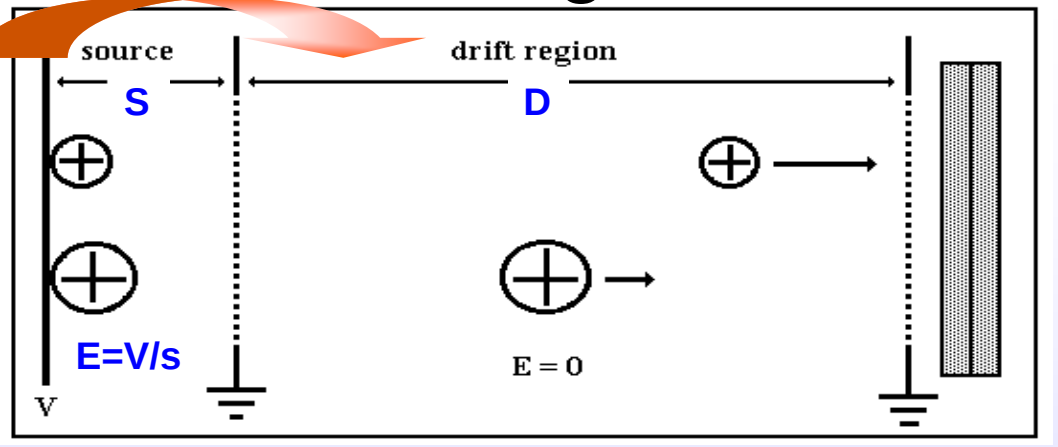
$$t = a(m)^{\frac{1}{2}} + b$$

The mass is independent of any instrumental parameters

- Internal calibration
- External calibration

# *m/z* is a function of flight time

Vacuum pump
$10^{-5}$ ~$10^{-8}$ torr



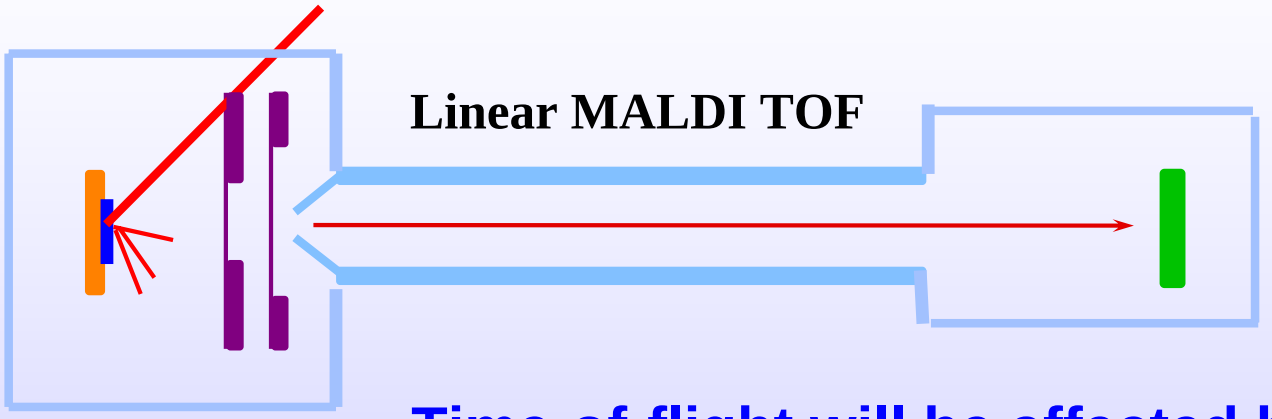

***In a electric field*** $E = \frac{V}{s}$ *where V = accelerating voltage*

***In drift region*** $\frac{1}{2}mv^2 = qV = zeEs \longrightarrow v = \left(\frac{2zeEs}{m}\right)^{\frac{1}{2}}$

$\longrightarrow$ ***Flight time*** $t = \left(\frac{m}{2zeEs}\right)^{\frac{1}{2}} D$

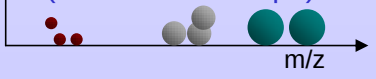
***Relation between m/z and t*** $\left(\frac{m}{z}\right) = \left(\frac{2eEs}{D^2}\right)t^2$

# Factors of Poor Resolution in TOF MS




**Linear MALDI TOF**

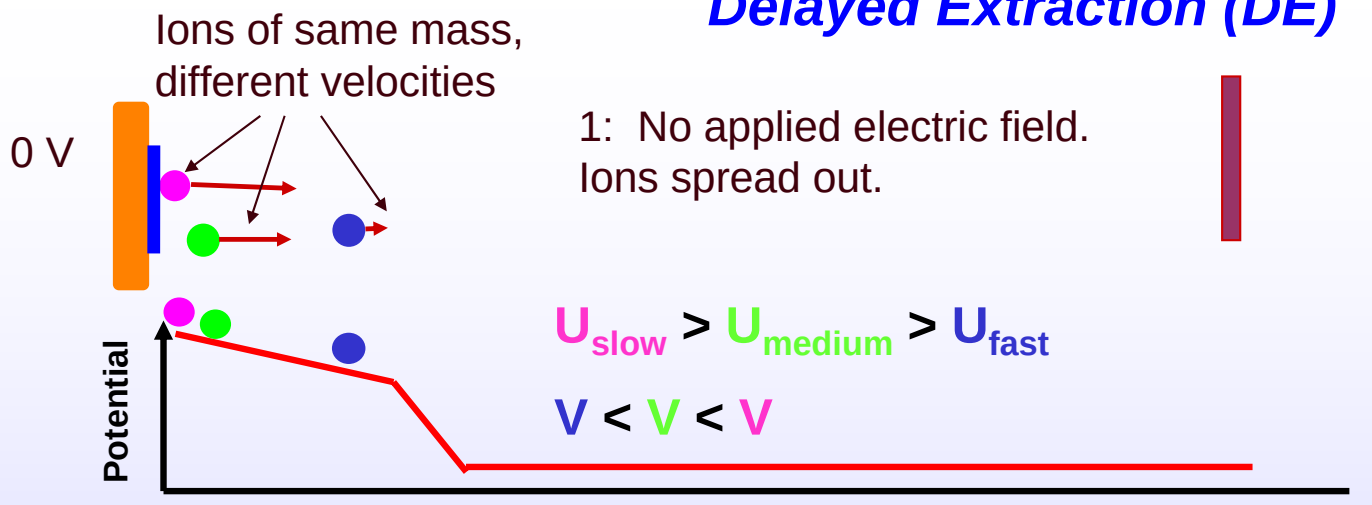
**Large kinetic energy spread, broad peaks**
poor resolution
(unresolved isotope)

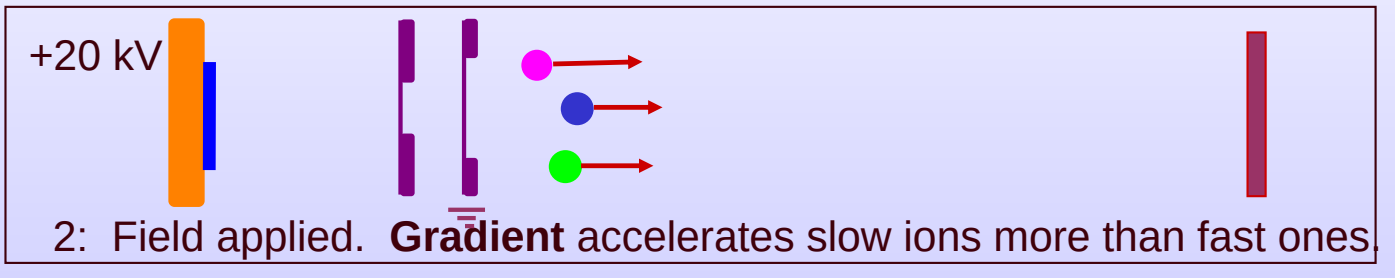


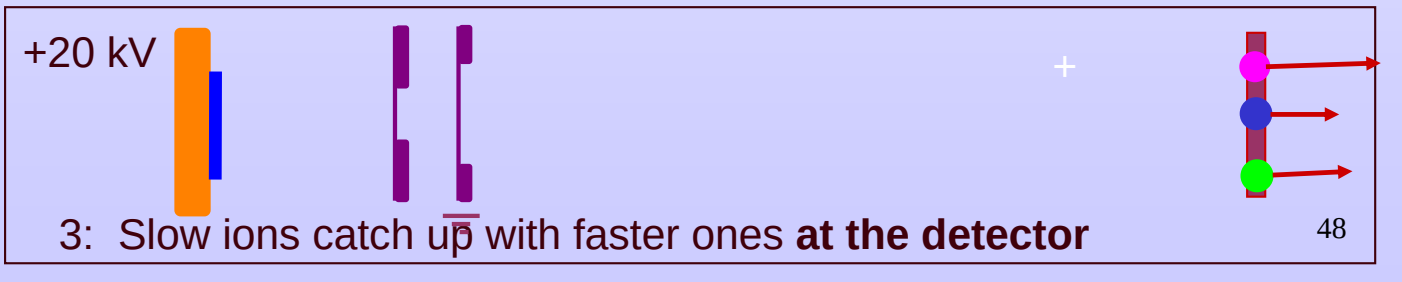

## Time-of-flight will be affected by:

- **Temporal Distribution**
  time of ion formation
- **Spatial Distribution**
  initial location in the extraction field
- **Kinetic Energy Distribution**
  different initial kinetic energy

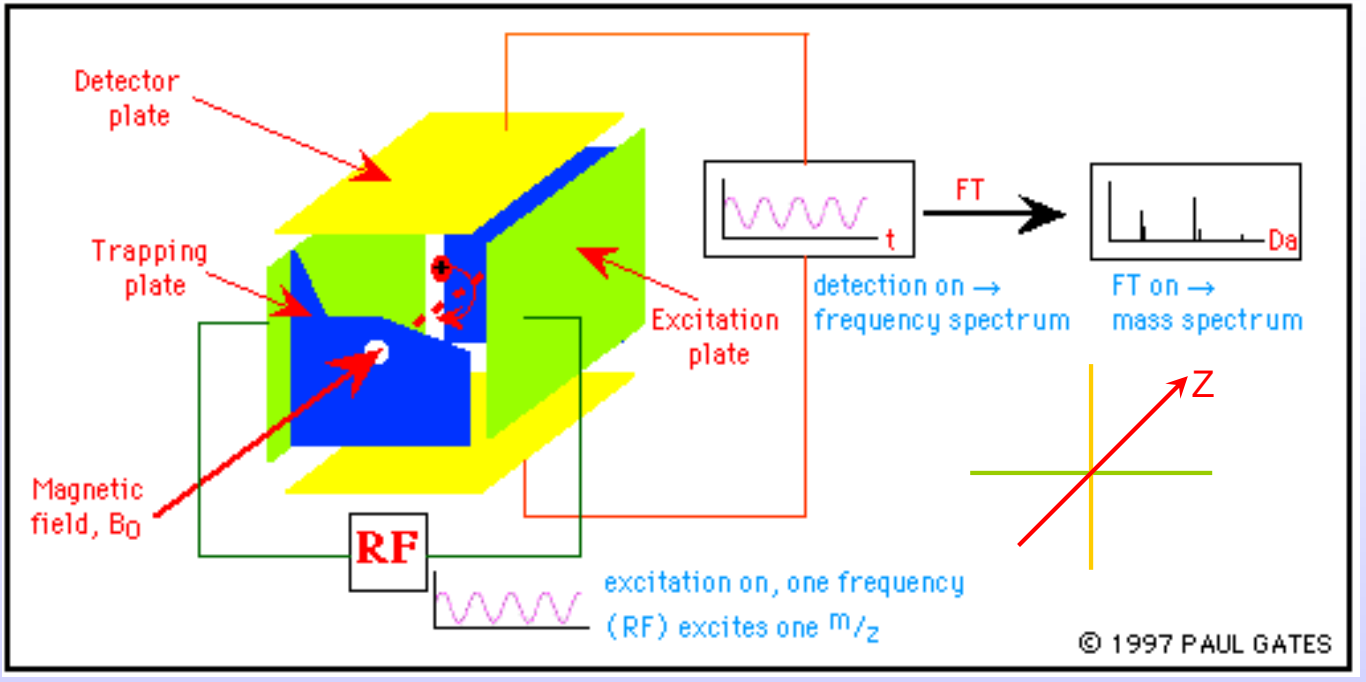
# *Delayed Extraction (DE)*



# Fourier Transform Ion Cyclotron Resonance Analyzer, FTICR

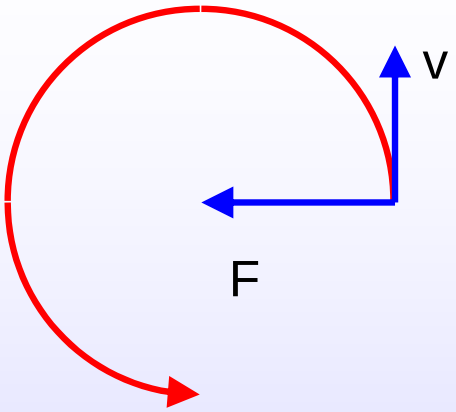



$$\mathbf{F} = z_i \mathbf{v} \times \mathbf{B}_0$$

$$\omega_c = \frac{z_i \mathbf{B}_0}{2\pi m_i}$$

$$\frac{m_i}{z_i} = \frac{\mathbf{B}_0}{2\pi\omega_c}$$

Comisarow, M. B.; Marshall, A. G. Fourier transform ion cyclotron resonance spectroscopy *Chem. Phys. Lett.* **1974**, *25*, 282-283.



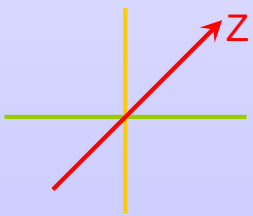

F=z v x B

F: Magnetic force on the ion

z: Charge of the ion

v: Velocity of the ion


B: Magnetic Field



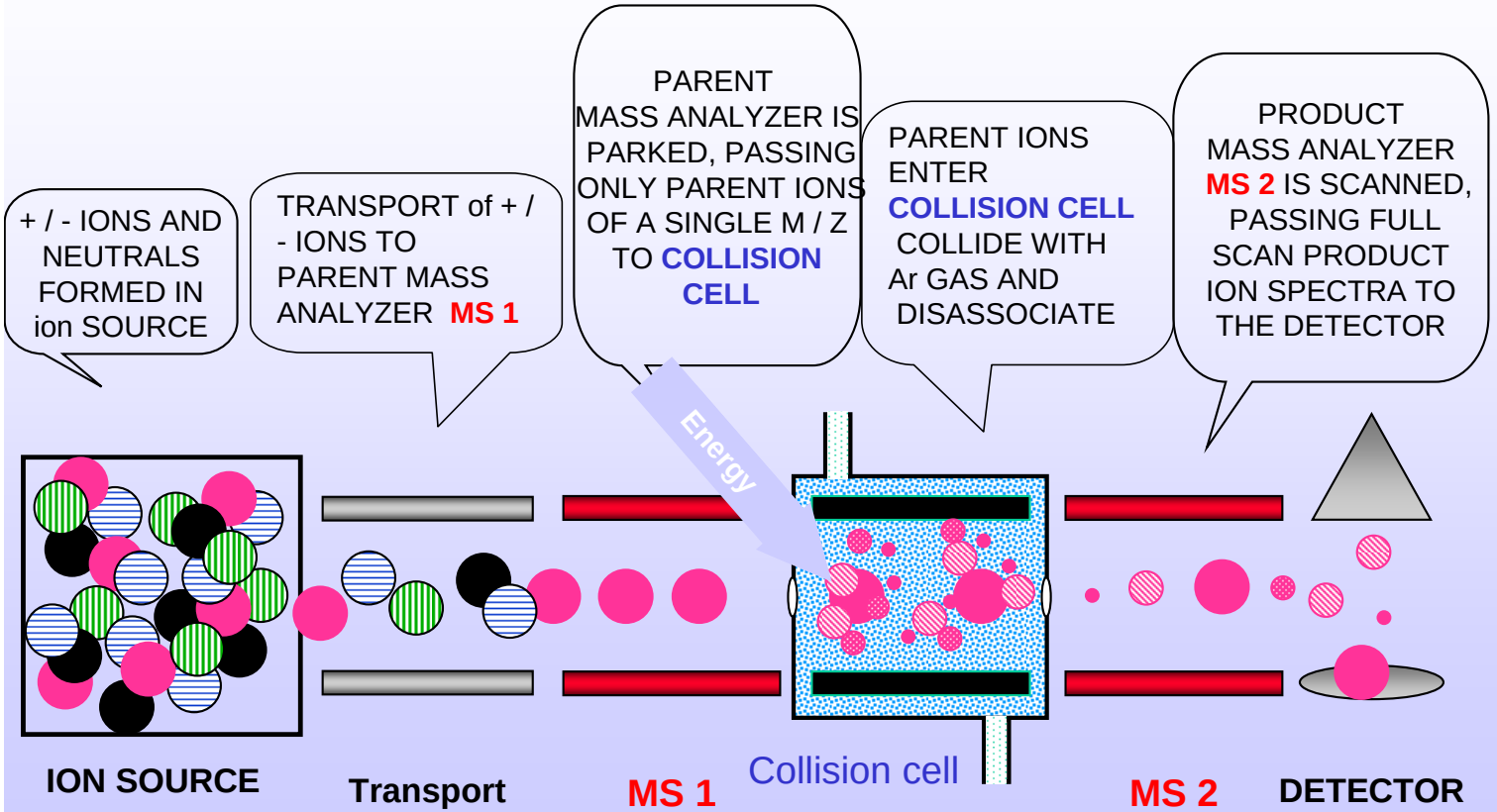

A charged particle (ion) will rotate around the magnetic field line of a homogeneous
magnetic field in a circular motion.

Large biological system | Exact mass measurement | LC-MS

# Comparison of Different Types of Mass Spectrometers

| Characteristic | Magnetic Sector | Quadrupole | QIT | TOF | FT-ICR |
|---|---|---|---|---|---|
| **Mass Range (Da)** | 15,000 | 4,000 | 100,000 | unlimited | > $10^6$ |
| **Resolution** | 200,000 | Unit | 30,000 | 15,000 | > $10^6$ |
| **Dynamic Range** | ++++ | +++ | +++ | +++ | ++ |
| **MS/MS** | ++++ | +++ | +++ | +++ | ++ |
| **LC or CE** | + | ++++ | +++ | ++ | ++ |
| **Cost** | $$$$ | + | + | ++ | ++++ |

# Tandem Mass Spectrometry

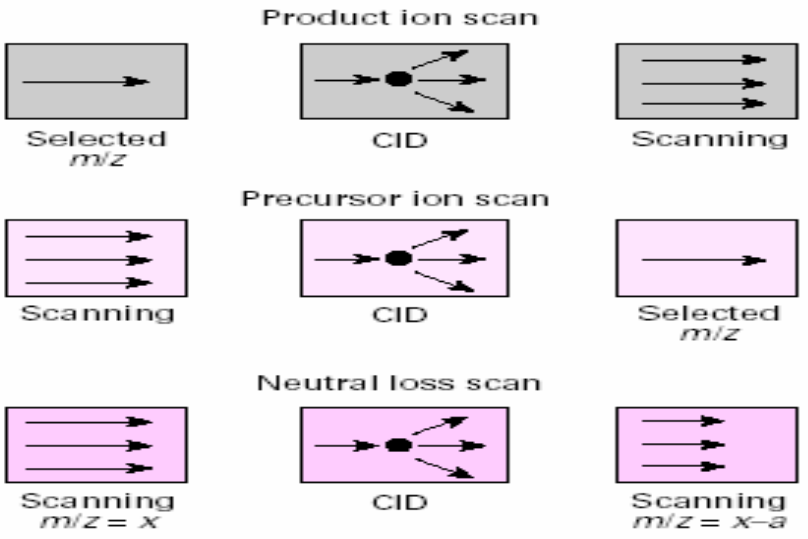


# Methods of MS-MS

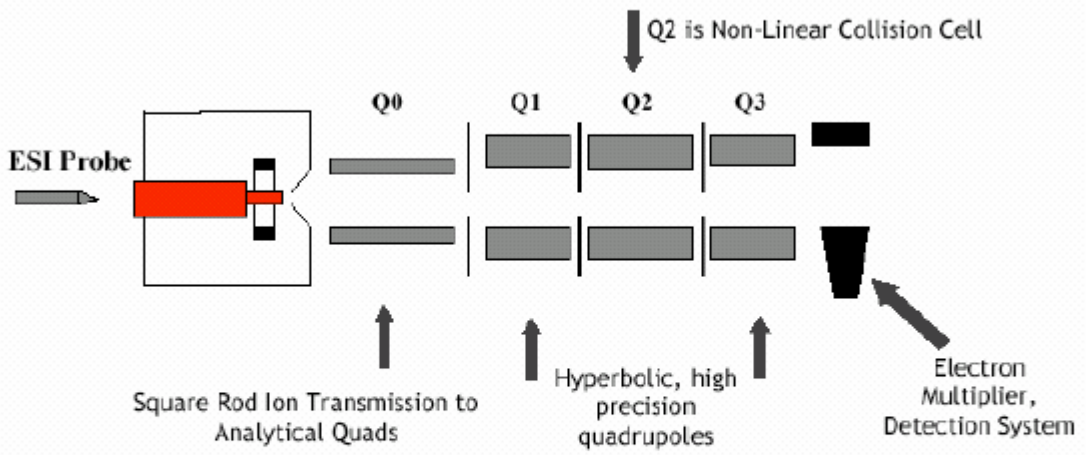
**In space**

MS 1 → Collision cell → MS 2

Precursor ion selection — Dissociation — Product ion scan

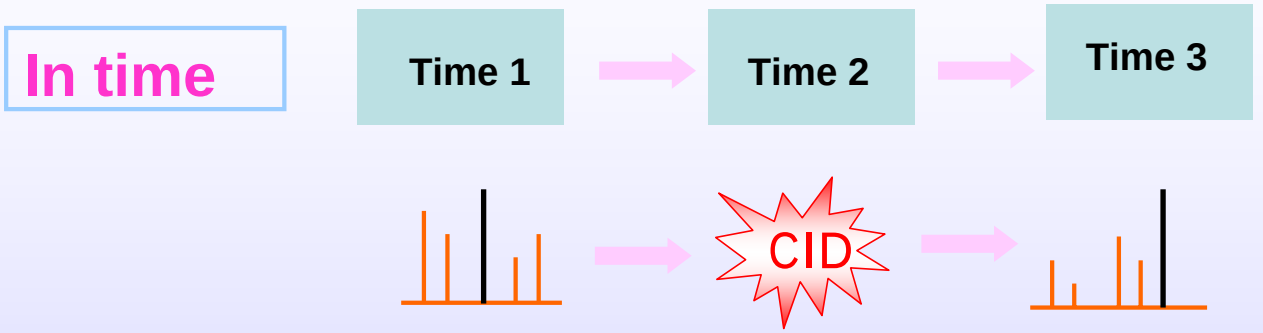
# Triple Quadrupole (QQQ)

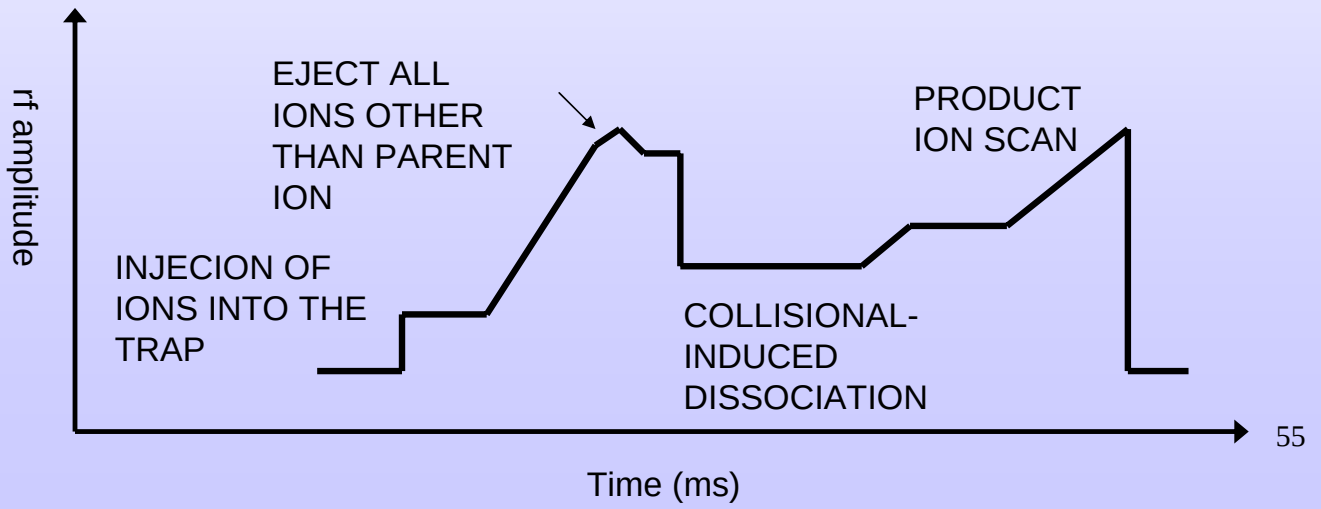



**Quad Functions during MS/MS**

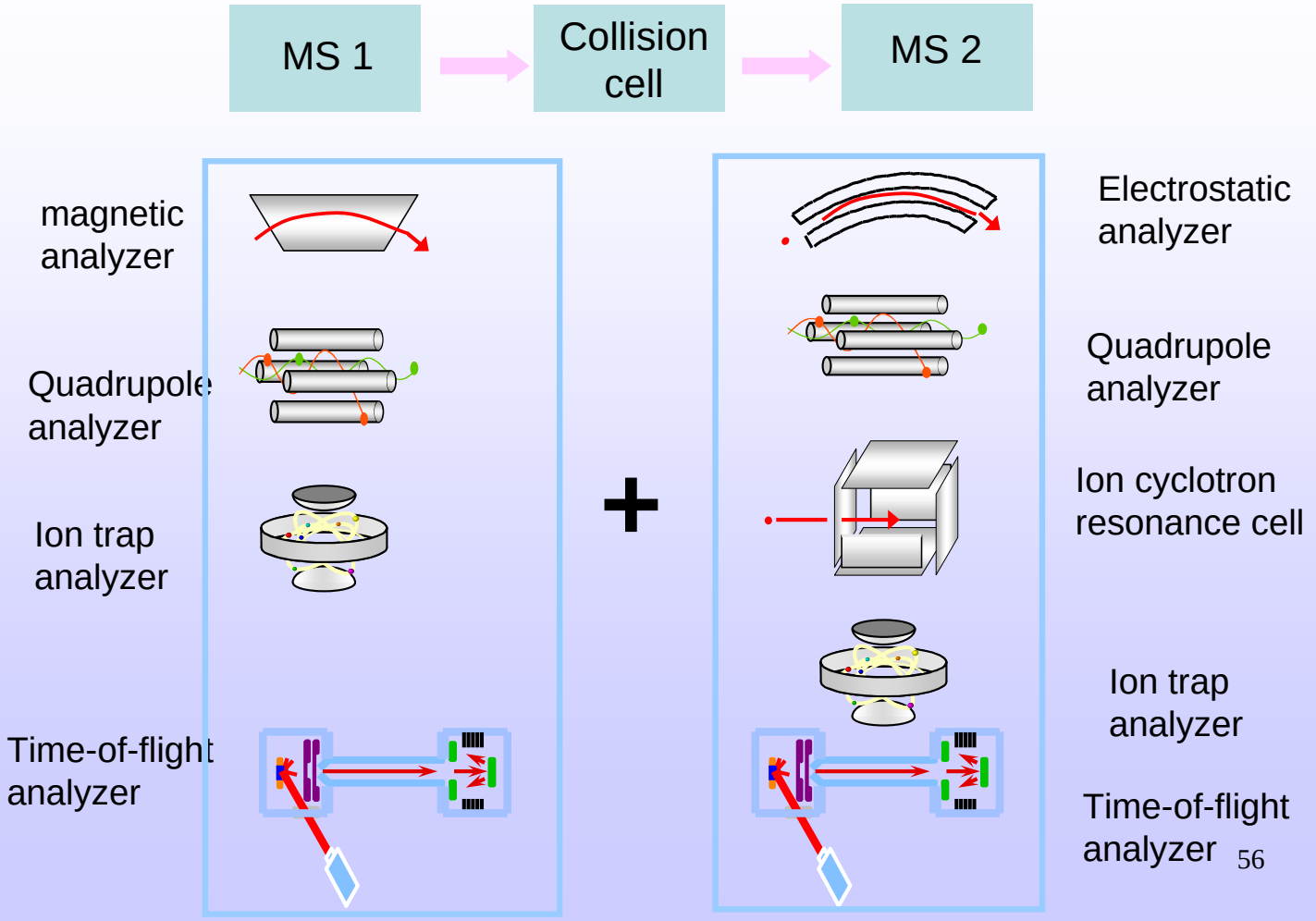
- Q1 scans a preset m/z range or selects ion of interest
- Q2 (collision cell) transmits ion and introduces collision gas into flight path
- Q3 analyses fragment ions generated in Q2

# Methods of MS-MS



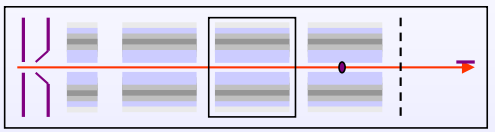



# New Types Mass spectrometer



# Orthogonal-Quadrupole Time-of-flight Mass Spectrometer, Q-TOF

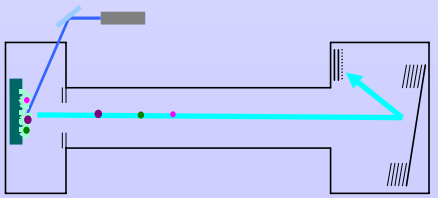
Triple quadruple MS



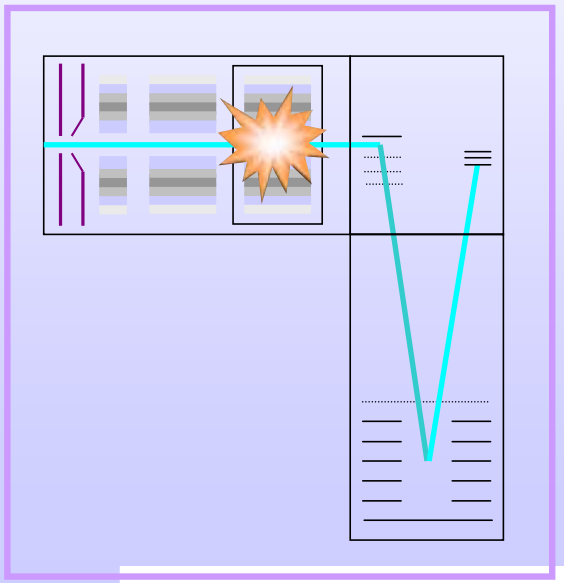
- Good precursor ion selection
- Efficient fragmentation

+

Time-of-flight MS

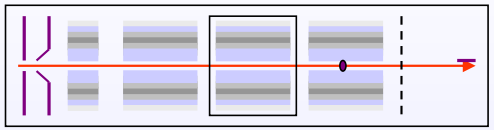


- Good sensitivity
- Good sensitivity
- Good resolution
- Speed



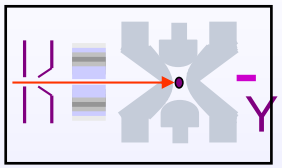
# 2D Linear Trap Instrument

Triple quadruple MS



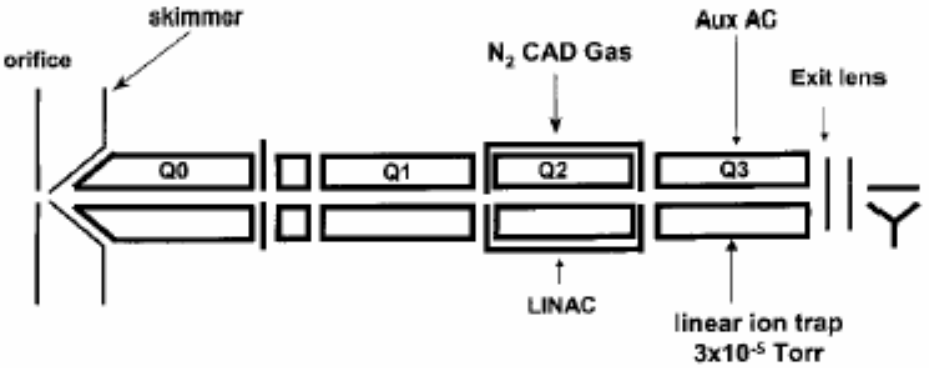
+

Quadrupole ion trap MS




- 4 scan modes
- Quantitative analysis

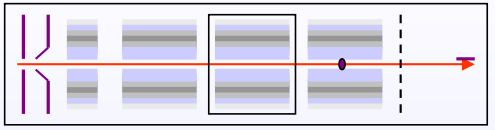
- Good sensitivity
- Sensitive product ion scans




Blanc *et al* proteomics, **3,** 859 (2003)

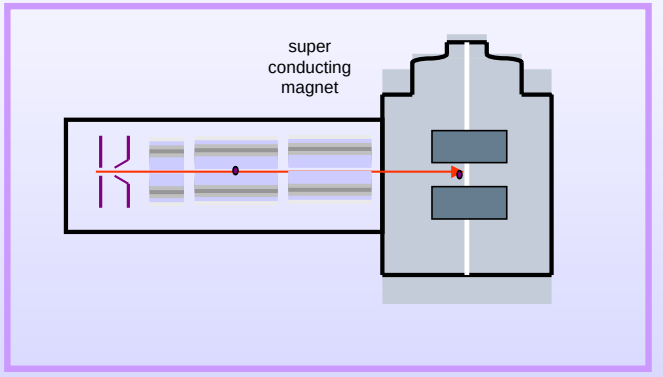
# FTICR Trap Instrument

Linear ion trap MS

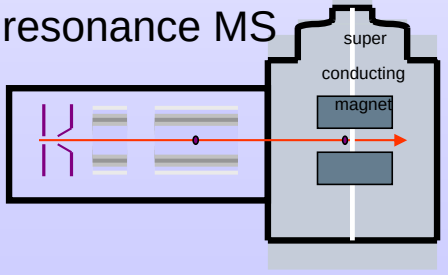


- Good sensitivity
- $MS^n$ capibility

+

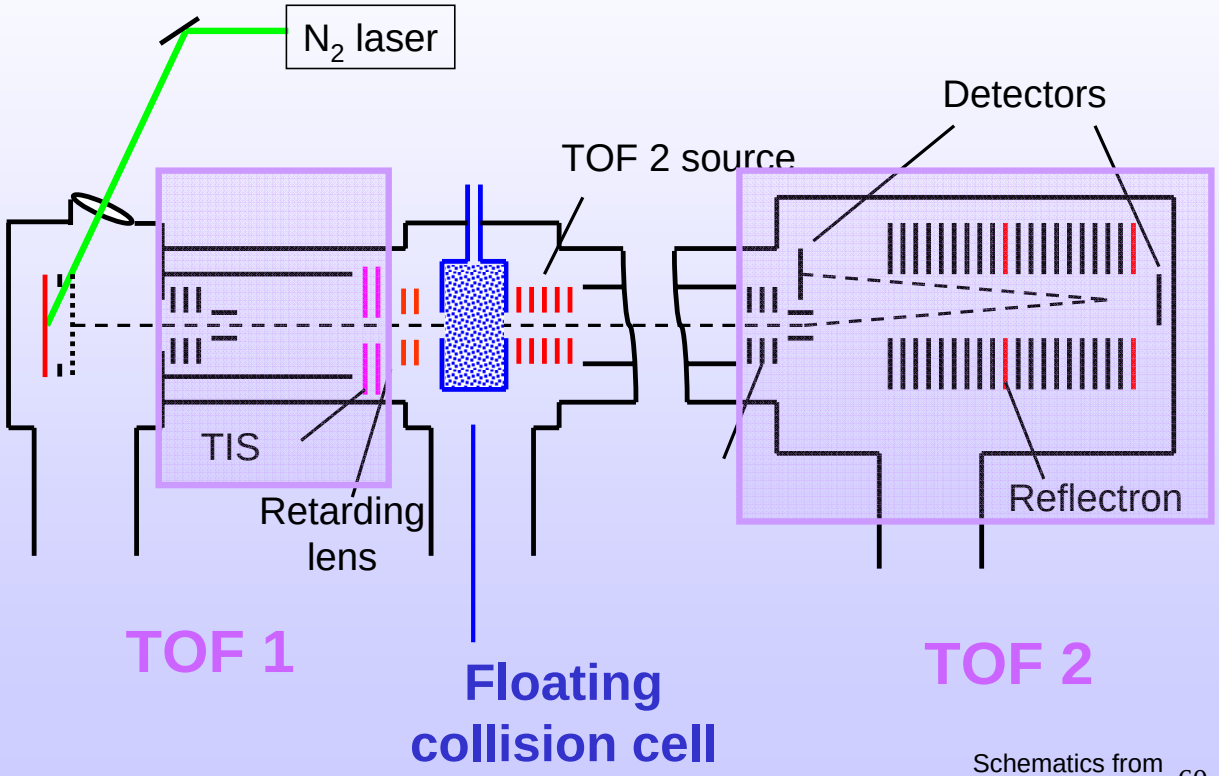



Fourier transform
ion cyclotron
resonance MS



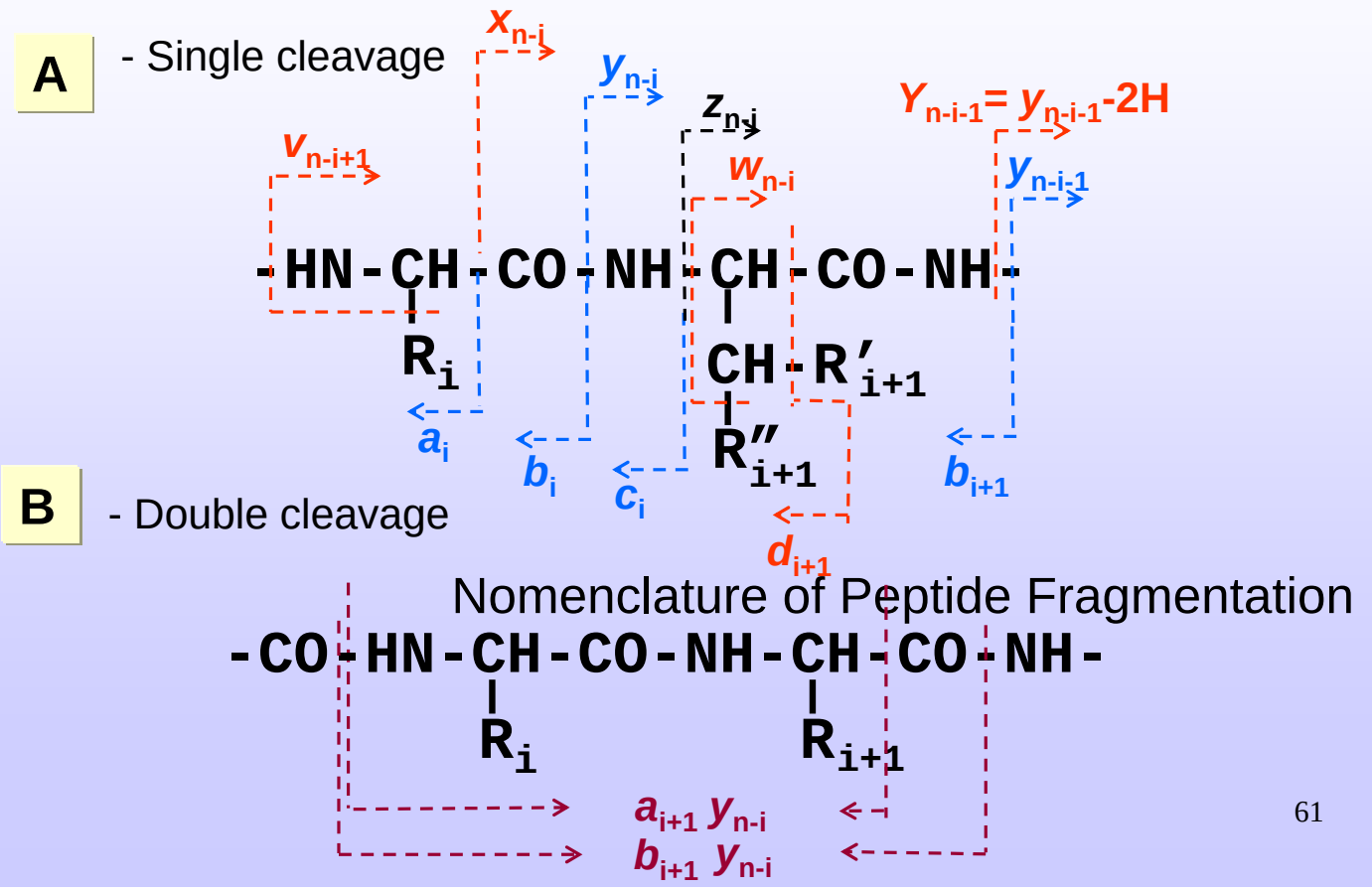

- Accurate Mass
- Ultra-high Resolution
- High dynamic range

# Time-of-flight / time-of-flight Mass Spectrometer, TOF-TOF




Schematics from Applied Biosystemss

# High Energy Peptide Fragment Pattern and Nomenclature



# Carbohydrates

- Carbohydrates are the most abundant organic molecules in nature
  - Photosynthesis energy stored in carbohydrates;
  - Carbohydrates are the metabolic precursors of all other biomolecules;
  - Important component of cell structures;
  - Important function in cell-cell recognition;
  - Carbohydrate chemistry:
    - Contains at least one asymmetric carbon center;
    - Favorable cyclic structures;
    - Able to form polymers

# Summary of Some Functions of Glycans

Providing structural components
- Cell walls
- Extracellular matrix

Modifying protein properties
- Solubility
- Stability

Intrinsic functions performed by glycans

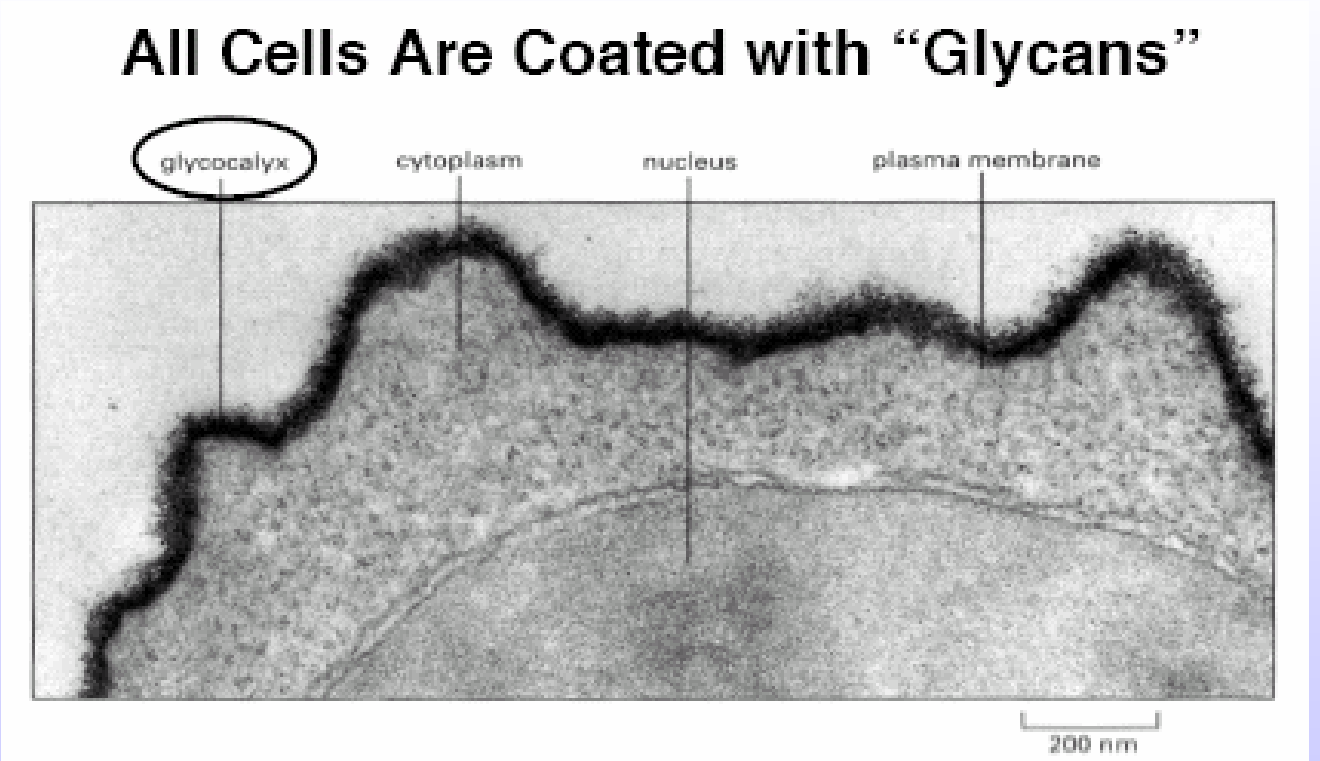
Directing trafficking of glycocconjugates
- Intracellular
- Extracellular

Mediating and modulating cell adhesion
- Cell–cell interactions
- Cell–matrix interactions

Mediating and modulating signalling
- Intracellular
- Extracellular

Extrinsic functions resulting from glycan–lectin interactions

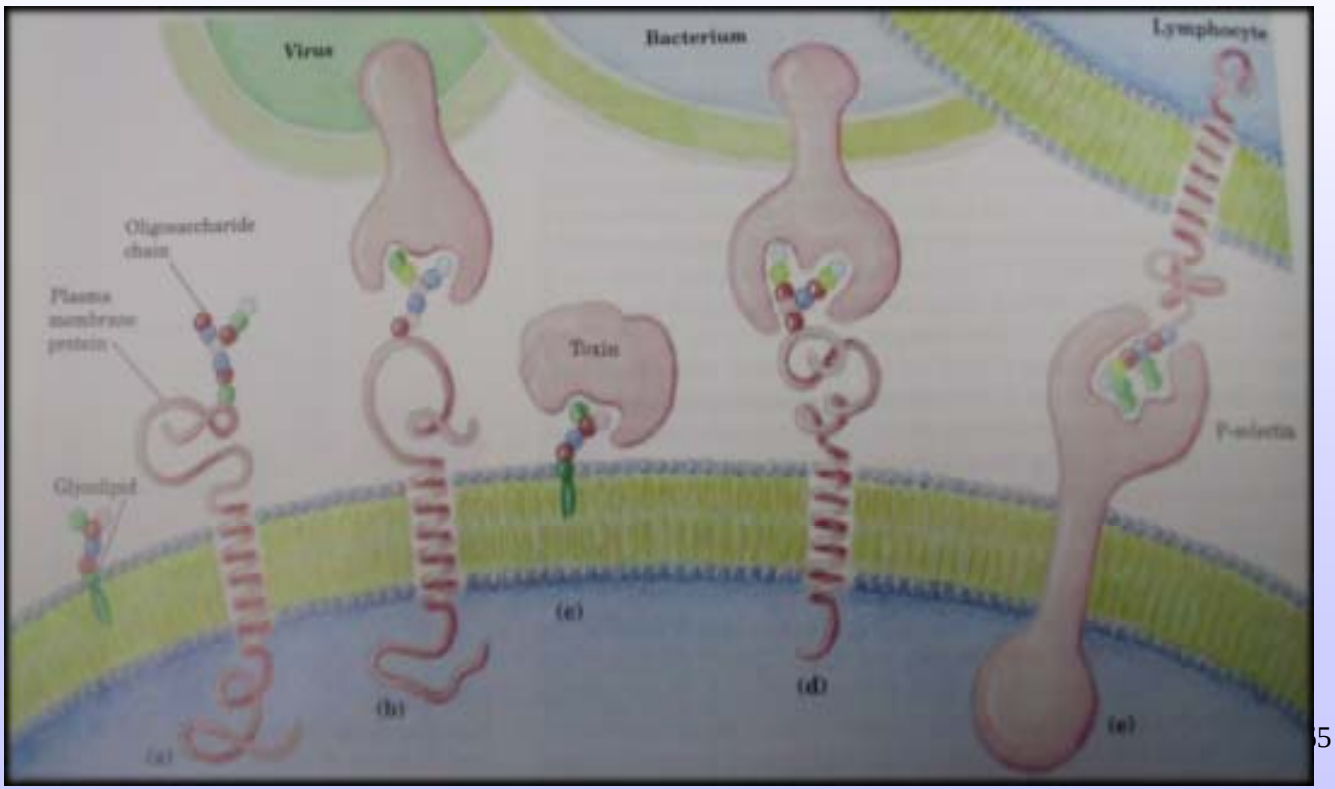
# The cell surface of bacterial is LPS-rich




Electron micrograph of a human lymphocyte (Ruthenium Red staining)

# Roles of Oligosaccharides in Recognition and Adhesion at the Cell Surface

Virus, Bacterium, Lymphocyte…….



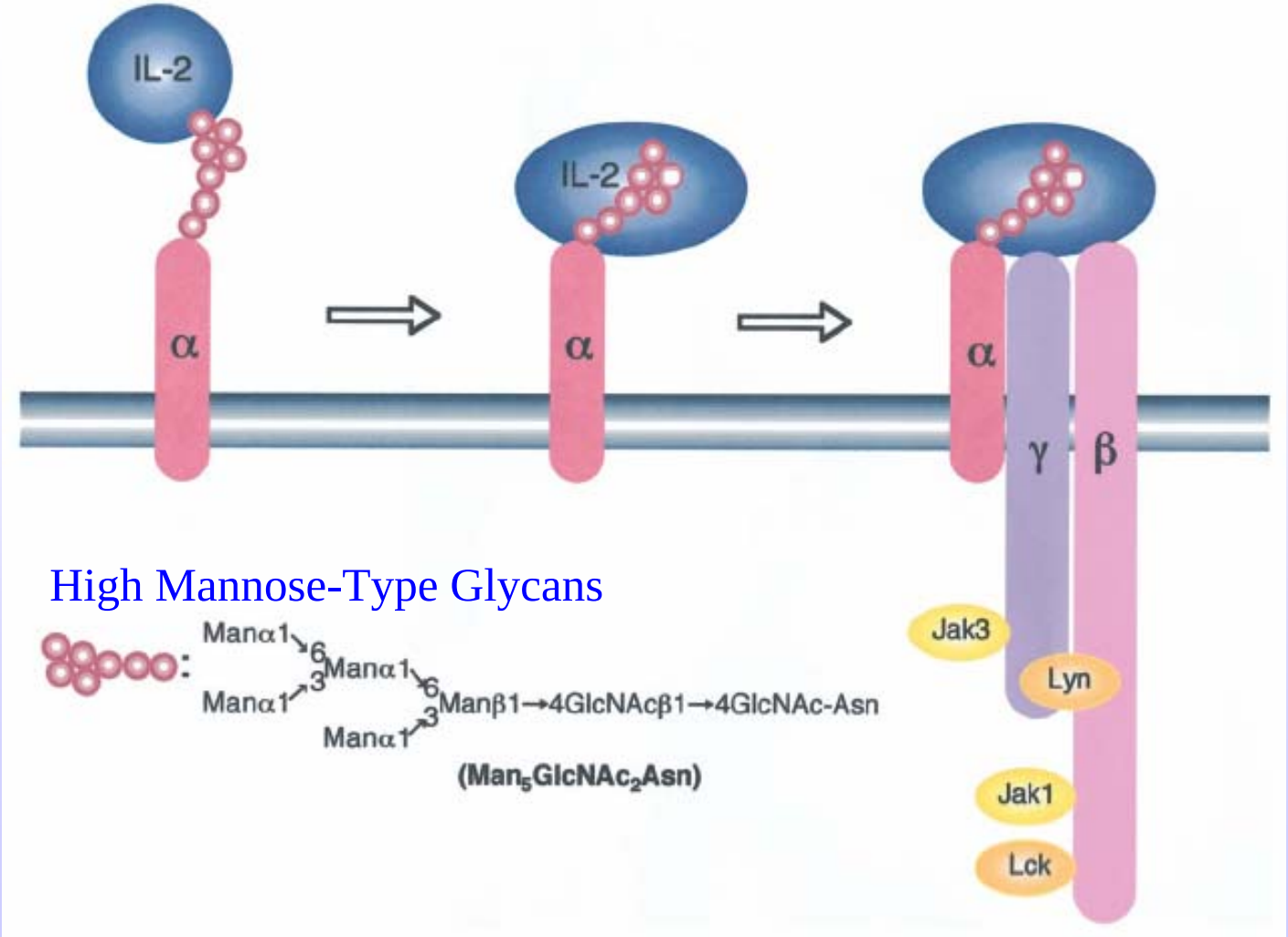
# Carbohydrates Have Diverse M.W.

Macromolecules

| Macromolecule | Building Block | Aproximate Mass | Possible Variations in a Trimer |
|---|---|---|---|
| Protein | Amino acids | 125 → $10^4$-$10^5$ | 6 |
| Nucleic Acid | Nucleotides | 330 → $10^3$-$10^9$ | 6 |
| Lipid | Fatty acids | 250 → $10^3$ | NA |
| Carbohydrate | Monosaccharides | 200 → $10^2$-$10^6$ | 1,056 to 27,648! |

# Monosaccharide Mass and Low Mass MS/MS (CID) Carbohydrate Marker Ions

| Monosaccharide / Disaccharide | Composition | Monoisotopic mass Intact | Residue | Low Mass CID |
|---|---|---|---|---|
| Hexose, Hex (Galactose, Mannose) | C6O6H12 | 180.0634 | 162.0528 | 163 |
| Deoxy-Hexose, dHex (Fucose) | C6O5H12 | 164.0685 | 146.0579 | 147 |
| N-acetylhexoseamine, HexNAc (N-acetylglucoseamine, N-acetylgalactoseamine) | C8O6NH15 | 221.0899 | 203.0794 | 204 |
| N-acetylneuraminic acid, Neu5Ac, NANA (sialic acid) | C11O9NH19 | 309.106 | 291.0954 | 292/274/256 |
| Hex-HexNAc | C14O11NH25 | 383.1428 | 365.1322 | 366 |
| Hex3HexNac2 (Common N-linked Core) | C46O35N2H76 | 910.3278 | 892.3172 | - |



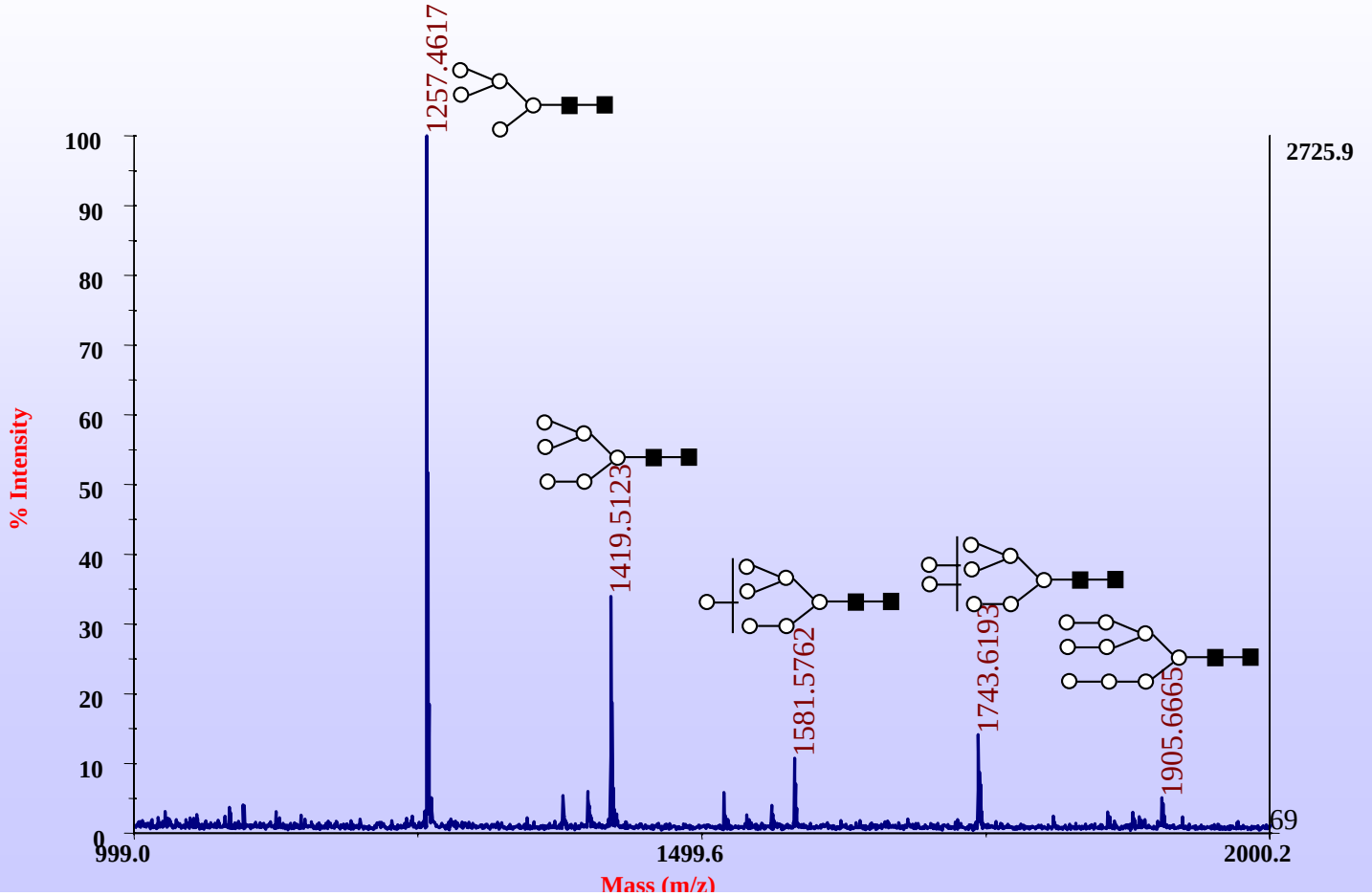

High Mannose-Type Glycans



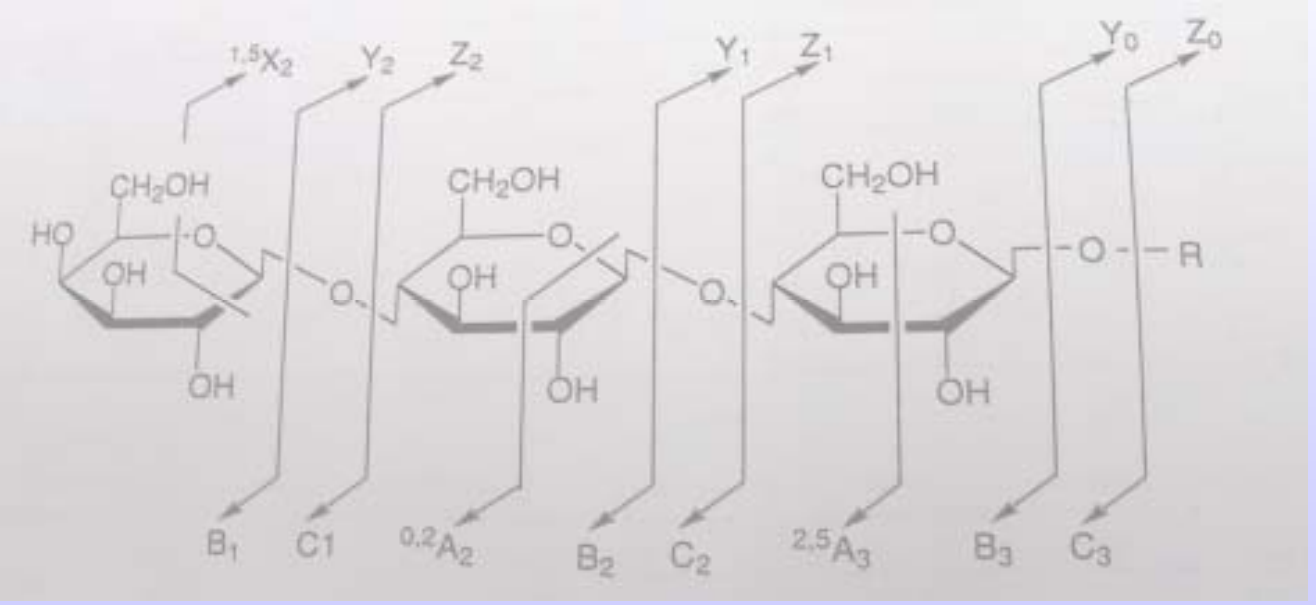



THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 276, No. 10, Issue of March 9, pp. 7351–7356, 2001

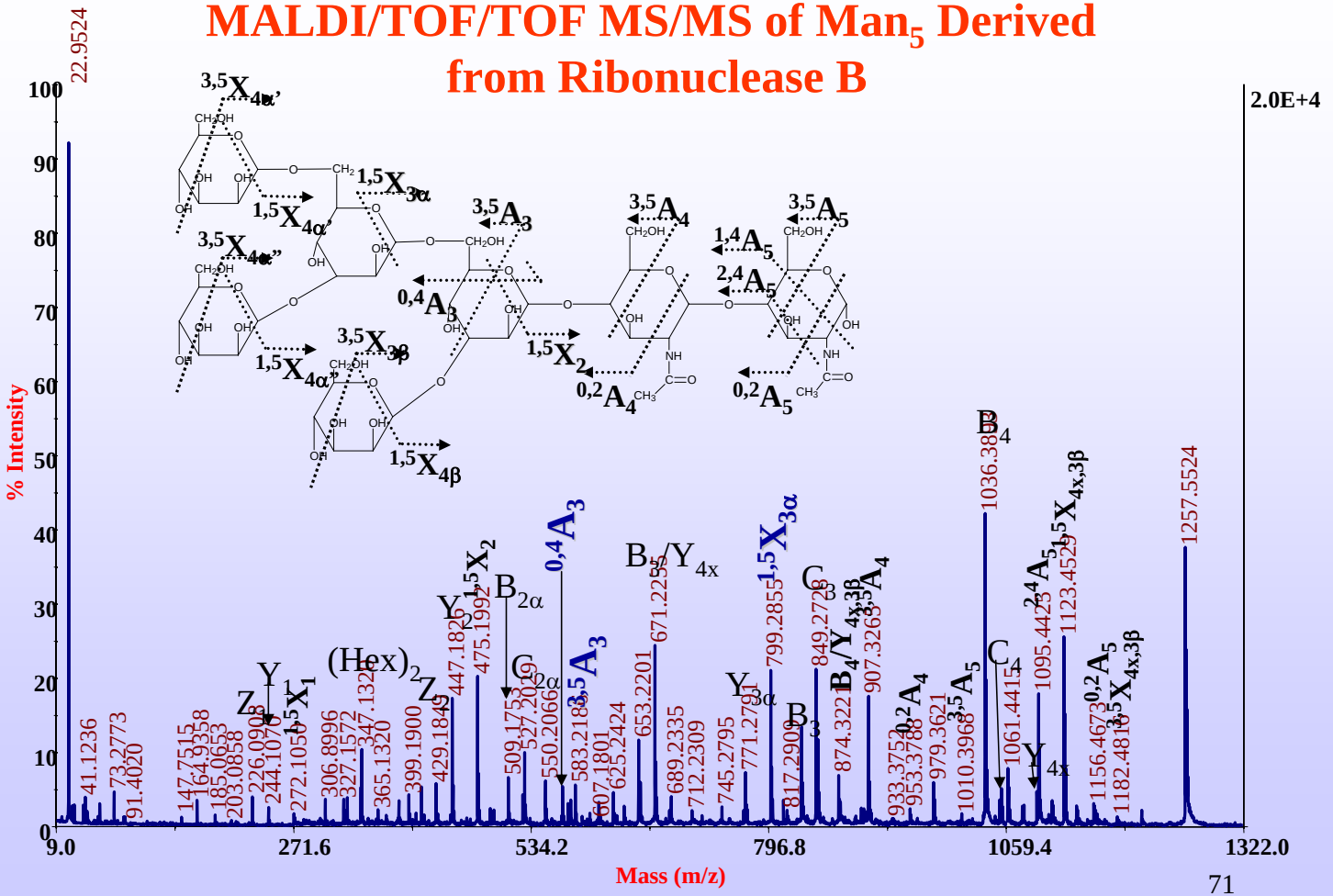
# MALDI/TOF/TOF MS Spectrum of N-Glycans Derived from Ribonuclease B



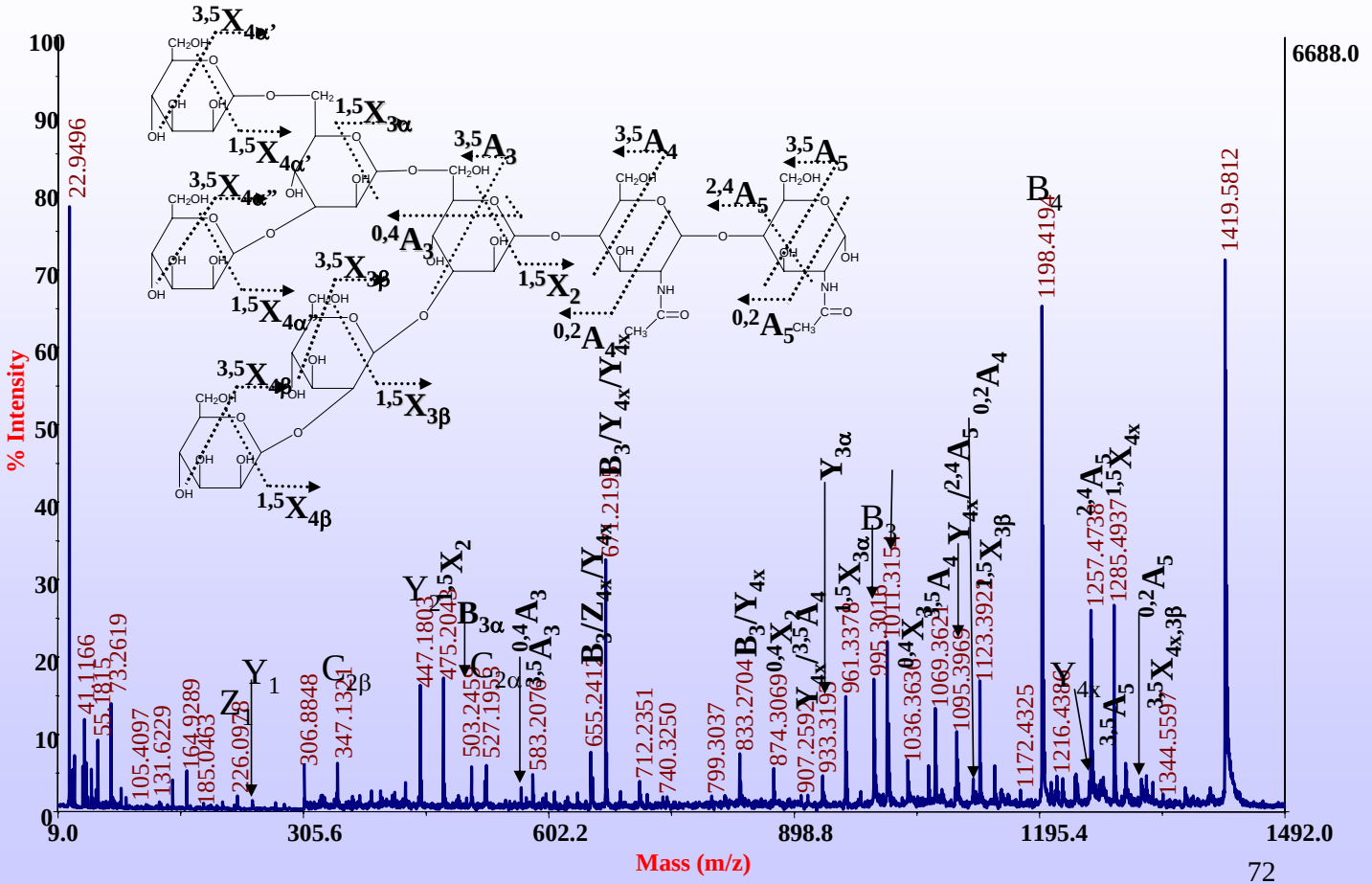
# Tandem Mass Spectrometry



# MALDI/TOF/TOF MS/MS of $Man_5$ Derived from Ribonuclease B

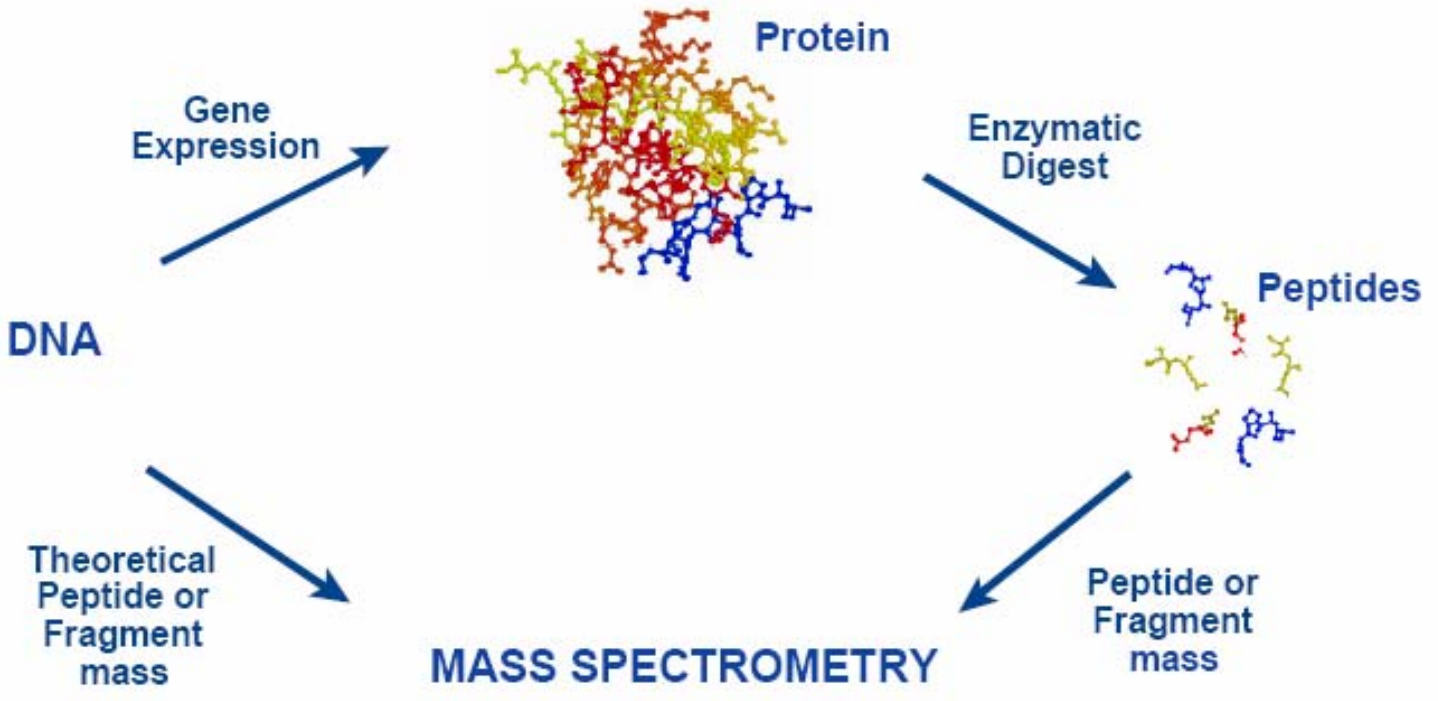


# MALDI/TOF/TOF MS/MS of $Man_6$ Derived from Ribonuclease B



# Identification and Quantification of Proteins

# Peptide and Protein Analysis



# Mass Spectrometry Methods


Molecular weight

Identification
Sequence information

- ✓ MALDI TOF MS
- ✓ ESI MS

- ✓Peptide mass fingerprinting (MALDI TOF MS)
- ✓Peptide Sequencing (Tandem MS)

# Peptide Mass Fingerprint

Left to right:

William J. Henzel; ***Protein Chemistry;*** Genentech

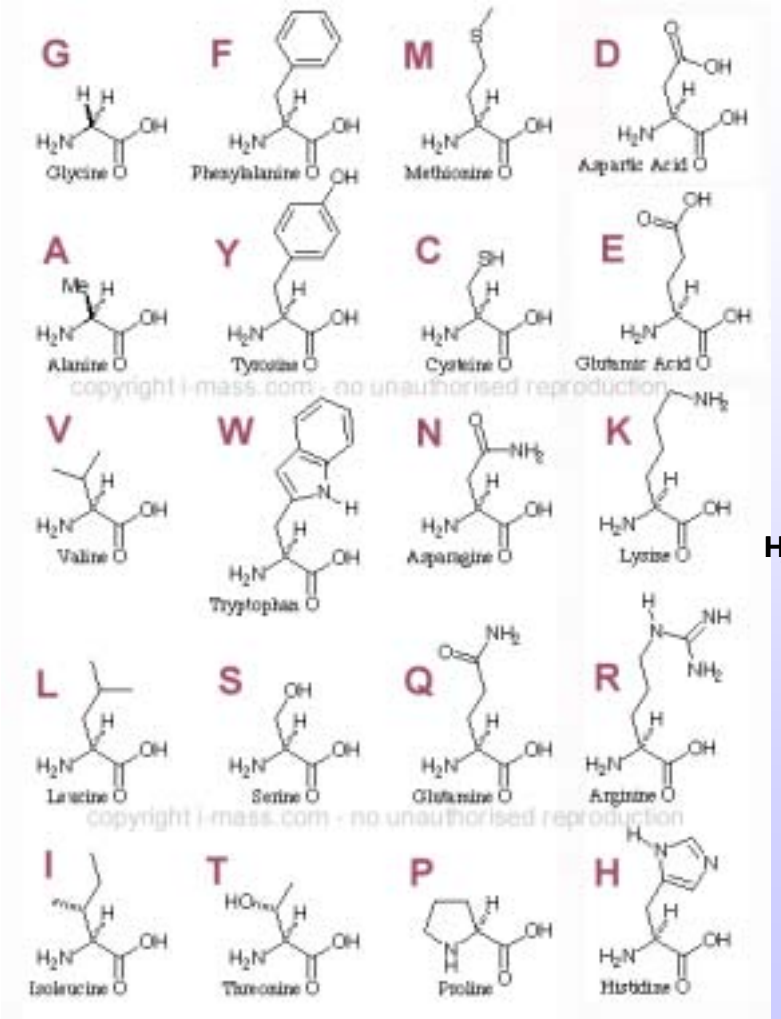
John T. Stults; ***Analytical Chemistry***; Genentech

Colin Watanabe; ***Software Engineer;*** Genentech

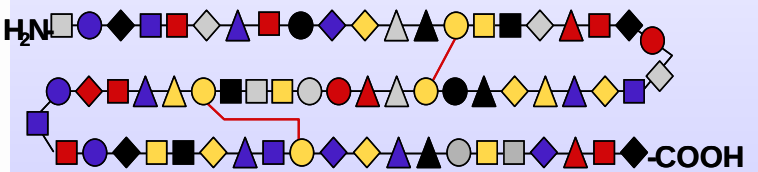
- This methodology was the first to allow protein identification without the need for time-consuming Edman sequencing or immunoaffinity probes..

- Landmark Study(1993)

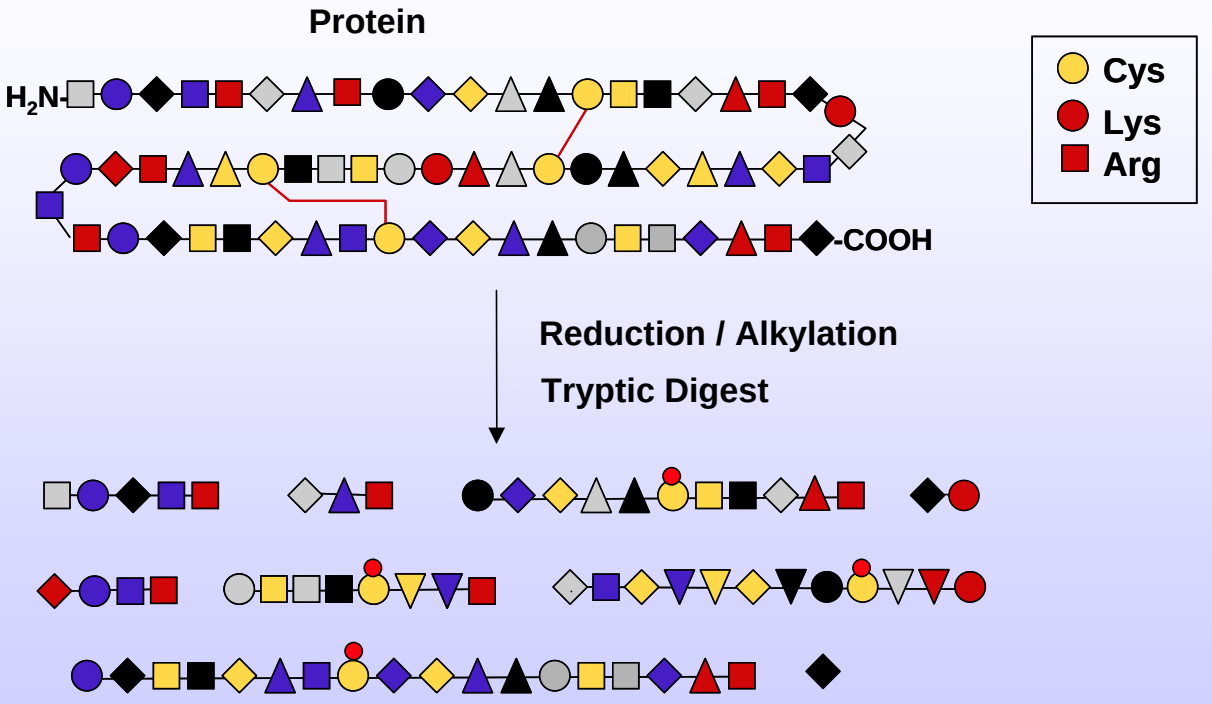
MS approaches alone could be used to analyze proteins from 2DE




# Peptide Mass Fingerprinting



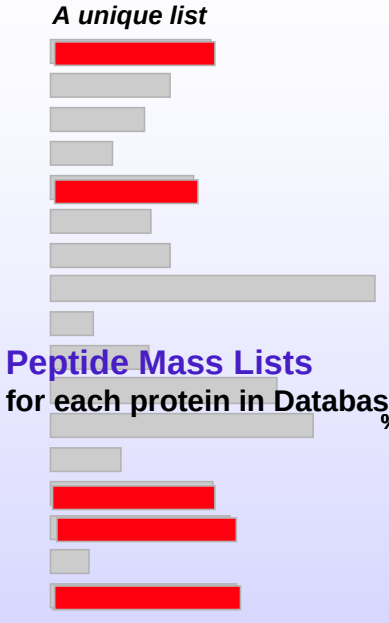
# Step 1: Enzyme digestion or chemical fragmentation



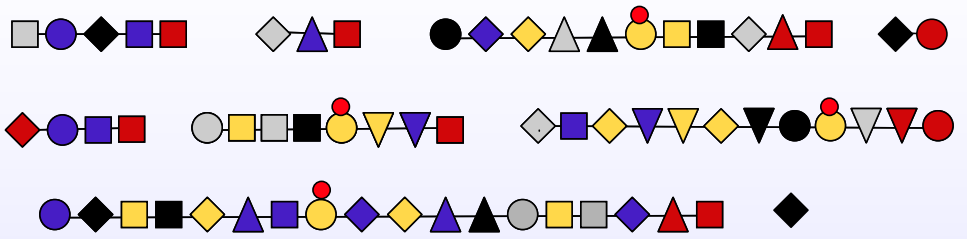

**Every protein generate a set of unique peptides**

每一個蛋白質有一套獨特的peptide碎片

# Step 2: Mass spectrometry analysis



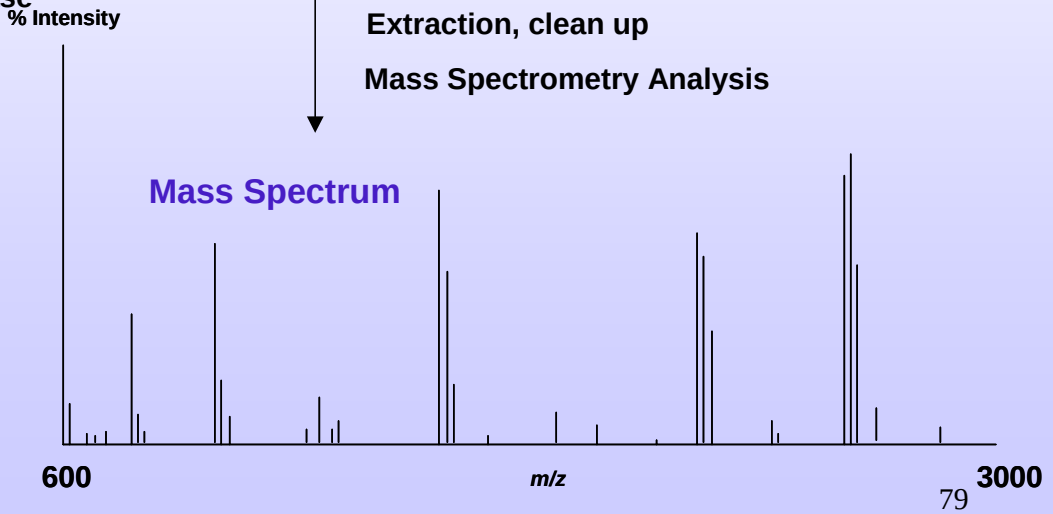

**Peptide Mass Lists**
**for each protein in Database**



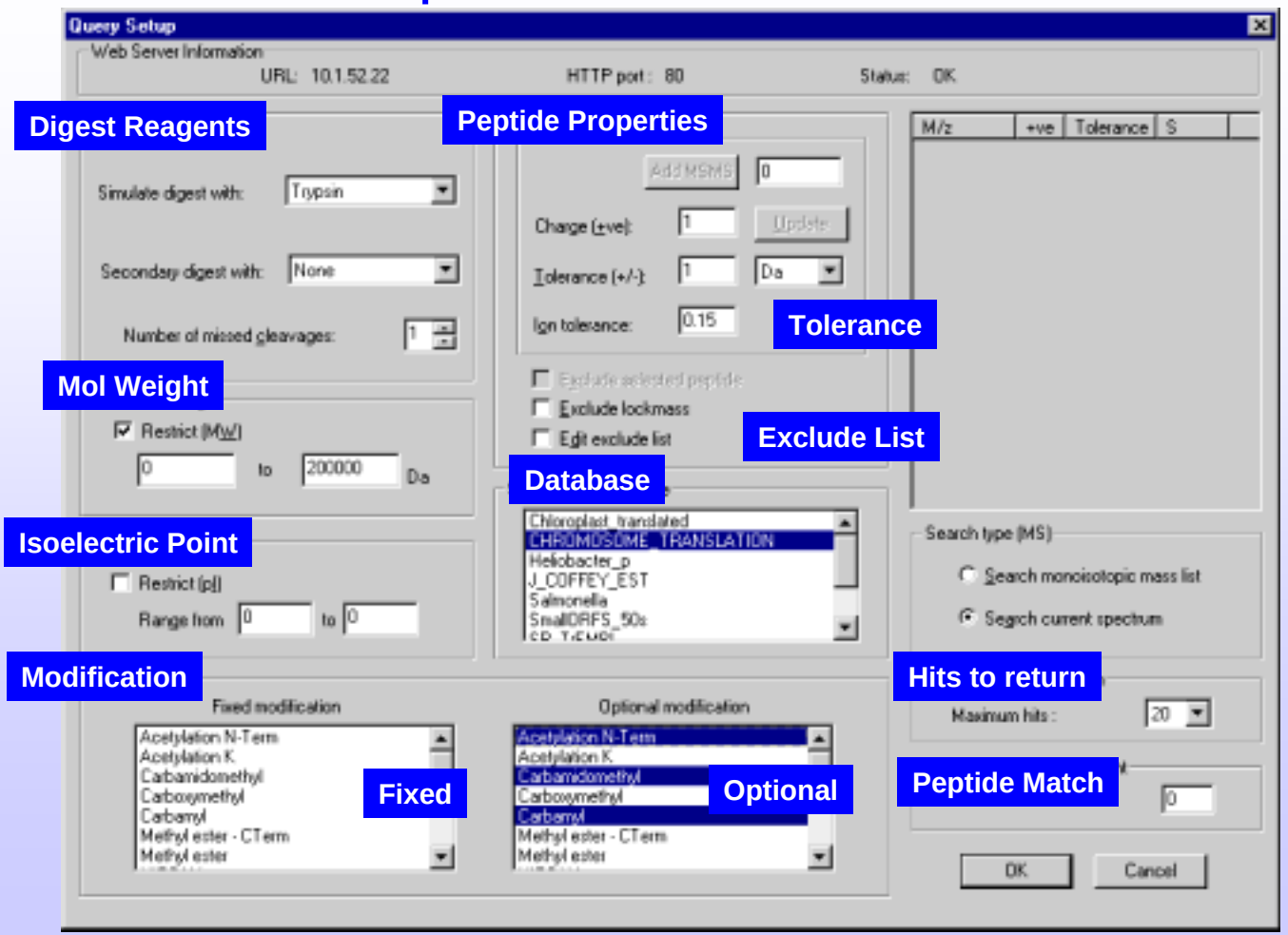
**Peptide Mixtures**

**Extraction, clean up**

**Mass Spectrometry Analysis**



# Step 3: Database match

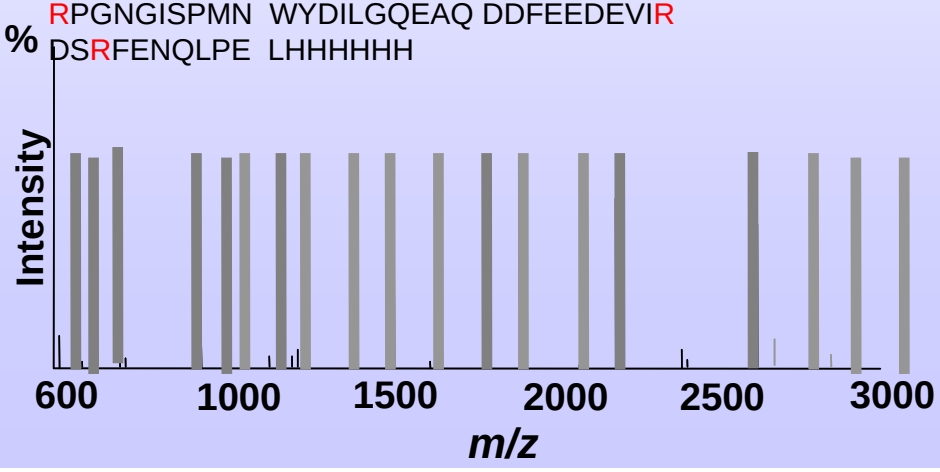


# Step 3: In-Silica digestion

## Protein Sequence

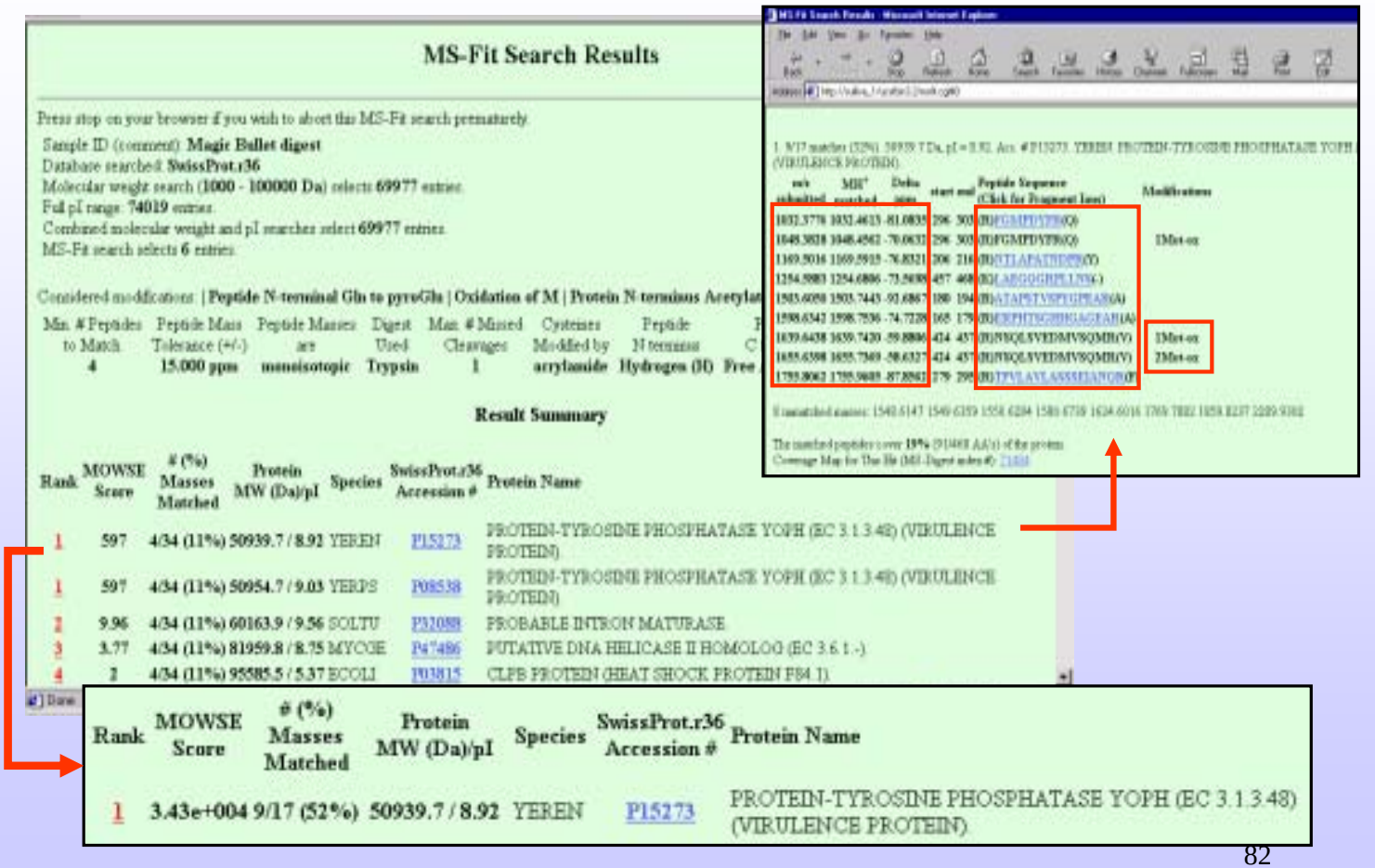
MVYIIAEIGC NHNGDINLAK KMVDVAVSCG
VDAVKFQTFK AEKLISKFAP KAEYQKATTG
TADSQLEMTK RLELSFEEYL EMRDYAISKG
VETFSTPFDE ESLEFLISTD MPIYKIPSGE
ITNLPYLEKI GKQQKKVILS TGMAVMEEIH
QAVNILRQNG TTDISILHCT TEYPTPYPSL
NLNVIHTLKD EFKDLTIGYS DHSIGSEVPI
AAAAMGAEVI EKHFTLDTNM
EVPDHKASAT
PDILAALVKG FALLNQALGR FEKIPDPVEE
KNKIVARKSV VALKPIKKGD IYSIENITVK
RPGNGISPMN WYDILGQEAQ DDFEEDEVIR
DSRFENQLPE LHHHHHH

Digestion



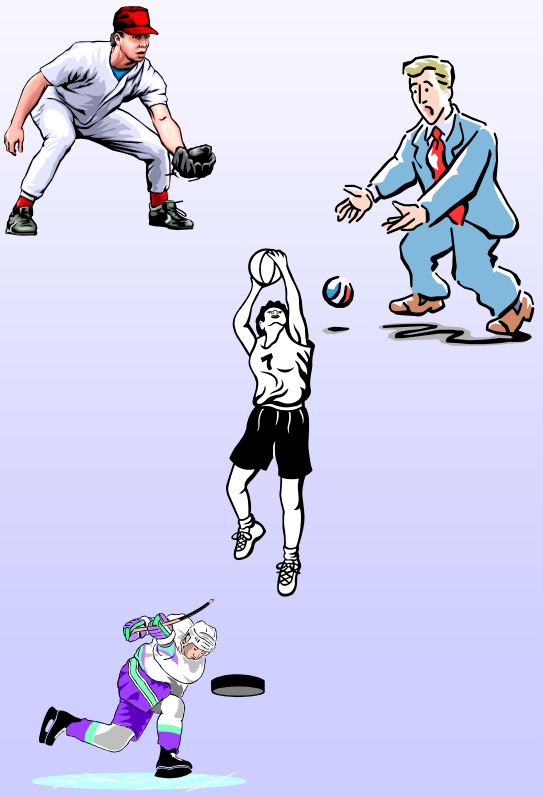

| Mass | Peptide Sequence |
|---|---|
| 3583.8 | QNGTTDISILHCTTEYPTPY PSLNLNVIHTLK |
| 3493.6 | RPGNGISPMNWYDILGQEAQ DDFEEDEVIR |
| 2995.4 | GVETFSTPFDEESLEFLIST DMPIYK |
| 2944.5 | DLTIGYSDHSIGSEVPIAAA AMGAEVIEK |
| 2324.2 | VILSTGMAVMEEIHQAVNIL R |
| 2188.1 | MVYIIAEIGCNHNGDINLAK |
| 1811.8 | FENQLPELHHHHHH |
| 1683.8 | HFTLDTNMEVPDHK |
| 1573.8 | IPSGEITNLPYLEK |
| 1558.7 | LELSFEEYLEMR |
| 1453.7 | ATTGTADSQLEMTK |
| 1392.7 | MVDVAVSCGVDAVK |
| 1351.7 | GDIYSIENITVK |
| 1269.7 | ASATPDILAALVK |
| 1159.7 | GFALLNQALGR |
| 954.63 | SVVALKPIK |
| 926.48 | IPDPVEEK |
| 696.36 | DYAISK |
| 670.36 | FQTFK |
| 638.31 | AEYQK |
| 538.25 | DEFK |

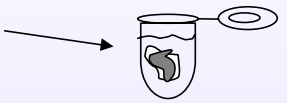
# Search Result

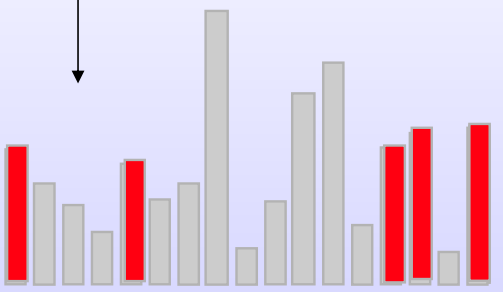


# Peptide Mass Fingerprinting (PMF)
# 蛋白質身份鑑定

In-Gel Digestion





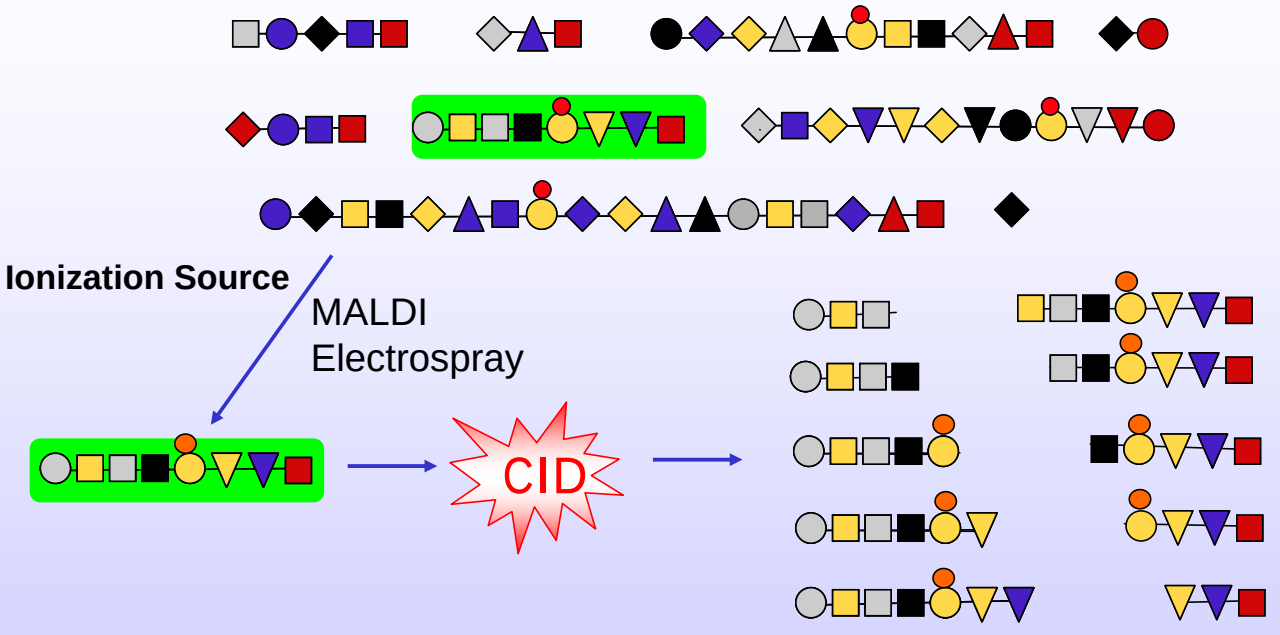


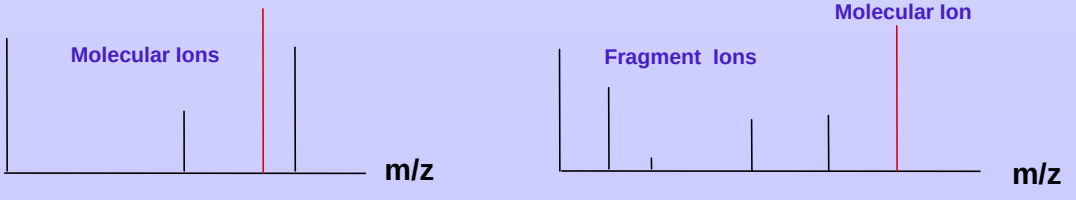
Extract peptides; mass analyze

Each protein has a unique peptide mass list

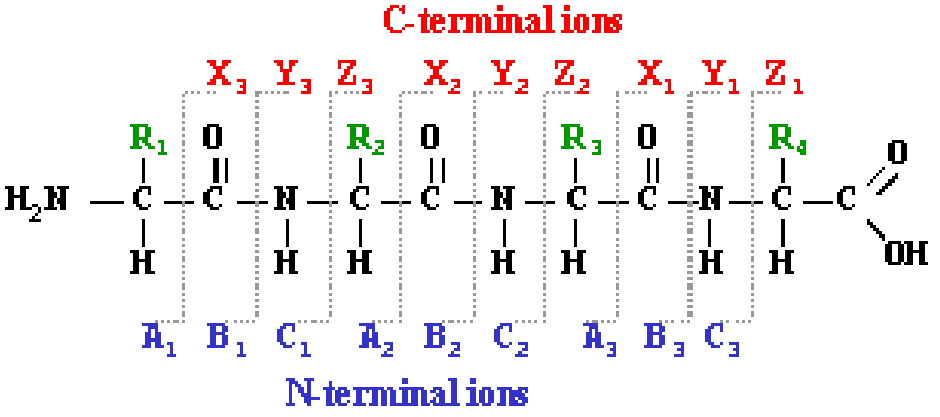
Database search

# Tandem Mass Spectrometry (MS/MS)



# Peptide Fragmentation




a,b,c – N-terminal side

x,y,z – C-terminal side

S-P-A-F-D-S-I-M-A-E-T-L-K

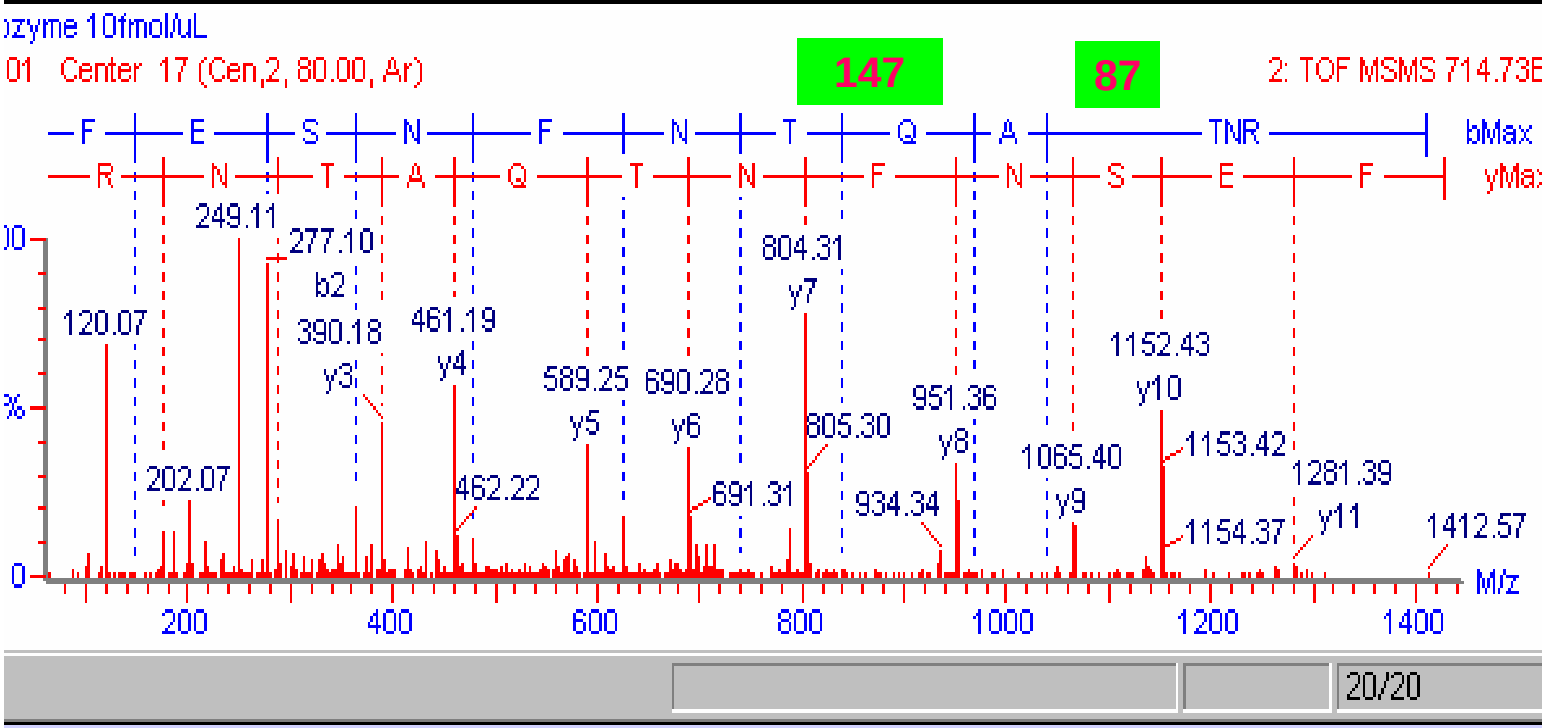
(protonated mass 1410.6)

| mass+ | b-ions | y-ions | mass+ |
|---|---|---|---|
| 88.1 | S | PAFDSIMAETLK | 1323.6 |
| 185.2 | SP | AFDSIMAETLK | 1226.4 |
| 256.3 | SPA | FDSIMAETLK | 1155.4 |
| 403.5 | SPAF | DSIMAETLK | 1008.2 |
| 518.5 | SPAFD | SIMAETLK | 893.1 |
| 605.6 | SPAFDS | IMAETLK | 806.0 |
| 718.8 | SPAFDSI | MAETLK | 692.3 |
| 850.0 | SPAFDSIM | AETLK | 561.7 |
| 921.1 | SPAFDSIMA | ETLK | 490.6 |
| 1050.2 | SPAFDSIMAE | TLK | 361.5 |
| 1151.3 | SPAFDSIMAET | LK | 260.4 |
| 1264.4 | SPAFDSIMAETL | K | 147.2 |

# Residue Mass of Amino Acids

| Table9.2 | Symbols and residue masses of the protein amino acids | | |
|---|---|---|---|
| Name | Symbol | Residue mass | Side-chain |
| Alanine | A,Ala | 71.079 | $CH_3$- |
| Arginine | R,Arg | 156.188 | $HN=C(NH_2)-N-(CH_2)_3$ |
| Asparagine | N,Asn | 114.104 | $H_2N-CO-CH_2$- |
| Aspartic acid | D,Asp | 115.089 | $HOOC-CH_2$- |
| Cysteine | C,Cys | 103.145 | $HS-CH_2$- |
| Glutamine | Q,Gln | 128.131 | $H_2N-CO-(CH_2)_2$- |
| Glutamic acid | E,Glu | 129.116 | $HOOC-(CH_2)_2$- |
| Glycine | G,Gly | 57.052 | H- |
| Histidine | H,His | 137.141 | Imidazole-$CH_2$- |
| Isoleucine | I,Ile | 113.16 | $CH_3-CH_2-CH(CH_3)$- |
| Leucine | L,Leu | 113.16 | $(CH_3)_2-CH-CH_2$- |
| Lysine | K,Lys | 128.17 | $H_2N-(CH_2)_4$- |
| Methionine | M,Met | 131.199 | $CH_3-S-(CH_2)_2$- |
| Metsulphoxide | Met.SO | 147.199 | $CH_3-S(O)-(CD_2)_2$- |
| Phenylalanine | F,Phe | 147.177 | Phenyl-$CH_2$- |
| Proline | P,Pro | 97.117 | Phrrolidone-CH- |
| Serine | S,Ser | 87.078 | $HO-CH_2$- |
| Threonine | T,Thr | 101.105 | $CH_3-CH(OH)$- |
| Tryptophan | W,Trp | 186.213 | ndole-NH-CH=C-CH |
| Tyrosine | Y,Tyr | 163.176 | 4-OH-Phenyl-$CH_2$- |
| Valine | V,Val | 99.133 | $CH_3-CH(CH_2)$- |

# Peptide Sequencing


F (phenylalanine): 147.177
S (Serine) : 87.078

# Separation and Quantification Techniques for Mass Spectrometry

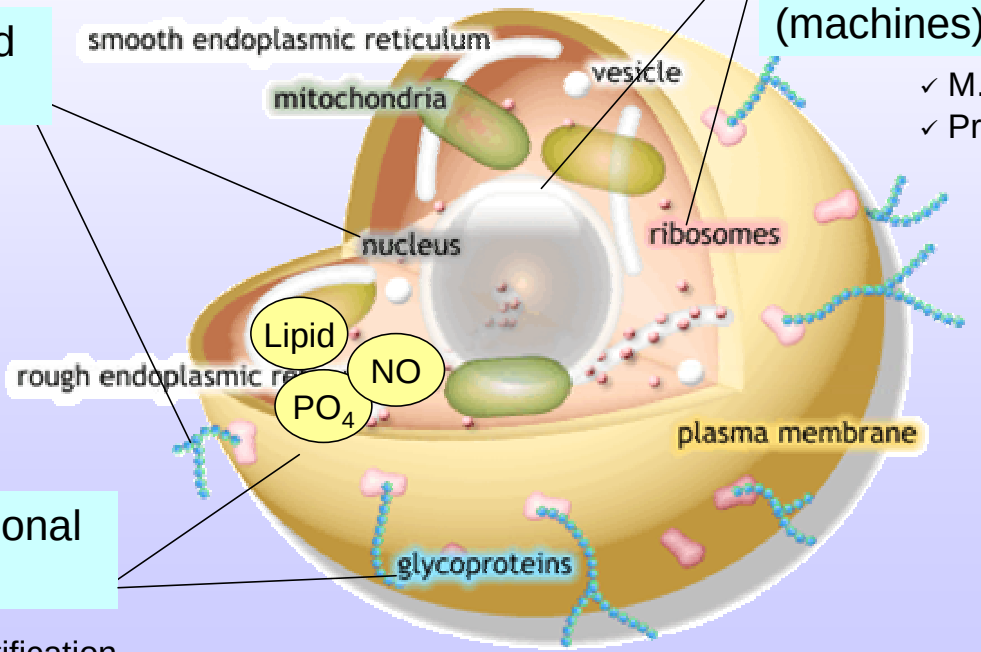
# Diverse Properties of Proteins v.s. Mass Spectrometry

- ✓ M.W. measurement
- ✓ Protein identification
- ✓ tandem mass spectrometry

Protein-ligand interaction

Protein Complex (machines)

- ✓ M.W. measurement
- ✓ Protein identification




Post-translational modification

- ✓ Protein identification
- ✓ Tandem mass spectrometry

# What Do We See?

***Technology Platform*** V.S. ***Complex Proteome***

# Proteomic technologies for biomarker detection & discovery

- Gel electrophoresis
  - ✓ Two-dimensional gel electrophoresis (2DE)
  - ✓ Two-dimensional differential gel electrophoresis (DIGE)
- Gel free ( Mass spectrometry)
  - ✓ Surface-enhanced laser desorption / ionization TOF MS (SELDI-TOF MS)
  - ✓ Multidimentional protein identification technology (MudPIT)
  - ✓ Isotope coded affinity tag (ICAT)
- Other technologies:
  - ✓ Protein microarrays
  - ✓ ELISA

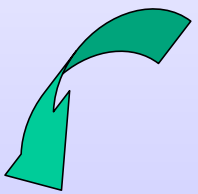
## Number of Proteins per Genome

- Haemophilus 1742
- *E. coli* 4413
- Yeast 6600
- Caenorhabditis 18000
- Drosophila 13000
- Human >100000

# A key technical challenge in proteomics

A more complex biological problem

Diverse solution
($10^5$ – $10^6$ for protein abundance)

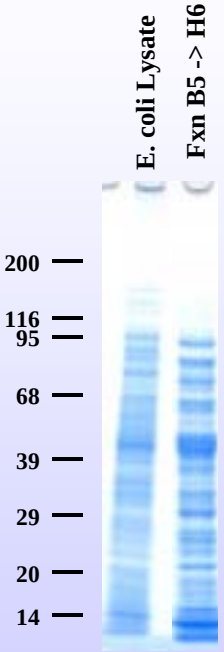


Mass-spectrometry strategy for more complex mixture ?

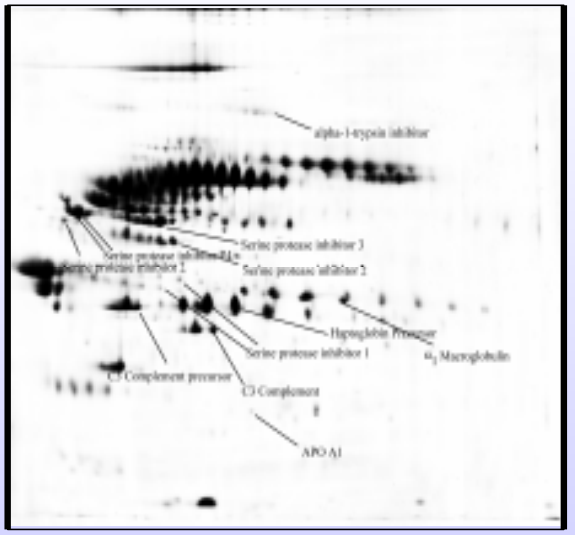
+

Sample prefractionation

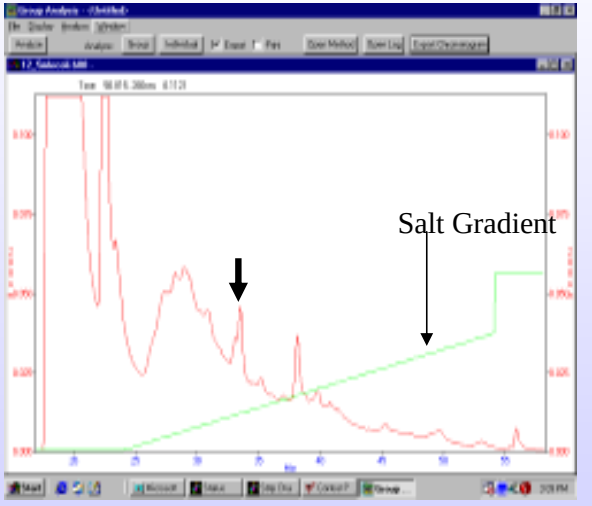
# Protein Separation




SDS-PAGE

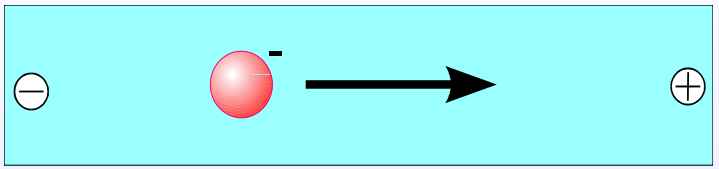


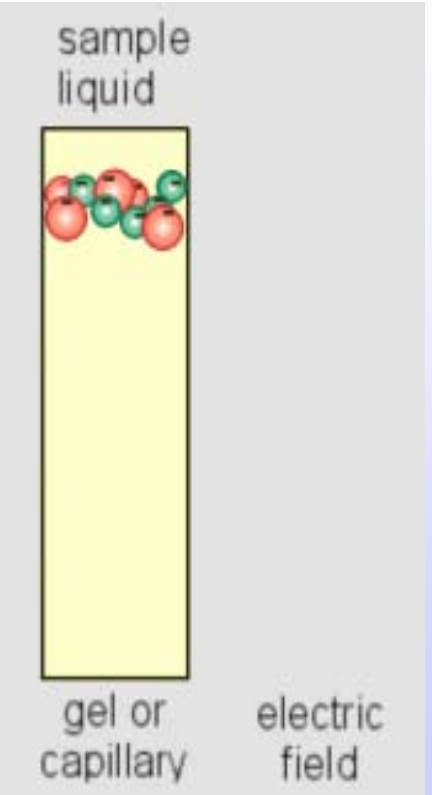

2D-gel




Chromatography

# Principle of Electrophoresis





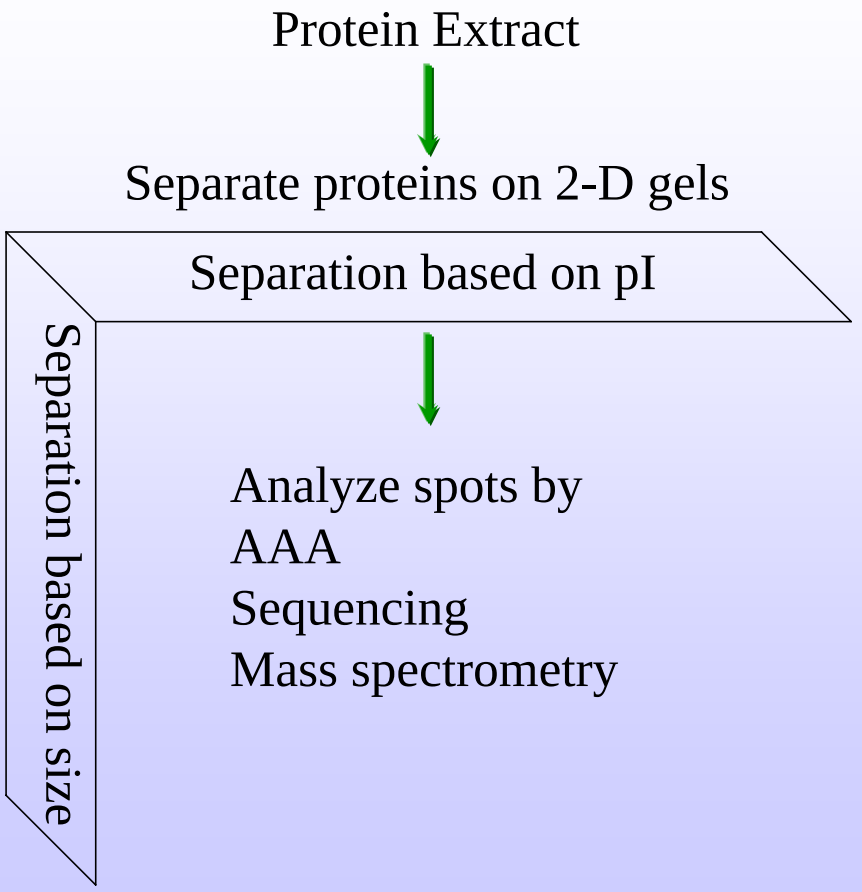

$$v = \mu ' E$$

v = migration velocity (cm/s)

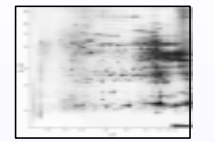
$\mu$ = electrophoretic mobility ($cm^2/V's$)

charge, size, shape of molecule, viscosity, pore size, buffer pH and ionic strength, temperature of medium

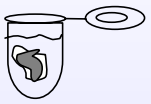
# Two-dimensional gel electrophoresis



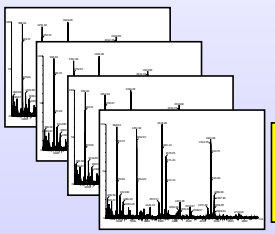
# 2D-PAGE　(二維電泳分離)



Run gel;
stain;

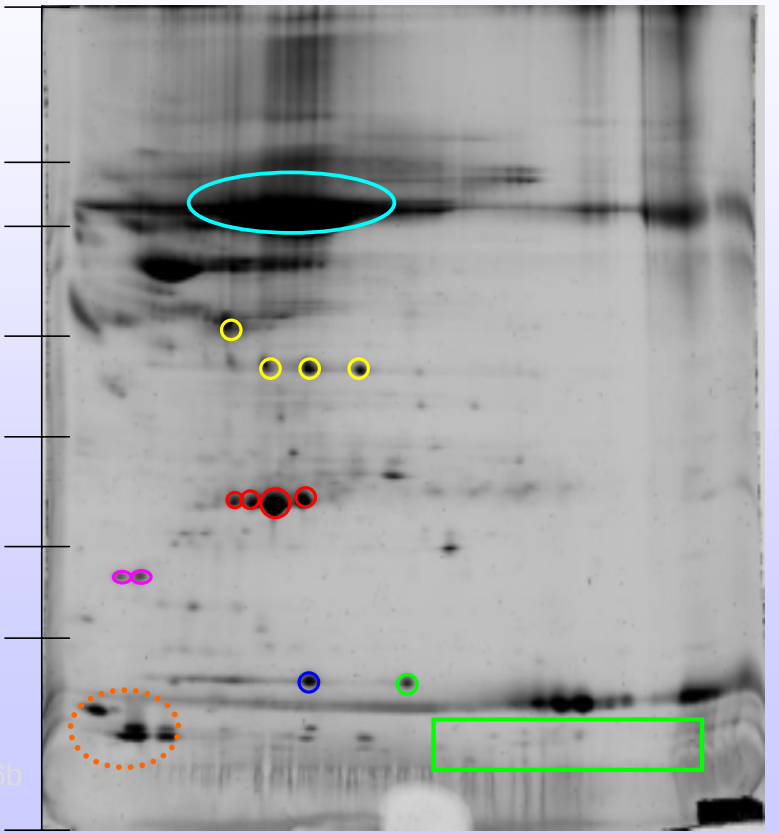


In-Gel Digestion
and sample prep
for MS



Extract peptides;
mass analyze

Database search



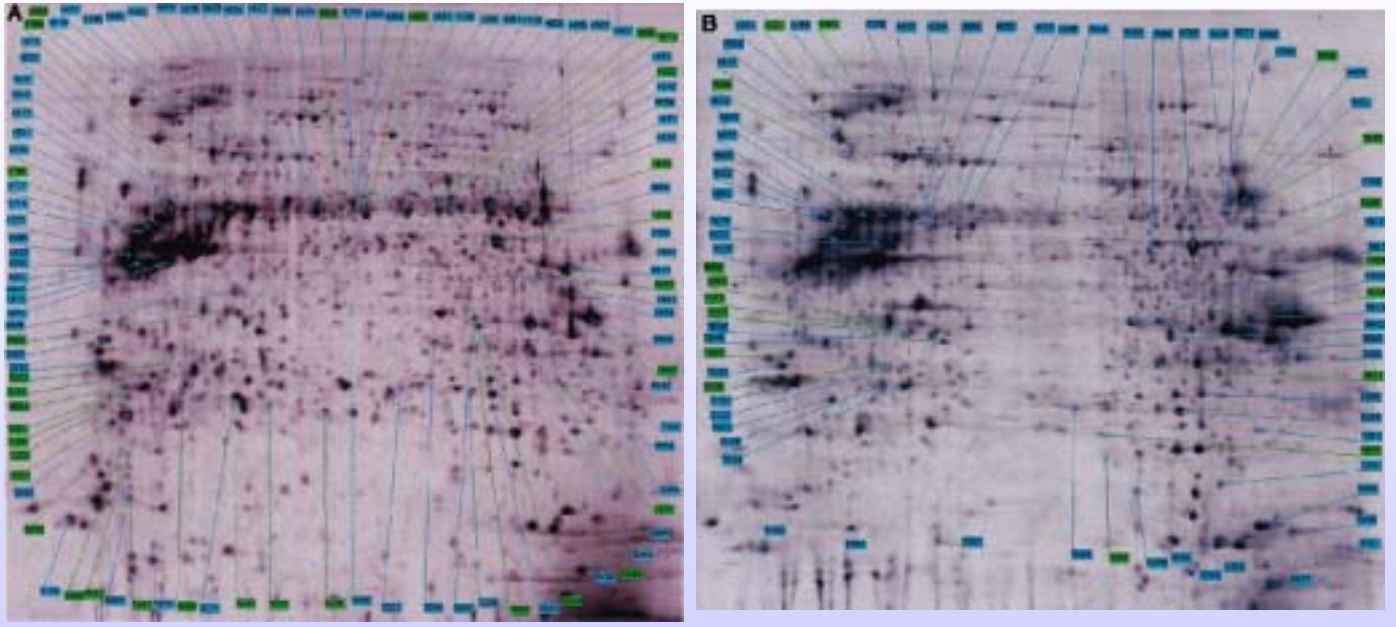
# Separation methods for protein and peptides

| Method | Basis of separation |
|---|---|
| **Chromatographic methods** | |
| size-exclusion chromatography | molecular weight |
| ion-exchange chromatography | charge |
| reverse-phase high-performance chromatograph | hydrophobic interaction between the sample and bonded |
| hydrophobic-interaction chromatography | salt-promoted adsorption chromatography |
| affinity chromatography | biomolecular interaction (DNA, ligand, antibody⋯) |
| **Electrophoretic methods** | |
| one-dimensional gel electrophoresis | molecular weight |
| two-dimensional gel electrophoresis | 1st dimension: charge; 2nd dimension: molecular weigh |
| gel-free isoelectric focusing | charge |
| **Subcellular fractionation** | subcellular fractionation |

# (二維電泳分離)

## Example:

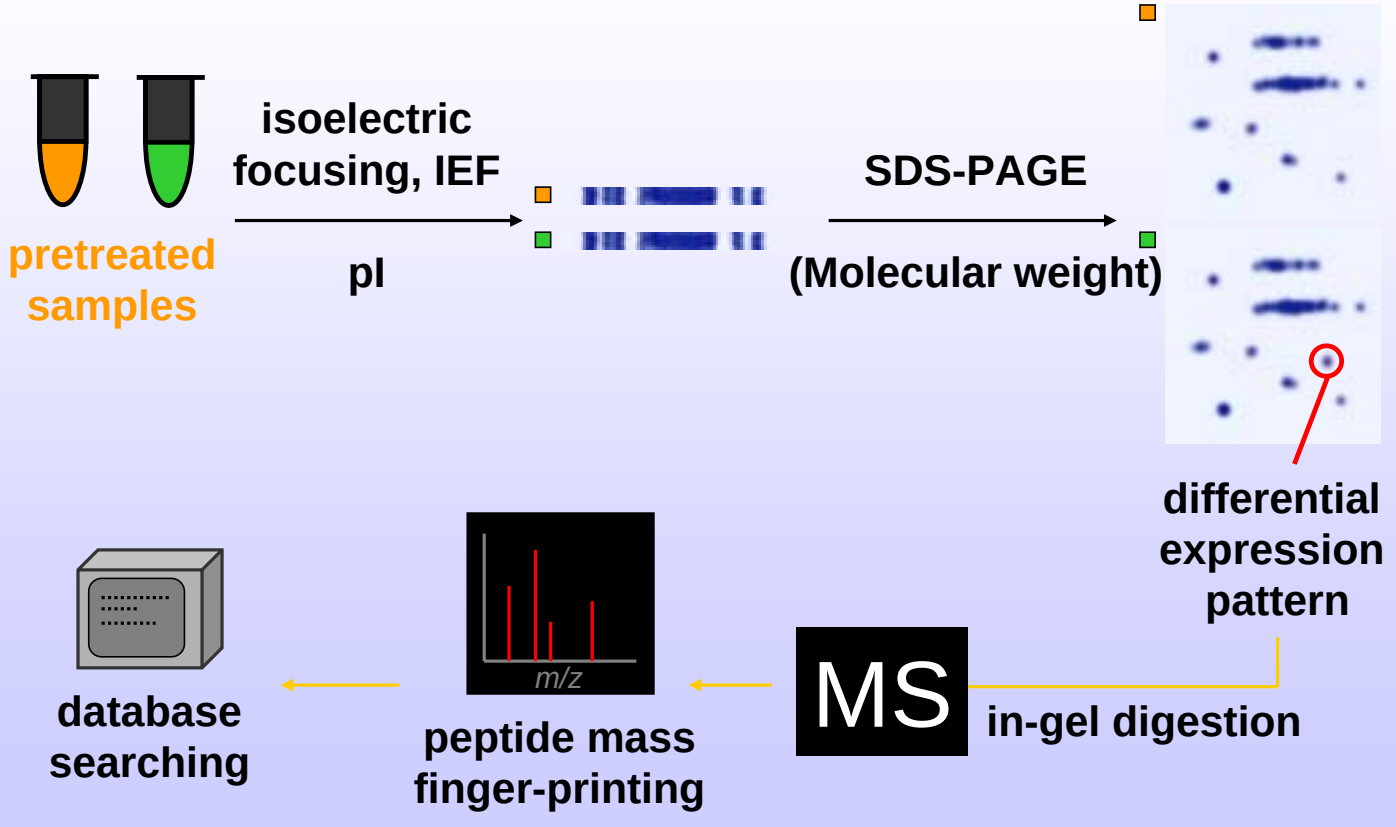
### *Today's 2D Gel-Based Proteomics*

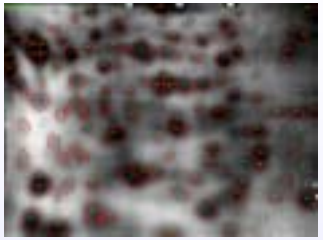



differentially expressed (170)

no differences

# 2D-PAGE & protein identification



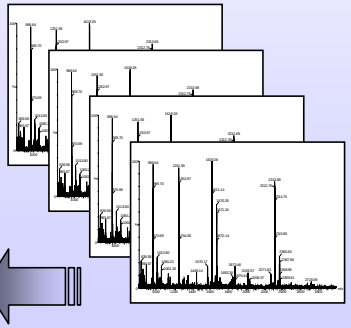


2D Gel

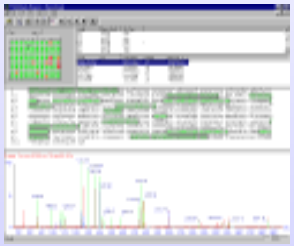
Robotic spot excision

In-Gel Digestion and sample prep for MS

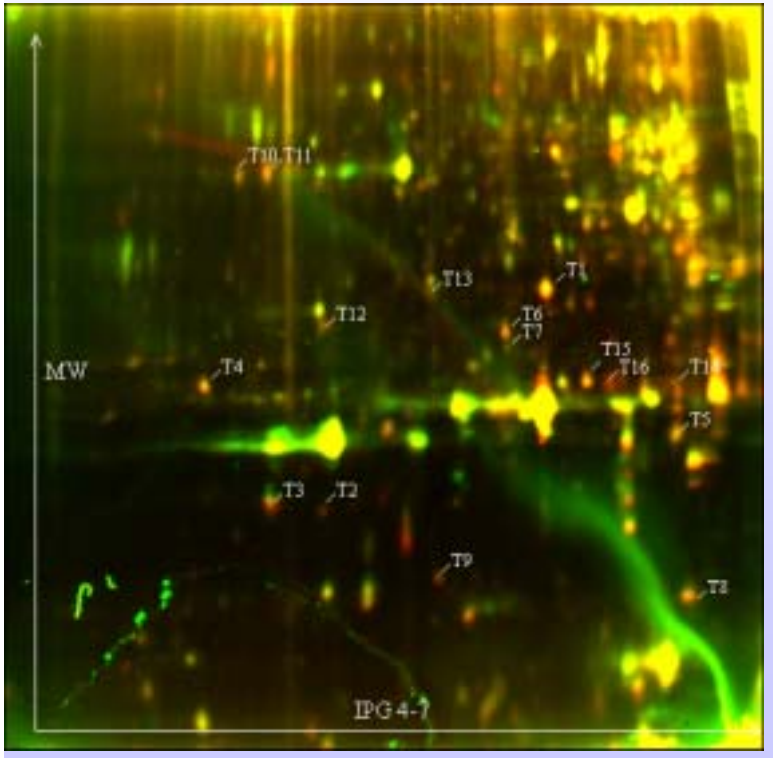
Advanced MS



Automated data processing and database searching

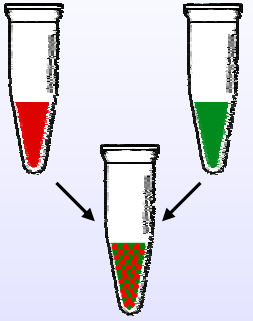


# Differential gel electrophoresis (2D-DIGE)




**Test labelled with propyl-Cy3**

**Control labelled with methyl-Cy5**



**2D-electrophoresis**

**IEF**

**SDS-PAGE**

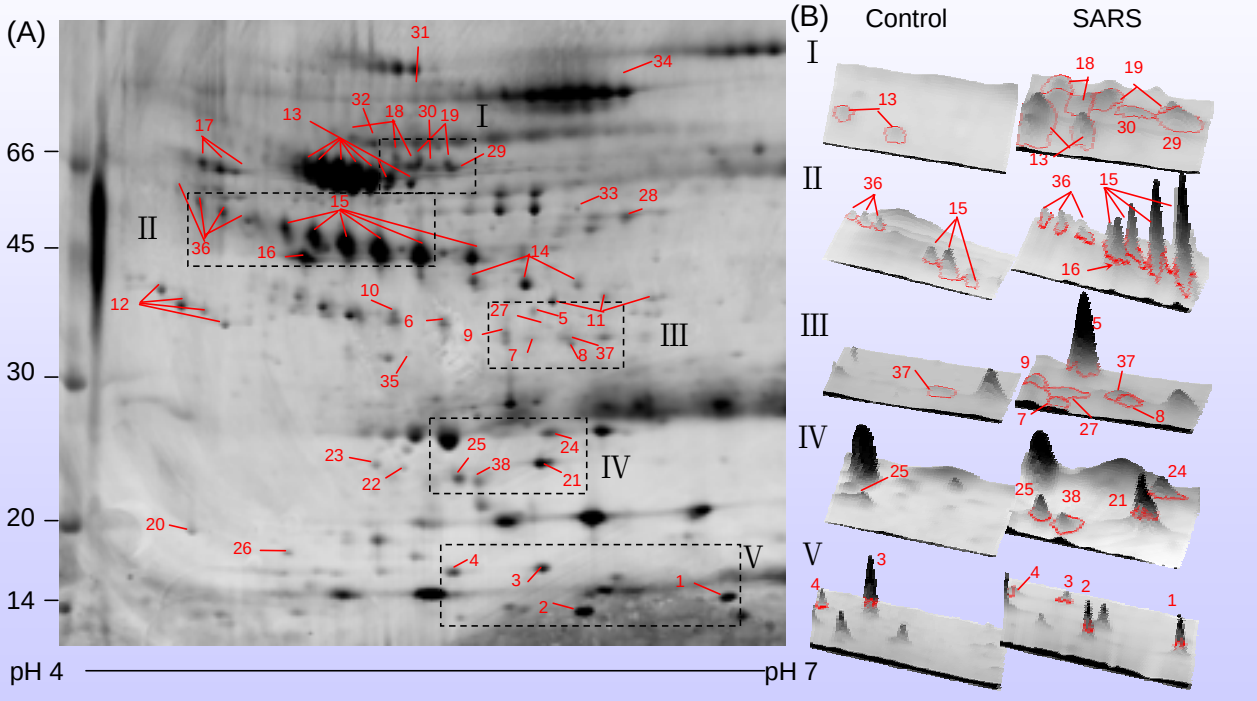
**λE Cy3 Image A**

**λE Cy5 Image B**

# 2D-PAGE & protein identification

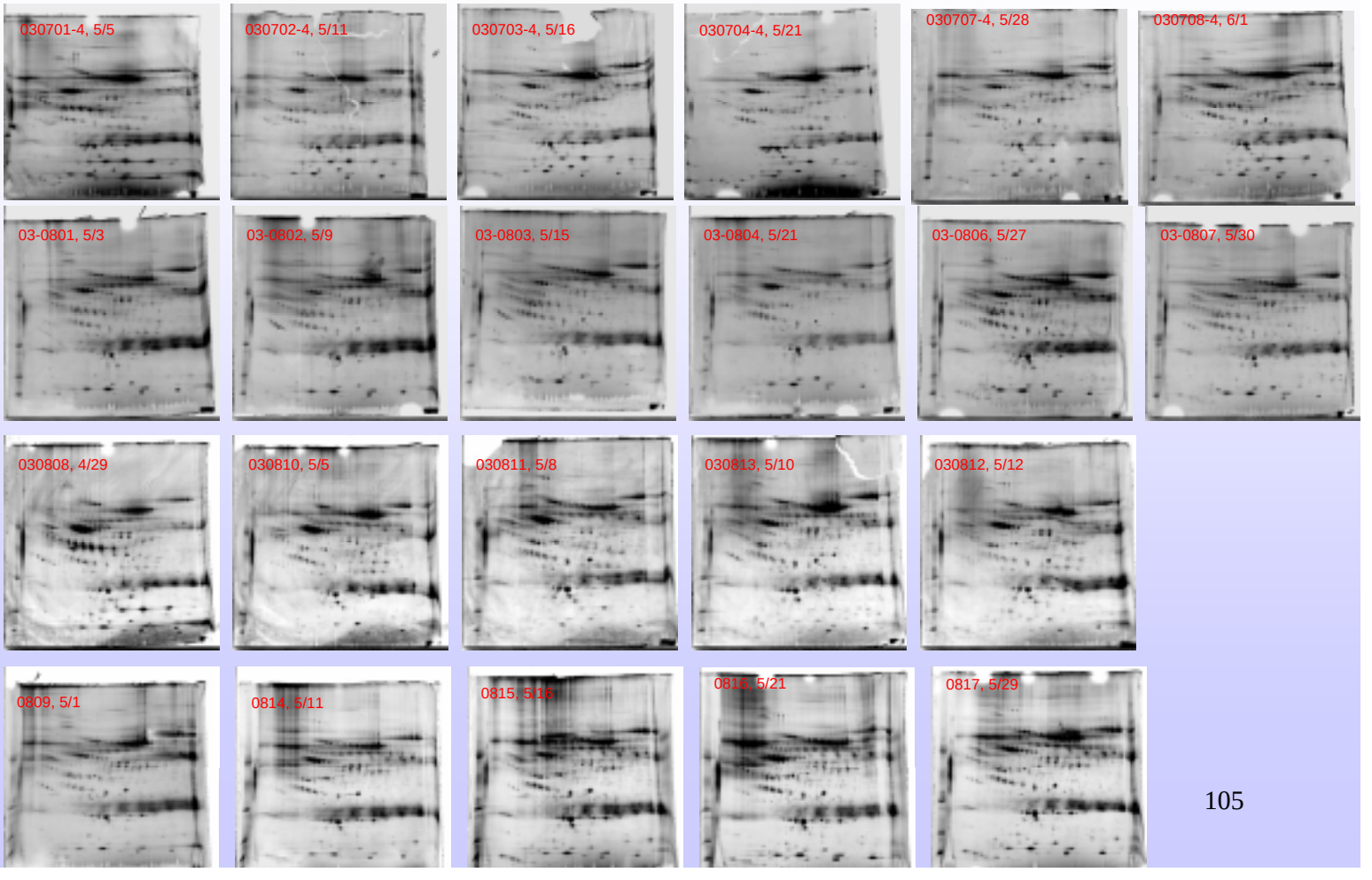
- MALDI-TOF MS is mostly used
- advantages:
  - discovering biomarker
  - high reproducibility
  - semi-quantitative
- drawbacks:
  - low sensitivity, particularly for less-abundant proteins

# Plasma proteome of Severe Acute Respiratory Syndrome (SARS) analyzed by 2-DE and MS



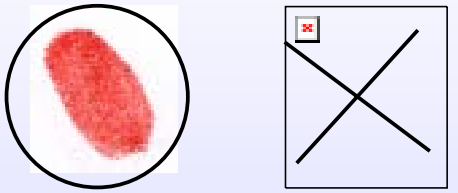

*Proc. Natl.* **Acad. Sci.** *USA.*, **101**, 17039 (2004)

# Protein Profiles of SARS Patients



## MALDI

Peptide Mass Fingerprinting (PMF)

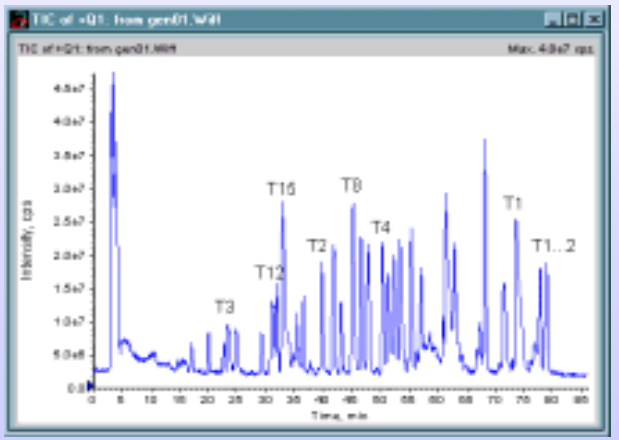


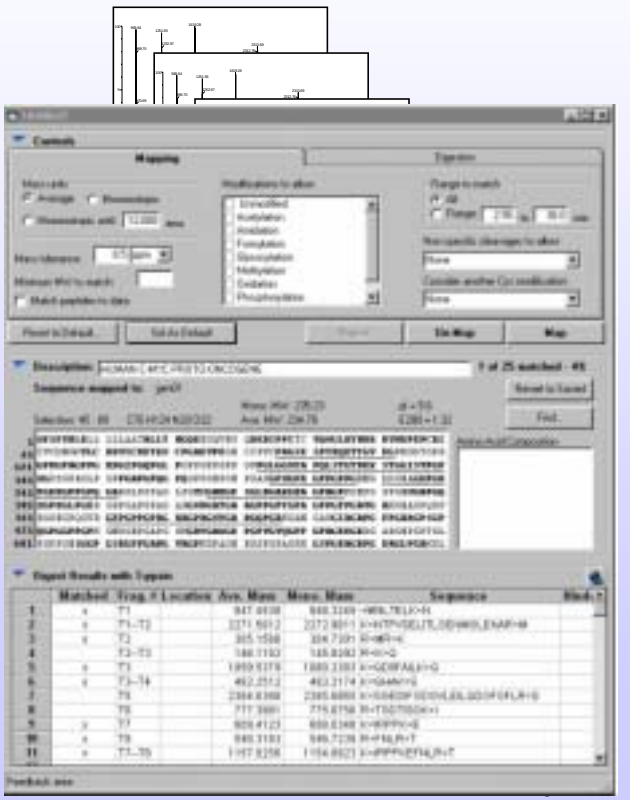
## Direct LC-MS/MS

Poor salt tolerance
Ion suppression

# 液相層析輔助質譜

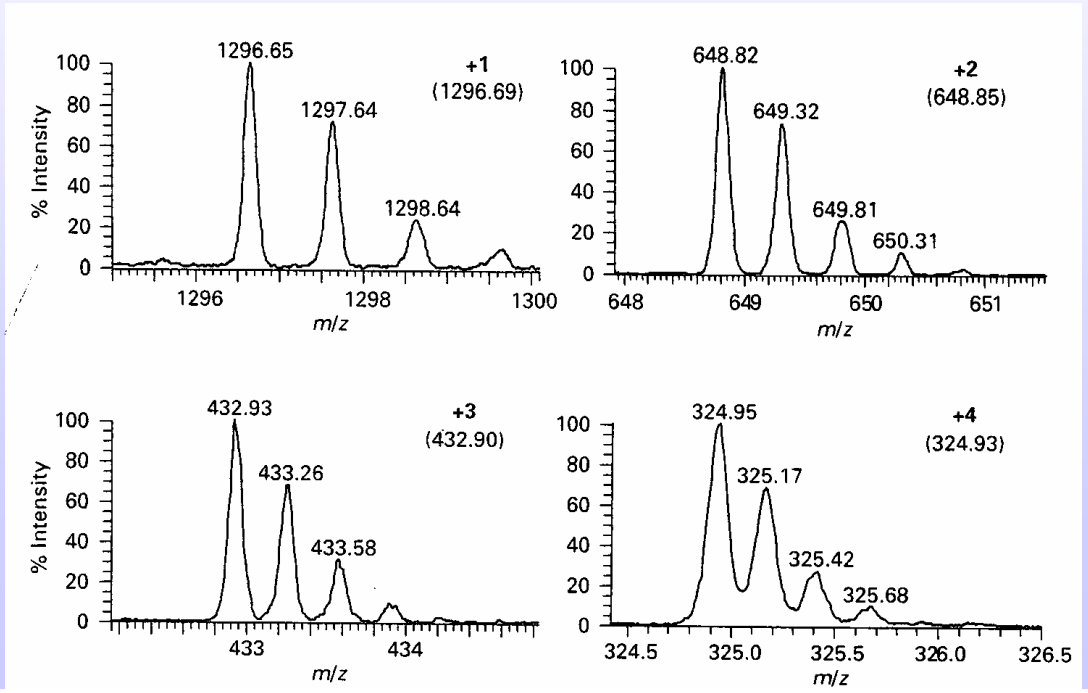
Ion suppression
Salt tolerance
Sample concentration

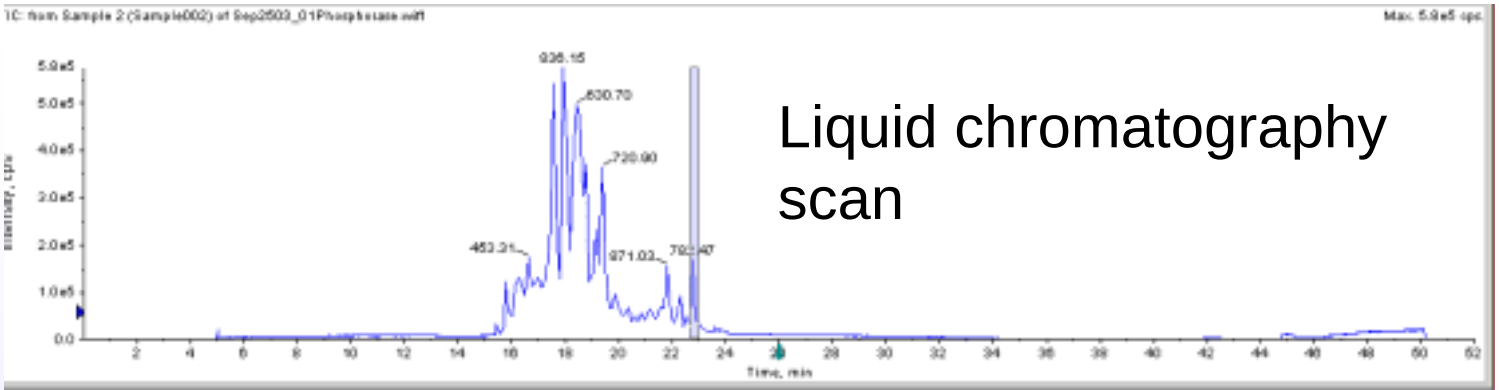


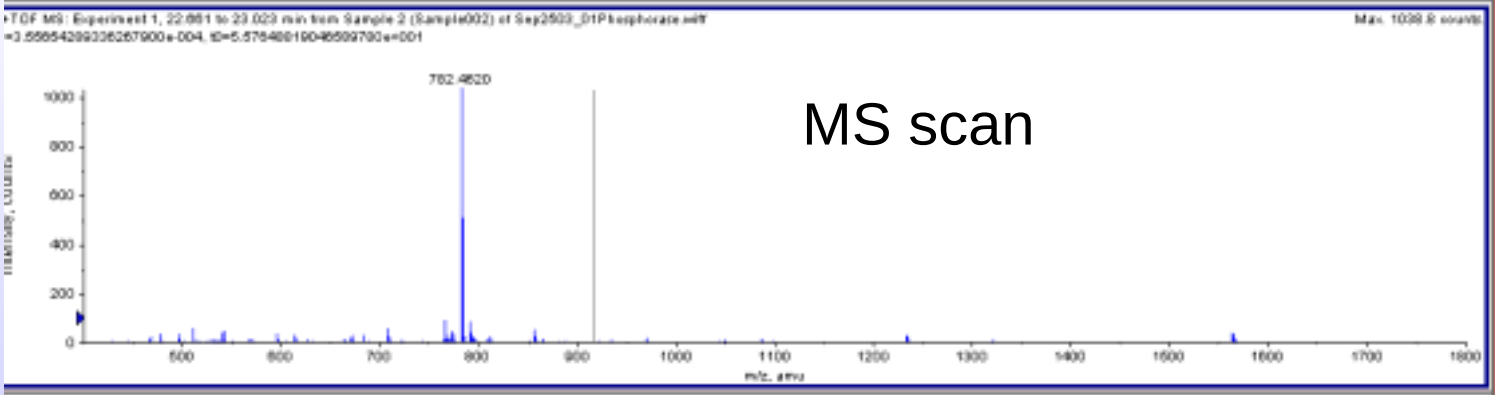
| Table9.3 | Mass differences due to isotopes in multiply charged p |
|---|---|

| Charge on peptide | Apparent mass | Mass difference between isotope |
|---|---|---|
| Single charge | [(M+H)/1] | 1Da |
| Double charge | [(M+2H)/2] | 0.5Da |
| Triple charge | [(M+3H)/3] | 0.33Da |
| n charges | [(M+nH)/n] | 1/nDa |

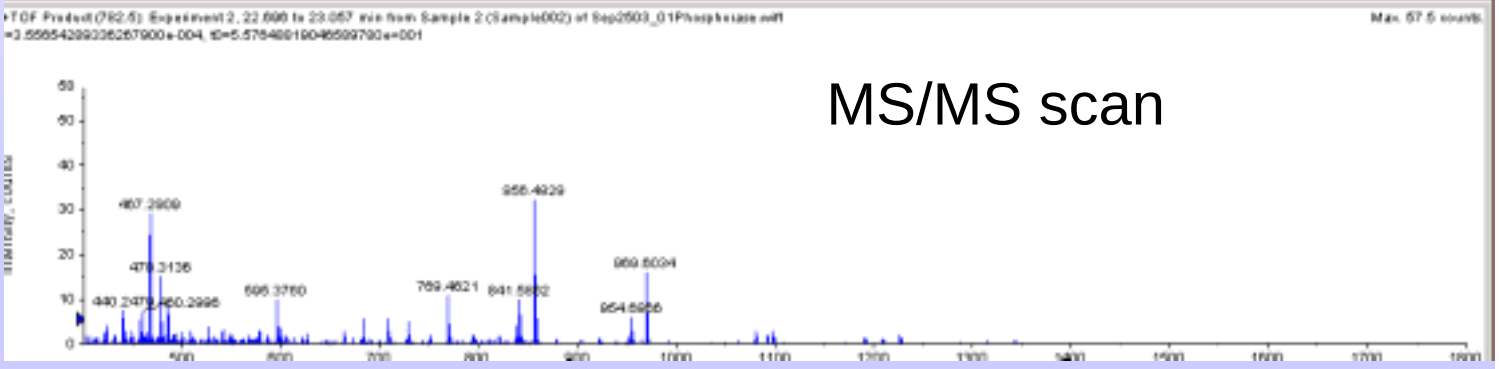




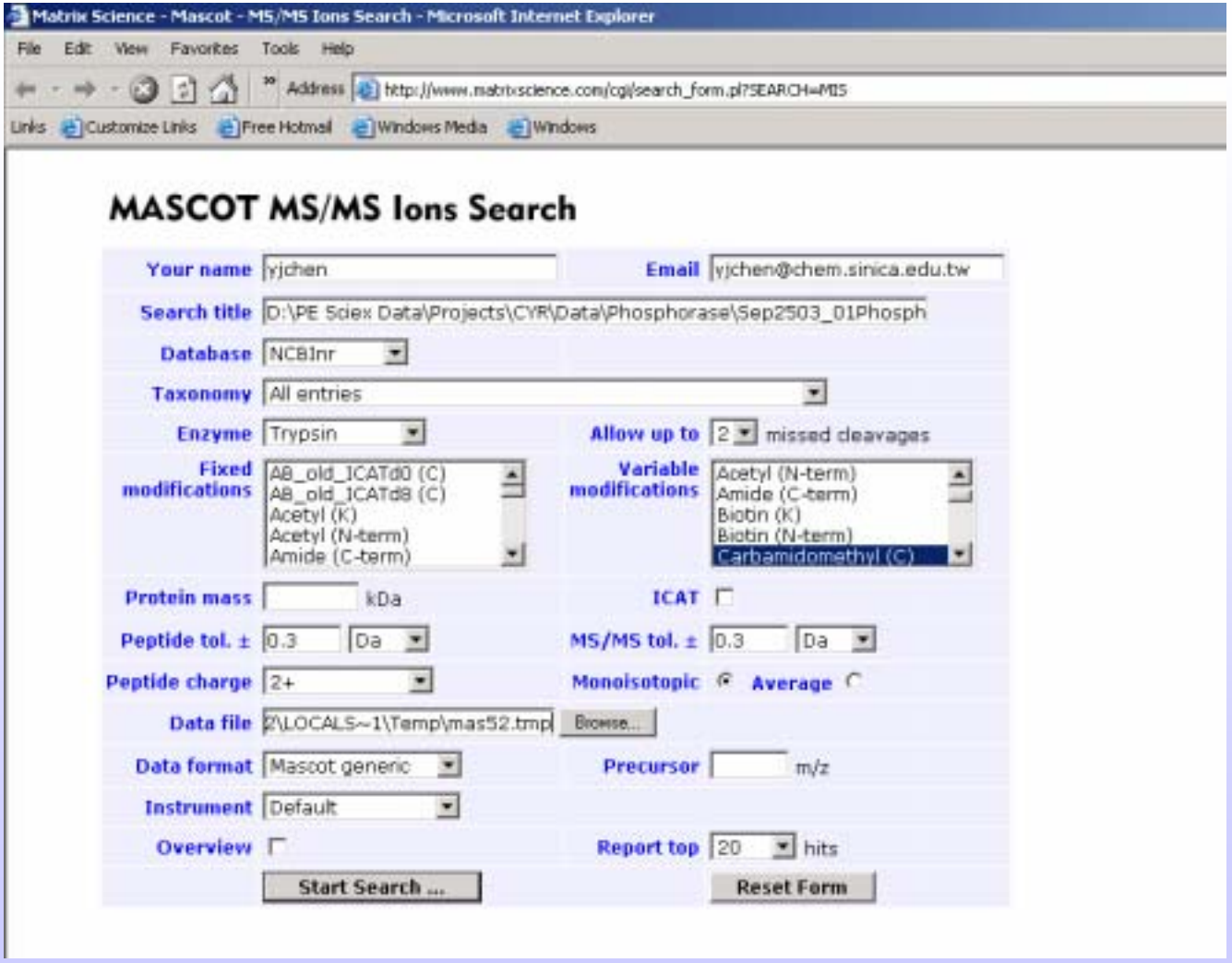

Liquid chromatography scan




MS scan




MS/MS scan

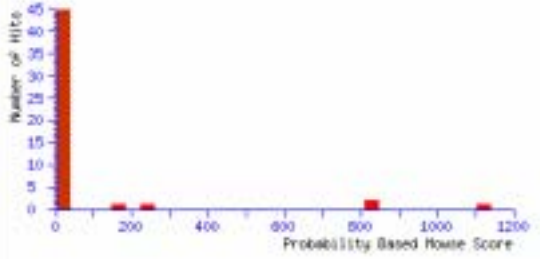

Matrix Science - Mascot - MS/MS Ions Search - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://www.matrixscience.com/cgi/search_form.pl?SEARCH=MIS
Links Customize Links Free Hotmail Windows Media Windows
MASCOT MS/MS Ions Search
Your name yjchen
Email yjchen@chem.sinica.edu.tw
Search title D:\PE Sciex Data\Projects\CYR\Data\Phosphorase\Sep2503_01Phosph
Database NCBInr
Taxonomy All entries
Enzyme Trypsin
Allow up to 2 missed cleavages
Fixed modifications
AB_old_ICATd0 (C)
AB_old_ICATd8 (C)
Acetyl (K)
Acetyl (N-term)
Amide (C-term)
Variable modifications
Acetyl (N-term)
Amide (C-term)
Biotin (K)
Biotin (N-term)
Carbamidomethyl (C)
Protein mass kDa
ICAT
Peptide tol. ± 0.3 Da
MS/MS tol. ± 0.3 Da
Peptide charge 2+
Monoisotopic Average
Data file 2\LOCALS~1\Temp\mas52.tmp Browse...
Data format Mascot generic
Precursor m/z
Instrument Default
Overview
Report top 20 hits
Start Search ...
Reset Form

# Mascot Search Results

```
User            : yjchen
Email           : yjchen@chem.sinica.edu.tw
Search title    : D:\PE Sciex Data\Projects\CYR\Data\Phosphorase\Sep2503_01Phosphorase.wiff : Sample002
MS data file    : C:\DOCUME~1\lab212\LOCALS~1\Temp\mas52.tmp
Database        : SwissProt 42.10 (184535 sequences; 83187710 residues)
Timestamp       : 24 Feb 2004 at 12:03:33 GMT
Significant hits: P00489          (PYG2_RABIT) Glycogen phosphorylase, muscle form (EC 2.4.1.1) (Myophosphorylase) Glycogen phosphorylase, musc
                  P11217-00-00-00 (PYG2_HUMAN) Splice isoform Displayed; Variant Displayed; Conflict Displayed; from P11217 Glycogen phosphoryl
                  O18751          (PYG2_SHEEP) Glycogen phosphorylase, muscle form (EC 2.4.1.1) (Myophosphorylase) Glycogen phosphorylase, musc
                  P11216          (PYG3_HUMAN) Glycogen phosphorylase, brain form (EC 2.4.1.1) Glycogen phosphorylase, brain form (EC 2.4.1.1)
                  P06737-00-02-00 (PYG1_HUMAN) Splice isoform Displayed; Variant GSD-VI-VAR_007909; Conflict Displayed; from P06737 Glycogen ph
                  P04191-00-00-00 (ATA1_RABIT) Splice isoform SERCA1B; Variant Displayed; Conflict Displayed; from P04191 Sarcoplasmic/endoplasm
                  Q85FB6          (RPOC_CYAME) Bifunctional DNA-directed RNA polymerase beta' and beta'' chain (EC 2.7.7.6) (PEP) [Includes: DN
                  P22703          (RPOD_AMASP) DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (RNAP beta' subunit) (Transcriptase beta' c
```

## Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 38 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.




## Peptide Summary Report

Switch to Protein Summary Report

To create a bookmark for this report, right click this link: Peptide Summary Report (D:\PE Sciex Data\Projects\CYR\Data\Phosphorase\Sep2503_01Phosphorase.wiff : Sample002)

Select All | Select None | Search Selected | ☐ Error tolerant

1. P00489 **Mass:** 97097 **Score:** 1121 **Peptides matched:** 31

(PHS2_RABIT) Glycogen phosphorylase, muscle form (EC 2.4.1.1) (Myophosphorylase) Glycogen phospl

☐ Check to include this hit in error tolerant search

| | Query | Observed | Mr(expt) | Mr(calc) | Delta | Miss | Score | Expect | Rank | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|
| ☑ | 12 | 490.33 | 978.64 | 978.54 | 0.10 | 1 | 18 | 8.1 | 1 | LPAPDEKIP |
| ☑ | 19 | 527.32 | 1052.63 | 1052.57 | 0.06 | 0 | 40 | 0.037 | 1 | VIFLENYR |
| ☑ | 20 | 527.80 | 1053.58 | 1053.48 | 0.11 | 0 | 65 | 0.00013 | 1 | TNFDAFPDK |
| | 22 | 536.84 | 1071.66 | 1071.53 | 0.13 | 0 | 11 | 39 | 5 | EIWGVEPSR |
| ☑ | 27 | 559.35 | 1116.68 | 1116.56 | 0.12 | 0 | 61 | 0.00044 | 1 | VAAAFPGDVDR |
| ☑ | 29 | 573.32 | 1144.62 | 1144.56 | 0.06 | 0 | 45 | 0.017 | 1 | YEFGIFNQK |
| ☑ | 40 | 615.40 | 1228.78 | 1228.67 | 0.11 | 0 | 82 | 3.5e-06 | 1 | GLAGVENVIELK |
| ☑ | 45 | 631.83 | 1261.64 | 1261.59 | 0.05 | 0 | 42 | 0.034 | 1 | VFADYEEYVK |
| ☑ | 46 | 421.95 | 1262.82 | 1262.70 | 0.12 | 2 | 30 | 0.45 | 1 | QRLPAPDEKIP |
| ☑ | 63 | 453.30 | 1356.89 | 1356.76 | 0.13 | 1 | 41 | 0.034 | 1 | GLAGVENVIELKK |
| ☑ | 73 | 713.95 | 1425.89 | 1425.77 | 0.11 | 0 | 70 | 4.1e-05 | 1 | MLQIIYEINQR |
| ☑ | 79 | 721.89 | 1441.76 | 1441.69 | 0.08 | 0 | 25 | 1.6 | 1 | VLYPNDNFFEGK |
| ☑ | 82 | 491.99 | 1472.95 | 1472.85 | 0.10 | 0 | 46 | 0.011 | 1 | WPVHLLETLLPR |
| ☑ | 83 | 737.50 | 1472.98 | 1472.85 | 0.13 | 0 | (45) | 0.013 | 1 | WPVHLLETLLPR |
| ☑ | 86 | 497.29 | 1488.84 | 1488.76 | 0.08 | 1 | 20 | 3.6 | 1 | LDWDKAWEVTVK |
| ☑ | 100 | 775.95 | 1549.88 | 1549.76 | 0.12 | 0 | 60 | 0.0004 | 1 | IGEEYISDLDQLR |
| ☑ | 105 | 522.98 | 1565.91 | 1565.78 | 0.12 | 0 | (63) | 0.0002 | 1 | DFNVGGYIQAVLDR |
| ☑ | 106 | 783.98 | 1565.94 | 1565.78 | 0.16 | 0 | 70 | 3.8e-05 | 1 | DFNVGGYIQAVLDR |
| ☑ | 109 | 790.98 | 1579.95 | 1579.82 | 0.12 | 0 | (61) | 0.00031 | 1 | QIIEQLSSGFFSPK |
| ☑ | 110 | 790.98 | 1579.95 | 1579.82 | 0.12 | 0 | 74 | 1.7e-05 | 1 | QIIEQLSSGFFSPK |
| ☑ | 112 | 537.32 | 1608.95 | 1608.86 | 0.09 | 0 | 45 | 0.012 | 1 | VHINPNSLFDVQVK |
| | 113 | 805.49 | 1608.97 | 1608.86 | 0.11 | 0 | (7) | 69 | 4 | VHINPNSLFDVQVK |
| ☑ | 127 | 560.32 | 1677.95 | 1677.86 | 0.09 | 1 | 55 | 0.0013 | 1 | IGEEYISDLDQLRK |
| ☑ | 128 | 840.00 | 1677.98 | 1677.86 | 0.12 | 1 | (48) | 0.0076 | 1 | IGEEYISDLDQLRK |
| | 131 | 565.67 | 1693.97 | 1693.81 | 0.16 | 1 | 11 | 35 | 2 | DFYELEPHKFQNK |
| | 134 | 570.99 | 1709.95 | 1709.86 | 0.09 | 1 | 16 | 9.3 | 3 | RIYYLSLEFYMGR |

Nominal mass ($M_r$): **97097**; Calculated pI value: **6.76**
NCBI BLAST search of P00489 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Oryctolagus cuniculus

Variable modifications: Carbamidomethyl (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 33%

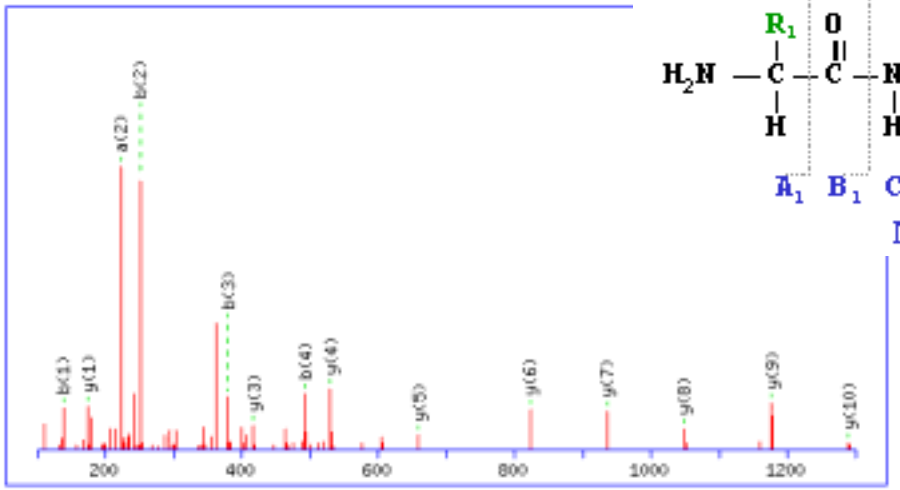
Matched peptides shown in **Bold Red**

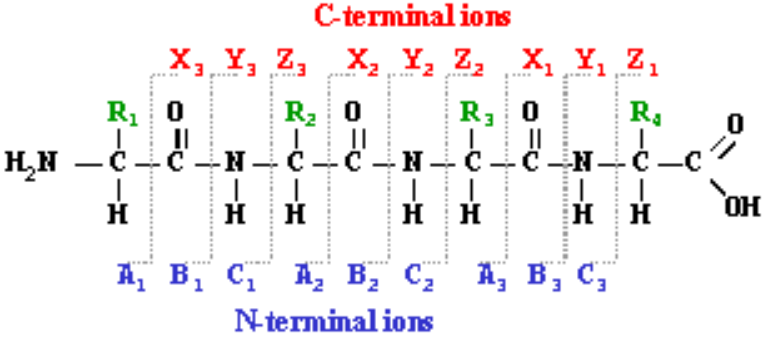
```
     1 SRPLSDQEKR KQISVRGLAG VENVTELKKN FNRHLHFTLV KDRNVATPRD
    51 YYFALAHTVR DHLVGRWIRT QQHYYEKDPK RIYYLSLEFY MGRTLQNTMV
   101 NLALENACDE ATYQLGLDME ELEEIEEDAG LGNGGLGRLA ACFLDSMATL
   151 GLAAYGYGIR YEFGIFNQKI CGGWQMEEAD DWLRYGNPWE KARPEFTLPV
   201 HFYGRVEHTS QGAKWVDTQV VLAMPYDTPV PGYRNNVVNT MRLWSAKAPN
   251 DFNLKDFNVG GYIQAVLDRN LAENISRVLY PNDNFFEGKE LRLKQEYFVV
   301 AATLQDIIRR FKSSKFGCRD PVRTNFDAFP DKVAIQLNDT HPSLAIPELM
   351 RVLVDLERLD WDKAWEVTVK TCAYTNHTVL PEALERWPVH LLETLLPRHL
   401 QIIYEINQRF LNRVAAAFPG DVDRLRRMSL VEEGAVKRIN MAHLCIAGSH
   451 AVNGVARIHS EILKKTIFKD FYELEPHKFQ NKTNGITPRR WLVLCNPGLA
   501 EIIAERIGEE YISDLDQLRK LLSYVDDEAF IRDVAKVKQE NKLKFAAYLE
   551 REYKVHINPN SLFDVQVKRI HEYKRQLLNC LHVITLYNRI KKEPNKFVVP
   601 RTVMIGGKAA PGYHMAKMII KLITAIGDVV NHDPVVGDRL RVIFLENYRV
   651 SLAEKVIPAA DLSEQISTAG TEASGTGNMK FMLNGALTIG TMDGANVEMA
   701 EEAGEENFFI FGMRVEDVDR LDQRGYNAQE YYDRIPELRQ IIEQLSSGFF
   751 SPKQPDLFKD IVNMLMHHDR FKVFADYEEY VKCQERVSAL YKNPREWTRM
   801 VIRNIATSGK FSSDRTIAQY AREIWGVEPS RQRLPAPDEK IP
```

Show predicted peptides also

Sort Peptides By (•) Residue Number ( ) Increasing Mass ( ) Decreasing Mass

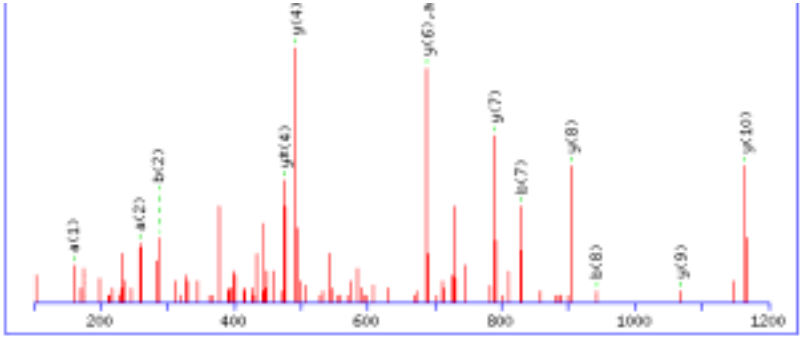
| Start - End | Observed | Mr(expt) | Mr(calc) | Delta | Miss | Sequence | |
|---|---|---|---|---|---|---|---|
| 17 - 28 | 615.40 | 1228.78 | 1228.67 | 0.11 | 0 | GLAGVENVTELK | (Ions score 82) |
| 17 - 29 | 453.30 | 1356.89 | 1356.76 | 0.13 | 1 | GLAGVENVTELKK | (Ions score 41) |
| 81 - 93 | 570.99 | 1709.95 | 1709.86 | 0.09 | 1 | RIYYLSLEFYMGR | (Ions score 16) |
| 161 - 169 | 573.32 | 1144.62 | 1144.56 | 0.06 | 0 | YEFGIFNQK | (Ions score 45) |
| 215 - 234 | 769.78 | 2306.32 | 2306.14 | 0.18 | 0 | WVDTQVVLAMPYDTPVPGYR | (Ions score 38) |
| 215 - 234 | 1154.16 | 2306.31 | 2306.14 | 0.17 | 0 | WVDTQVVLAMPYDTPVPGYR | (Ions score 25) |
| 256 - 269 | 783.98 | 1565.94 | 1565.78 | 0.16 | 0 | DFNVGGYIQAVLDR | (Ions score 70) |
| 256 - 269 | 522.98 | 1565.91 | 1565.78 | 0.12 | 0 | DFNVGGYIQAVLDR | (Ions score 63) |
| 278 - 289 | 721.89 | 1441.76 | 1441.69 | 0.08 | 0 | VLYPNDNFFEGK | (Ions score 25) |




**Monoisotopic mass of neutral peptide Mr(calc):** 1425.77
**Ions Score:** 70 **Expect:** 4.1e-05
**Matches (Bold Red):** 14/112 fragment ions using 21 most intense peaks

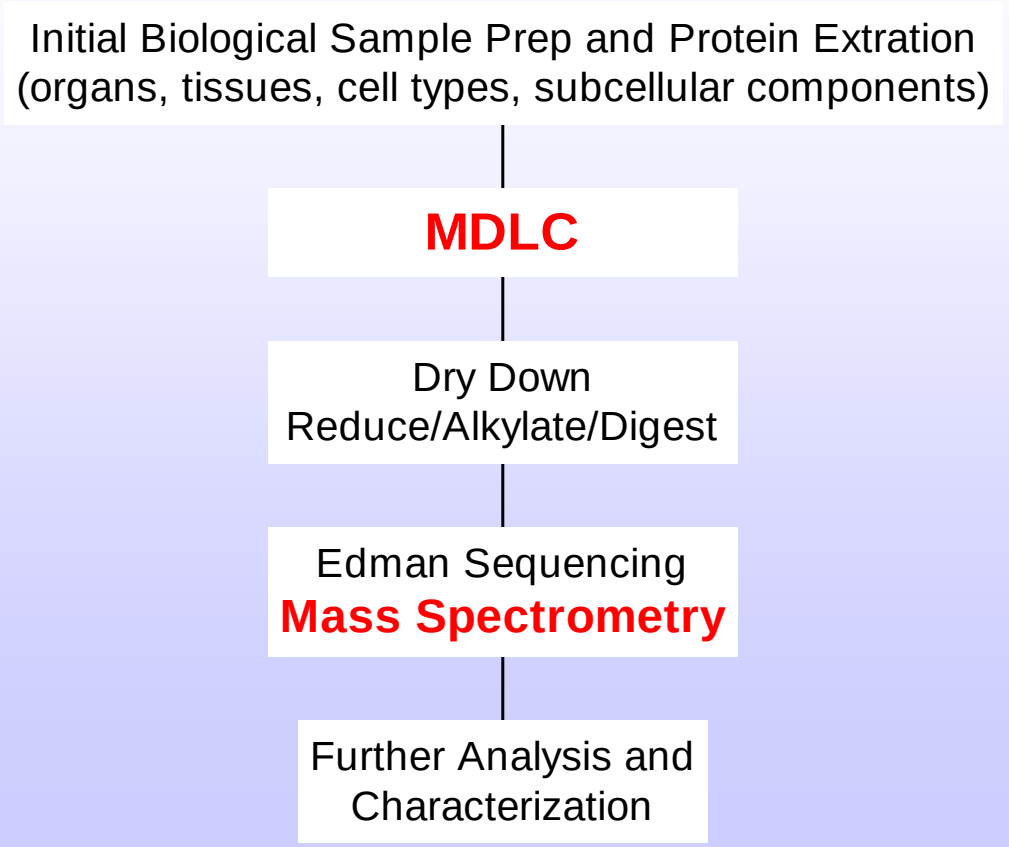
| # | a | $a^{++}$ | $a^*$ | $a^{*++}$ | b | $b^{++}$ | $b^*$ | $b^{*++}$ | Seq. | y | $y^{++}$ | $y^*$ | $y^{*++}$ | # |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 110.07 | 55.54 | | | **138.07** | 69.54 | | | H | | | | | 11 |
| 2 | **223.16** | 112.08 | | | **251.15** | 126.08 | | | L | **1289.72** | 645.36 | 1272.69 | 636.85 | 10 |
| 3 | 351.21 | 176.11 | 334.19 | 167.60 | **379.21** | 190.11 | 362.18 | 181.59 | Q | **1176.64** | 588.82 | 1159.61 | 580.31 | 9 |
| 4 | 464.30 | 232.65 | 447.27 | 224.14 | **492.29** | 246.65 | 475.27 | 238.14 | I | **1048.58** | 524.79 | 1031.55 | 516.28 | 8 |
| 5 | 577.38 | 289.19 | 560.36 | 280.68 | 605.38 | 303.19 | 588.35 | 294.68 | I | **935.49** | 468.25 | 918.47 | 459.74 | 7 |
| 6 | 740.45 | 370.73 | 723.42 | 362.21 | 768.44 | 384.72 | 751.41 | 376.21 | Y | **822.41** | 411.71 | 805.38 | 403.20 | 6 |
| 7 | 869.49 | 435.25 | 852.46 | 426.73 | 897.48 | 449.25 | 880.46 | 440.73 | E | **659.35** | 330.18 | 642.32 | 321.66 | 5 |
| 8 | 982.57 | 491.79 | 965.55 | 483.28 | 1010.57 | 505.79 | 993.54 | 497.27 | I | **530.30** | 265.66 | 513.28 | 257.14 | 4 |
| 9 | 1096.61 | 548.81 | 1079.59 | 540.30 | 1124.61 | 562.81 | 1107.58 | 554.30 | N | **417.22** | 209.11 | 400.19 | 200.60 | 3 |
| 10 | 1224.67 | 612.84 | 1207.65 | 604.33 | 1252.67 | 626.84 | 1235.64 | 618.32 | Q | 303.18 | 152.09 | 286.15 | 143.58 | 2 |



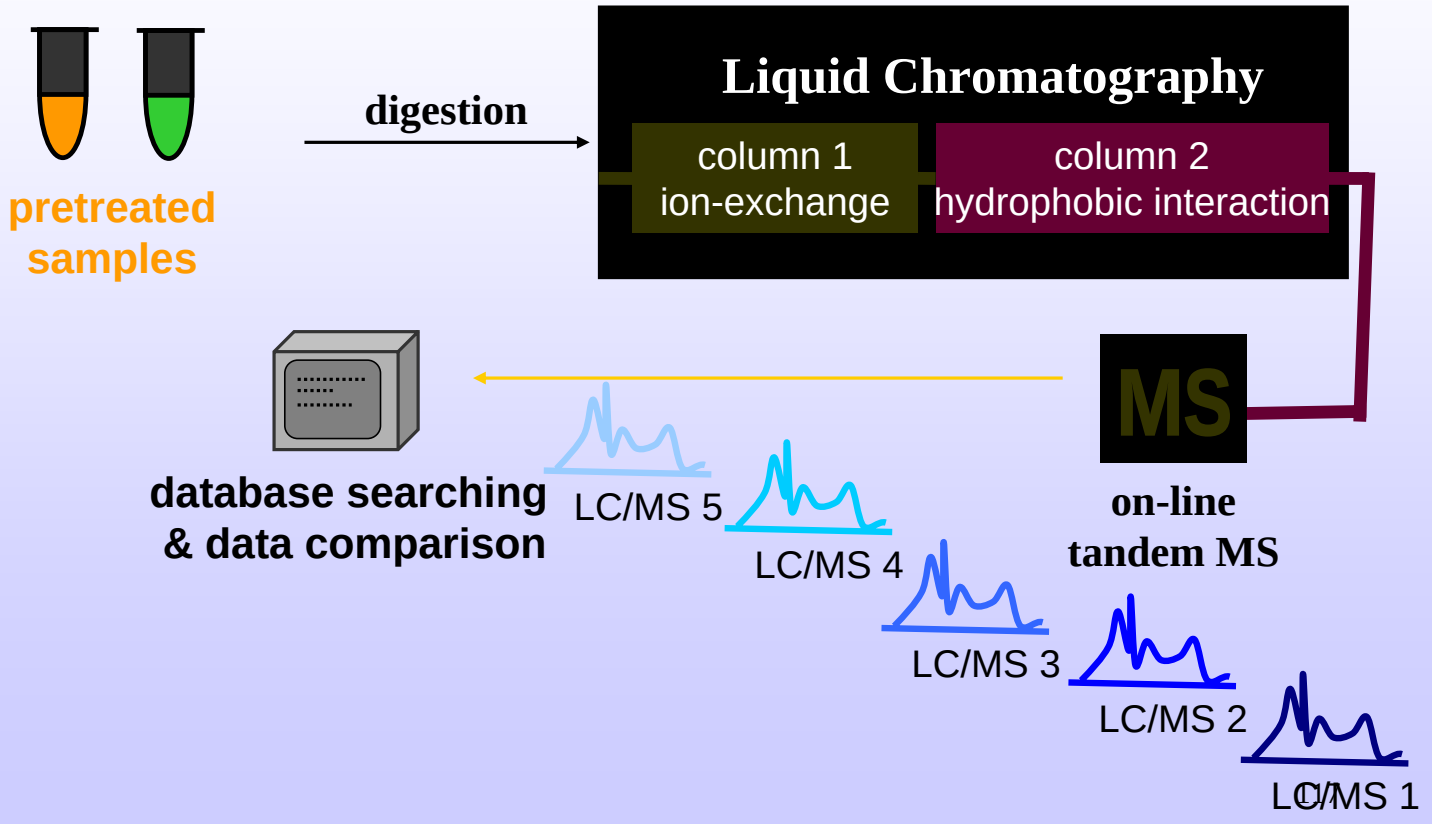

**Monoisotopic mass of neutral peptide Mr(calc):** 2306.14
**Ions Score:** 38 **Expect:** 0.041
**Matches (Bold Red):** 13/212 fragment ions using 22 most intense peaks

| # | a | $a^{++}$ | a* | $a^{*++}$ | b | $b^{++}$ | b* | $b^{*++}$ | Seq. | y | $y^{++}$ | y* | $y^{*++}$ | # |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **159.09** | 80.05 | | | 187.09 | 94.05 | | | W | | | | | 20 |
| 2 | **258.16** | 129.58 | | | **286.15** | 143.58 | | | V | 2121.07 | 1061.04 | 2104.04 | 1052.52 | 19 |
| 3 | 373.19 | 187.10 | | | 401.18 | 201.09 | | | D | 2022.00 | 1011.50 | 2004.97 | 1002.99 | 18 |
| 4 | 474.23 | 237.62 | | | 502.23 | 251.62 | | | T | 1906.97 | 953.99 | 1889.95 | 945.48 | 17 |
| 5 | 602.29 | 301.65 | 585.27 | 293.14 | 630.29 | 315.65 | 613.26 | 307.13 | Q | 1805.93 | 903.47 | 1788.90 | 894.95 | 16 |
| 6 | 701.36 | 351.18 | 684.34 | 342.67 | 729.36 | 365.18 | 712.33 | 356.67 | V | 1677.87 | 839.44 | 1660.84 | 830.92 | 15 |
| 7 | 800.43 | 400.72 | 783.40 | 392.21 | **828.43** | 414.72 | 811.40 | 406.20 | V | 1578.80 | 789.90 | 1561.77 | 781.39 | 14 |
| 8 | 913.51 | 457.26 | 896.49 | 448.75 | **941.51** | 471.26 | 924.48 | 462.74 | L | 1479.73 | 740.37 | 1462.70 | 731.86 | 13 |
| 9 | 984.55 | 492.78 | 967.52 | 484.27 | 1012.55 | 506.78 | 995.52 | 498.26 | A | 1366.65 | 683.83 | 1349.62 | 675.31 | 12 |
| 10 | 1115.59 | 558.30 | 1098.57 | 549.79 | 1143.59 | 572.30 | 1126.56 | 563.78 | M | 1295.61 | 648.31 | 1278.58 | 639.79 | 11 |
| 11 | 1212.64 | 606.83 | 1195.62 | 598.31 | 1240.64 | 620.82 | 1223.61 | 612.31 | P | **1164.57** | 582.79 | 1147.54 | 574.27 | 10 |
| 12 | 1375.71 | **688.36** | 1358.68 | 679.84 | 1403.70 | 702.36 | 1386.68 | 693.84 | Y | **1067.52** | 534.26 | 1050.49 | 525.75 | 9 |
| 13 | 1490.73 | 745.87 | 1473.71 | 737.36 | 1518.73 | 759.87 | 1501.70 | 751.36 | D | **904.45** | 452.73 | 887.43 | 444.22 | 8 |
| 14 | 1591.78 | 796.39 | 1574.76 | 787.88 | 1619.78 | 810.39 | 1602.75 | 801.88 | T | **789.43** | 395.22 | 772.40 | 386.70 | 7 |
| 15 | 1688.84 | 844.92 | 1671.81 | 836.41 | 1716.83 | 858.92 | 1699.80 | 850.41 | P | **688.38** | 344.69 | 671.35 | 336.18 | 6 |
| 16 | 1787.90 | 894.46 | 1770.88 | 885.94 | 1815.90 | 908.45 | 1798.87 | 899.94 | V | 591.32 | 296.17 | 574.30 | 287.65 | 5 |
| 17 | 1884.96 | 942.98 | 1867.93 | 934.47 | 1912.95 | 956.98 | 1895.92 | 948.47 | P | **492.26** | 246.63 | **475.23** | 238.12 | 4 |
| 18 | 1941.98 | 971.49 | 1924.95 | 962.98 | 1969.97 | 985.49 | 1952.95 | 976.98 | G | 395.20 | 198.11 | 378.18 | 189.59 | 3 |

# **M**ulti-**d**imensiona **L**iquid **C**hromatography(MDLC) Coupled Directly to MS

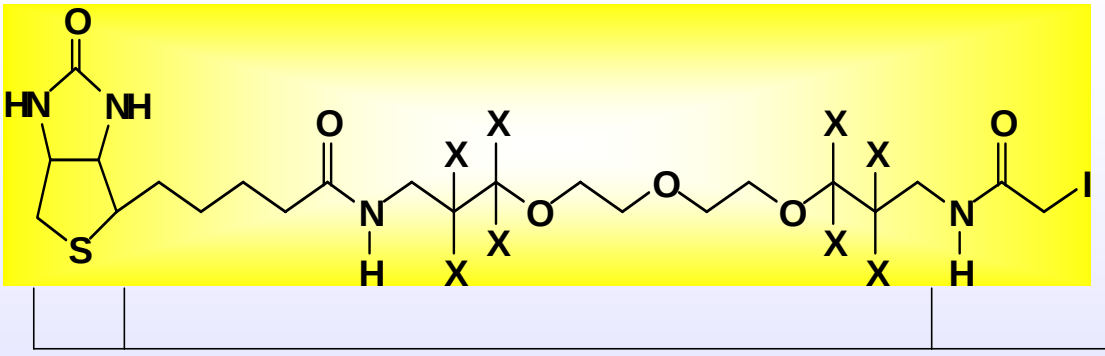


# MudPIT (Multidimentional Protein Identification Dechnology )



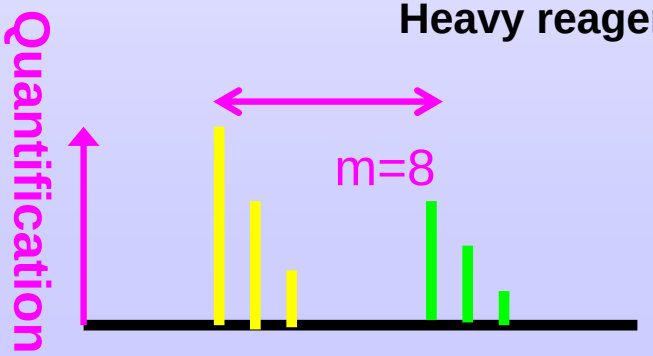
# Isotope Coded Affinity Tag

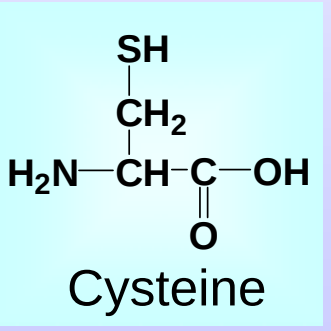
*Nature biotechnology* (1999)17:994-999



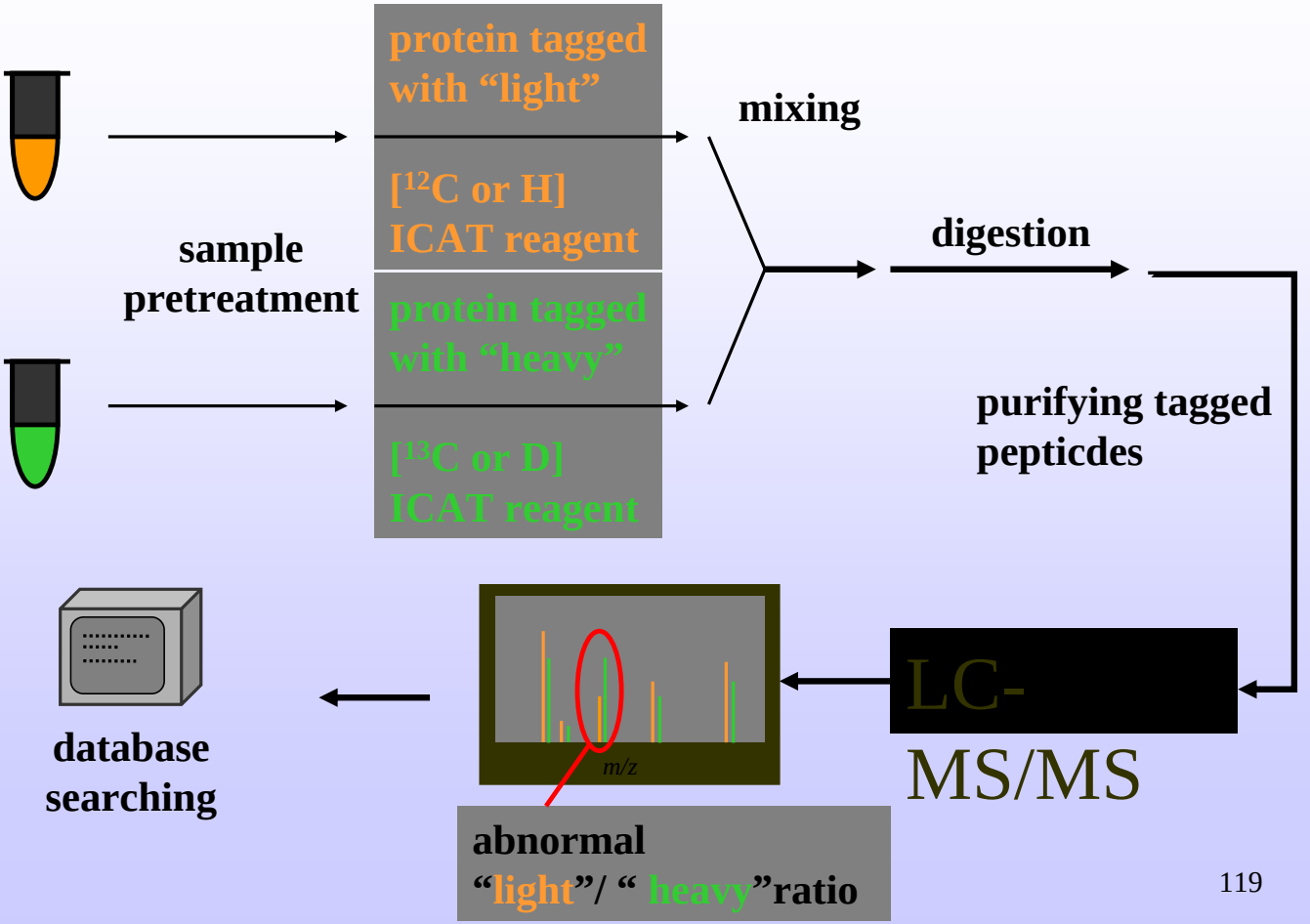
**Biotin** **Linker (light or heavy)** **Thiol-specific reactive group**

**Light reagent** : D0-ICAT (X=H)
**Heavy reagent** : D8-ICAT (X=D)



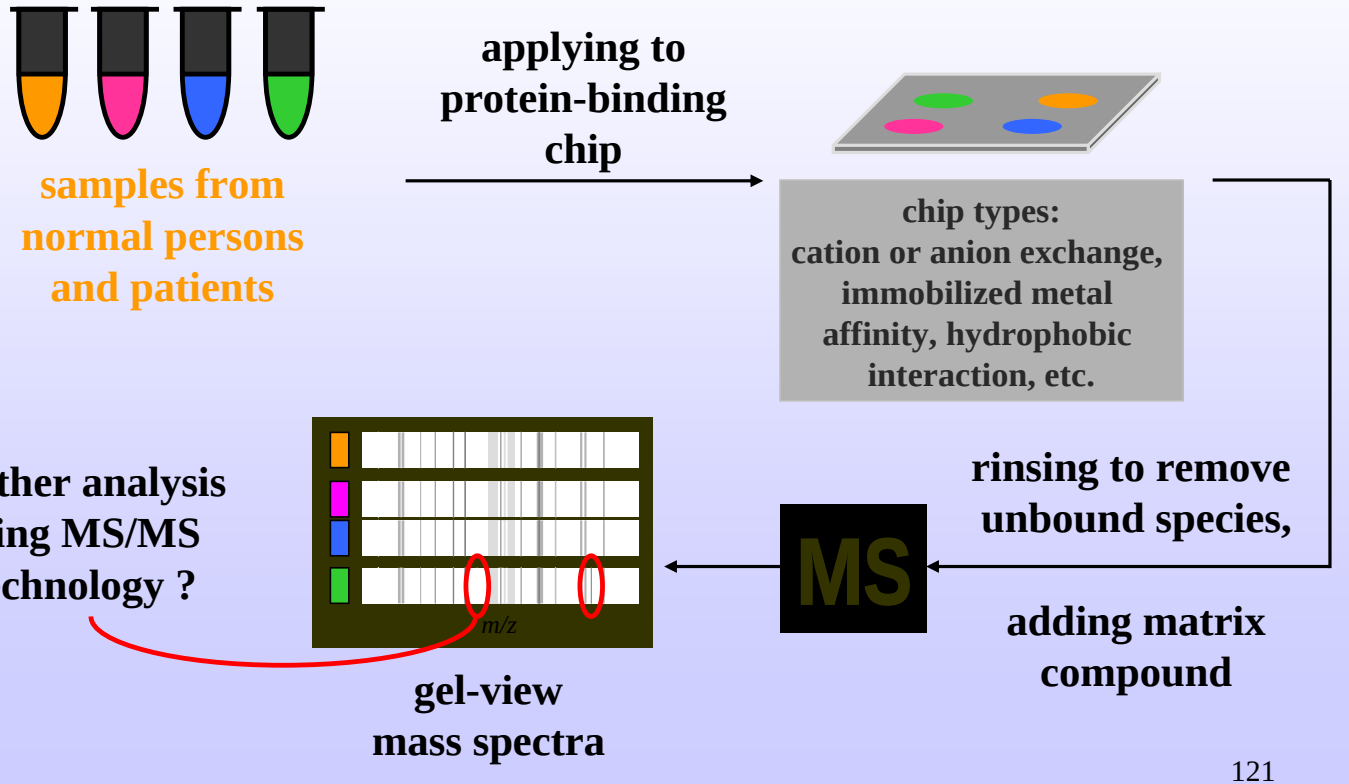
# Isotope Coded Affinity Tag (ICAT) strategy



# (四)　應用實例

# Proteomic pattern diagnostics

surface-enhanced laser desorption / ionization-time-of-flight, SELDI-TOF MS



# Disease Proteomics

## 一滴血可以診斷出疾病嗎?

**Use of proteomic patterns in serum to identify ovarian cancer**

*Emanuel F Petricoin III, Ali M Ardekani, Ben A Hitt, Peter J Levine, Vincent A Fusaro, Seth M Steinberg, Gordon B Mills, Charles Simone, David A Fishman, Elise C Kohn, Lance A Liotta*

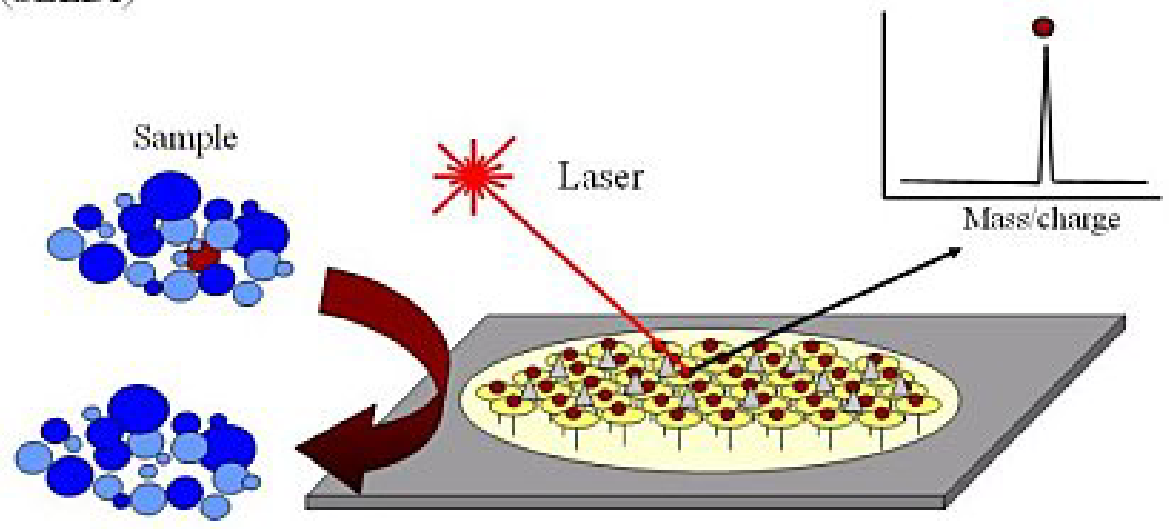
**Background** New technologies for the detection of early-stage ovarian cancer are urgently needed. Pathological changes within an organ might be reflected in proteomic patterns in serum. We developed a bioinformatics tool and used it to identify proteomic patterns in serum that distinguish neoplastic from non-neoplastic disease within the ovary.

572 THE LANCET · Vol 359 · February 16, 2002 · www.thelancet.com

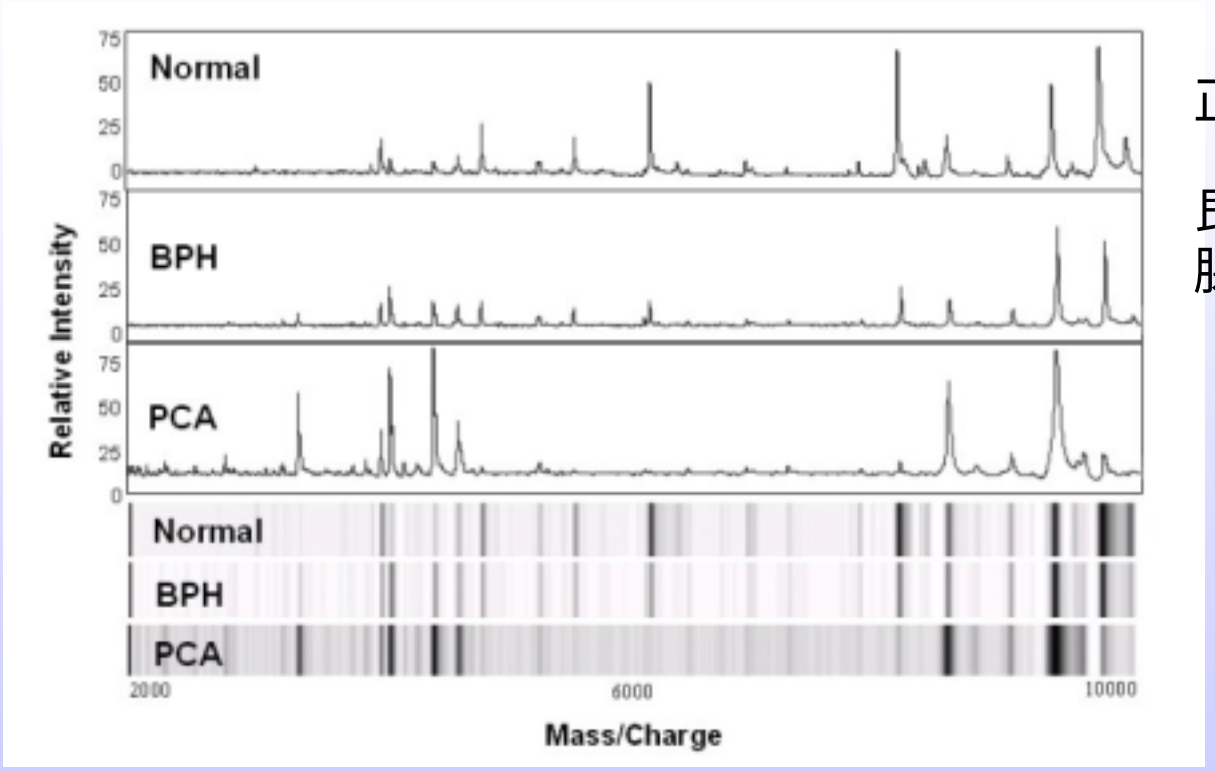
# *Protein Chips* 蛋白質晶片

## SELDI ProteinChip™ Technology Process

- Sample goes *directly* onto the ProteinChip™ Array
- Proteins● are captured and *retained* on the chip (affinity capture )
- EAM△ is added to the chip
- Retentate map is "read" by Surface-Enhanced Laser Desorption/Ionization (SELDI)



# 找尋腫瘤標記蛋白質




正常人

良性攝護腺腫瘤

攝護腺癌

# Imaging Mass Spectrometry

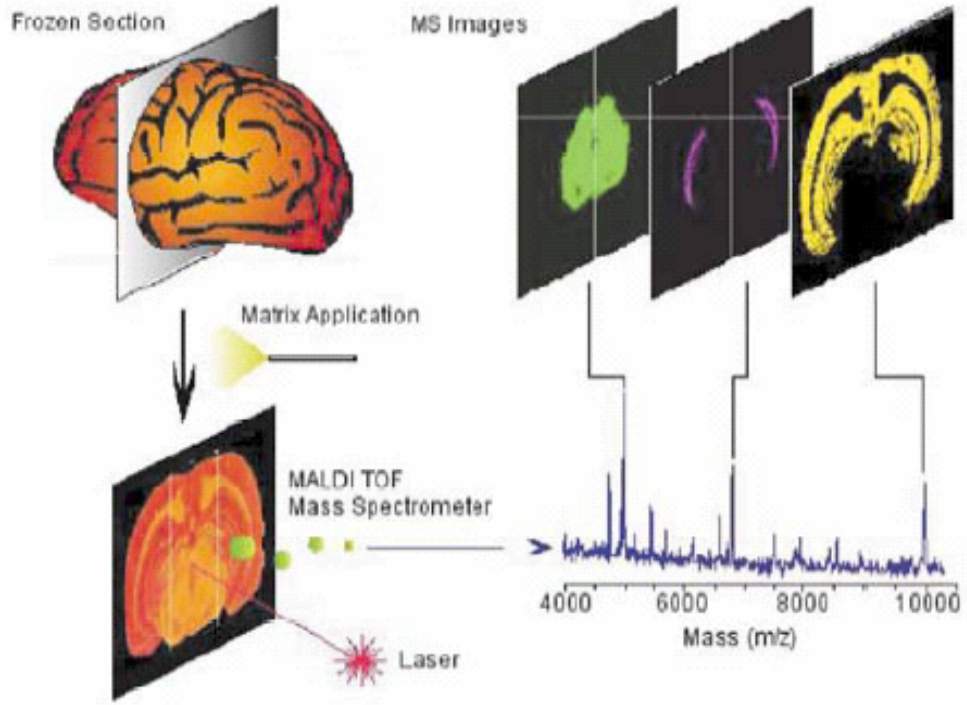



**Fig. 1** Methodology developed for the spatial analysis of tissue by MALDI mass spectrometry. Frozen sections are mounted on a metal plate, coated with an UV-absorbing matrix and placed in the mass spectrometer. A pulsed UV laser desorbs and ionizes analytes from the tissue and their m/z values are determined using a time-of-flight analyzer. From a raster over the tissue and measurement of the peak intensities over thousands of spots, mass spectrometric images are generated at specific molecular weight values.

Nat. Med. **2001**, *7*, 493-496

# MALDI Imaging Strategies

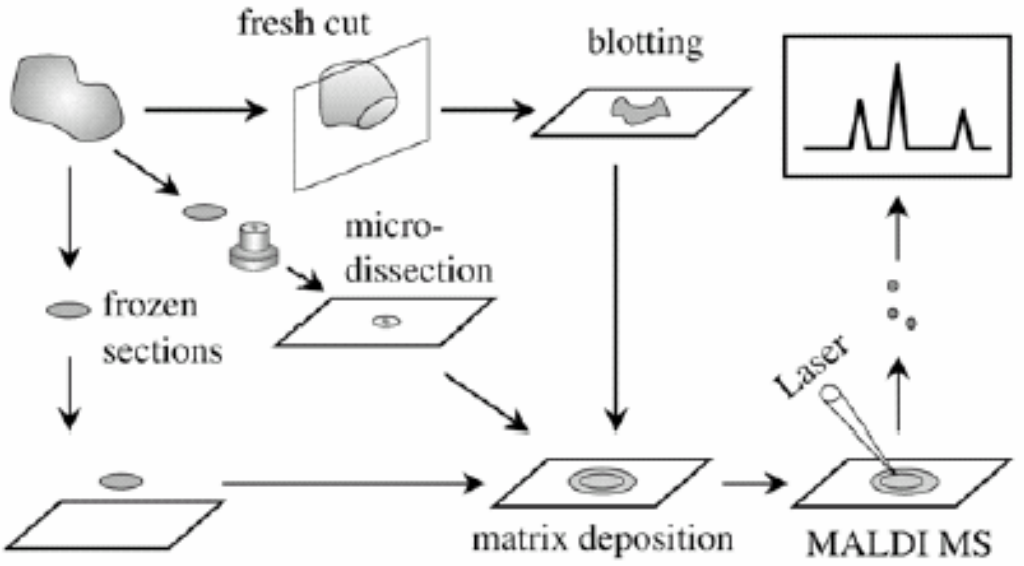
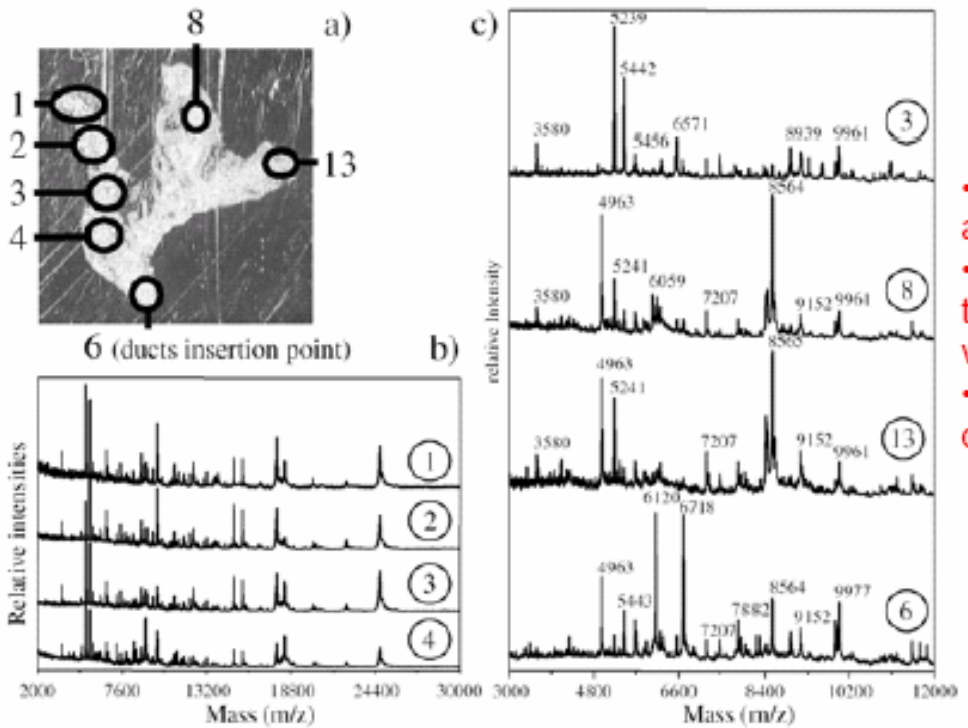



**Figure 6.** Schematics of the different strategies and sample preparation methods employed to investigate protein populations in tissue samples.

1. Membrane blotting of fresh tissue section, transferring some of the proteins present
2. Direct analysis of a frozen and dried tissue section by coating it with MALDI matrix
3. Laser capture microdissection followed by MALDI MS of selected cell

*JMS* **2001**, *36*, 355-369

# MALDI MS Analysis of Prostate Branches




- Matrix deposited using a narrow capillary
- Within the duct (1-4) the protein profiles are very similar
- Variability seen in other regions

**Figure 10.** MALDI-MS analysis of a 12 μm tissue section obtained from the mouse anterior prostate: (a) optical image of the section. The drops of matrix are circled. (b) Comparison of the protein profiles obtained in the left branch of the section. (c) Comparison of the protein profiles obtained in all three major branches and also near the point of embranchment.

*JMS* **2001**, *36*, 355-369

# Functional organization of the yeast proteome by systematic analysis of protein complexes

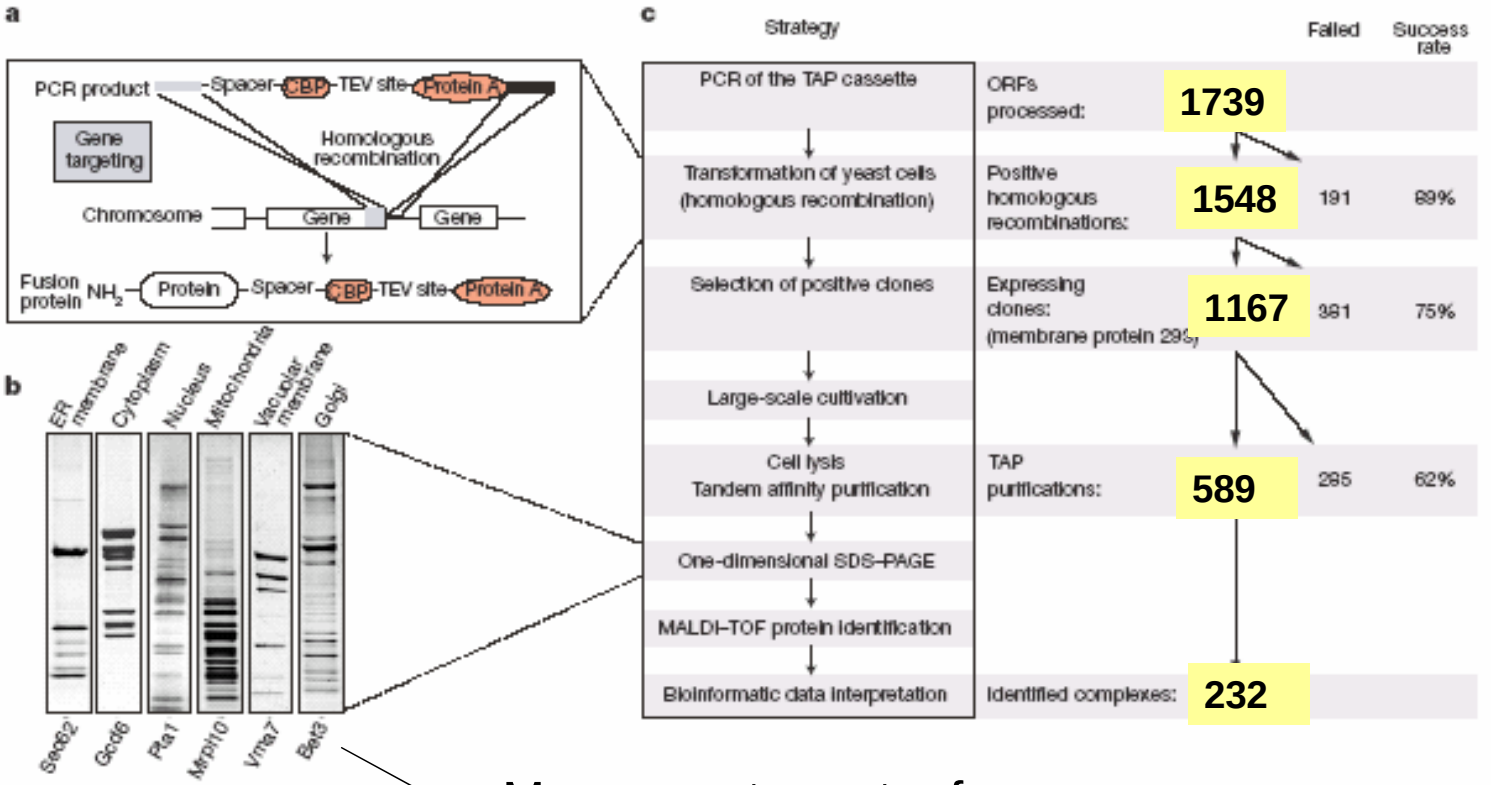
Anne-Claude Gavin*, Markus Bösche*, Roland Krause*, Paola Grandi*, Martina Marzioch*, Andreas Bauer*, Jörg Schultz*, Jens M. Rick*, Anne-Marie Michon*, Cristina-Maria Cruciat*, Marita Remor*, Christian Höfert*, Malgorzata Schelder*, Miro Brajenovic*, Heinz Ruffner*, Alejandro Merino*, Karin Klein*, Manuela Hudak*, David Dickson*, Tatjana Rudi*, Volker Gnau*, Angela Bauch*, Sonja Bastuck*, Bettina Huhse*, Christina Leutwein*, Marie-Anne Heurtier*, Richard R. Copley†, Angela Edelmann*, Erich Querfurth*, Vladimir Rybin*, Gerard Drewes*, Manfred Raida*, Tewis Bouwmeester*, Peer Bork†, Bertrand Seraphin†‡, Bernhard Kuster*, Gitte Neubauer* & Giulio Superti-Furga*†

* *Cellzome AG, Meyerhofstrasse 1, 69117 Heidelberg, Germany*
† *European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany*
‡ *CGM-CNRS, 91198 Gif sur Yvette Cedex, France*

Most cellular processes are carried out by multiprotein complexes. The identification and analysis of their components provides insight into how the ensemble of expressed proteins (proteome) is organized into functional units. We used tandem-affinity purification (TAP) and mass spectrometry in a large-scale approach to characterize multiprotein complexes in *Saccharomyces cerevisiae*. We processed 1,739 genes, including 1,143 human orthologues of relevance to human biology, and purified 589 protein assemblies. Bioinformatic analysis of these assemblies defined 232 distinct multiprotein complexes and proposed new cellular roles for 344 proteins, including 231 proteins with no previous functional annotation. Comparison of yeast and human complexes showed that conservation across species extends from single proteins to their molecular environment. Our analysis provides an outline of the eukaryotic proteome as a network of protein complexes at a level of organization beyond binary interactions. This higher-order map contains fundamental biological information and offers the context for a more reasoned and informed approach to drug discovery.

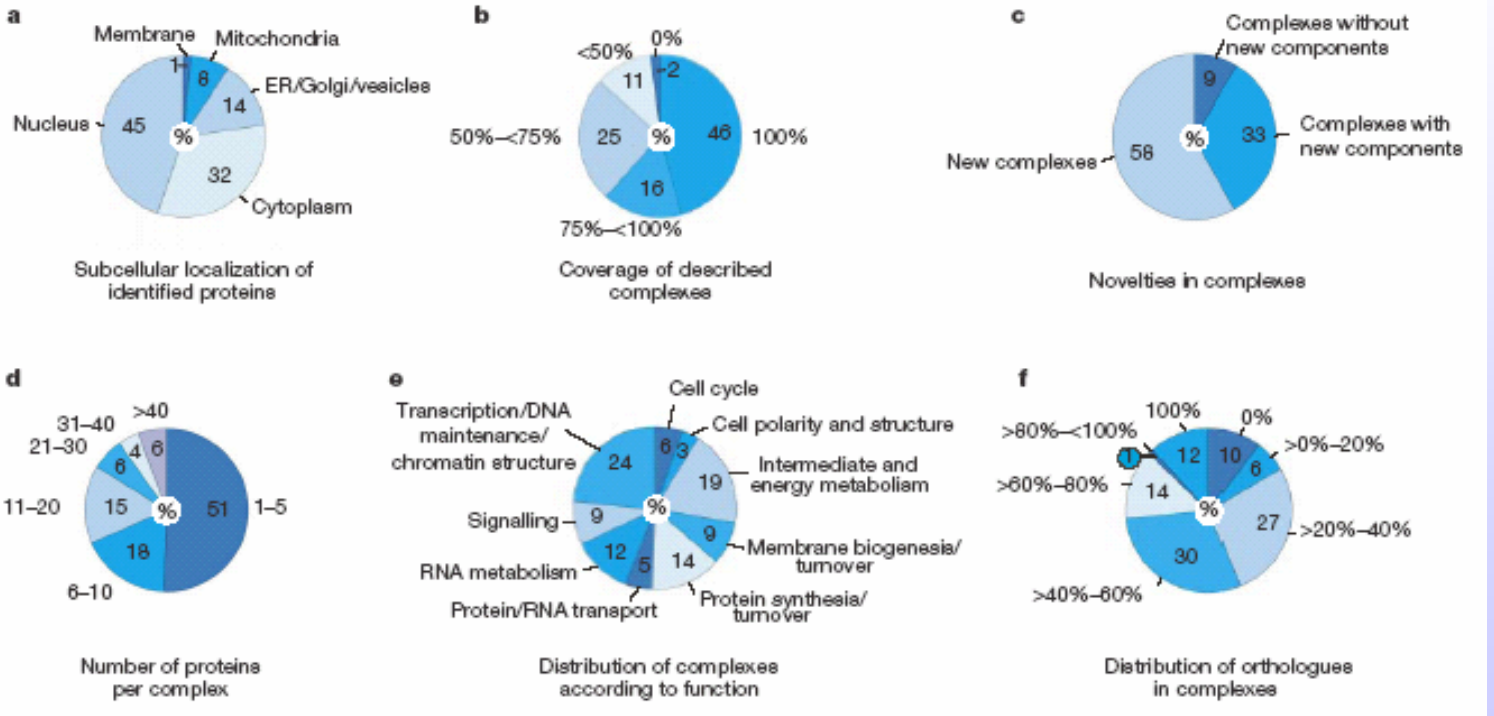
NATURE | VOL 415 | 10 JANUARY 2002

# Tandem-affinity purification (TAP)

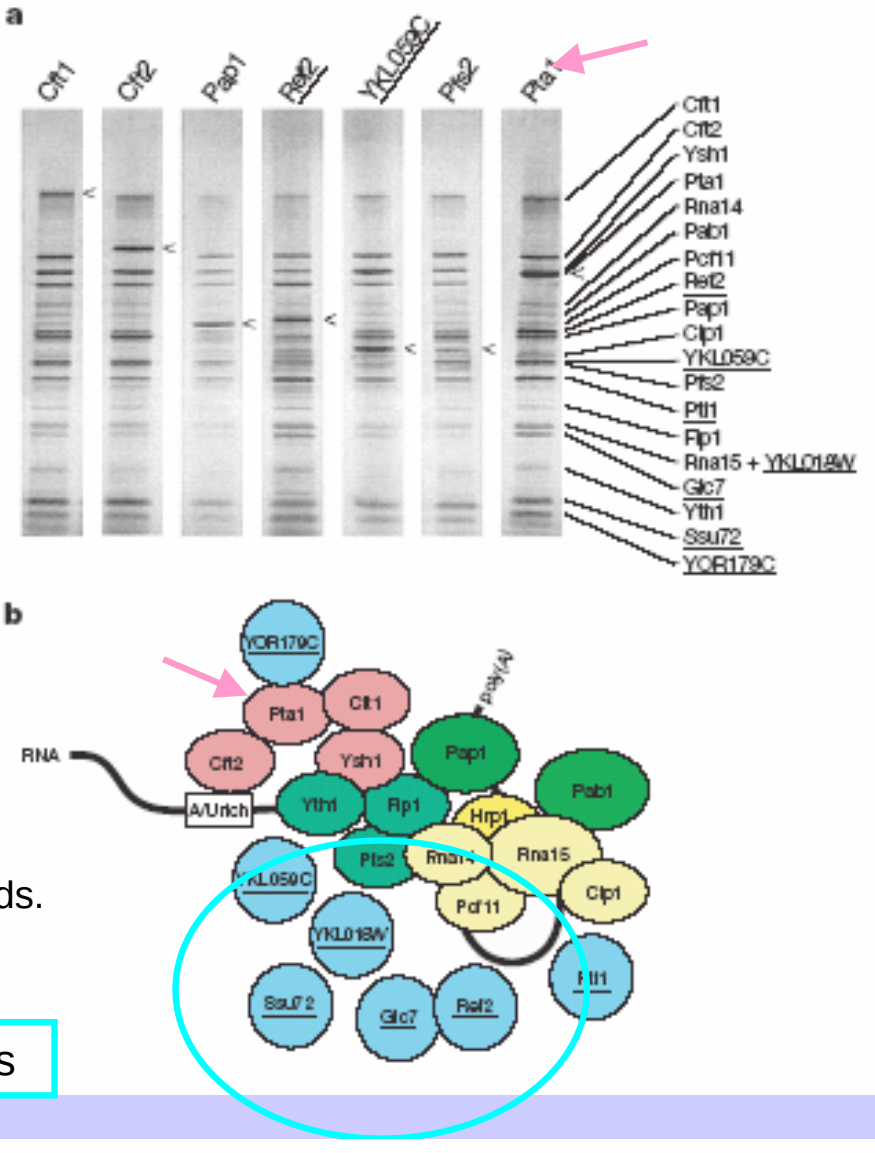



Mass spectrometry for component identification

# Statistics of identified proteins and complexes

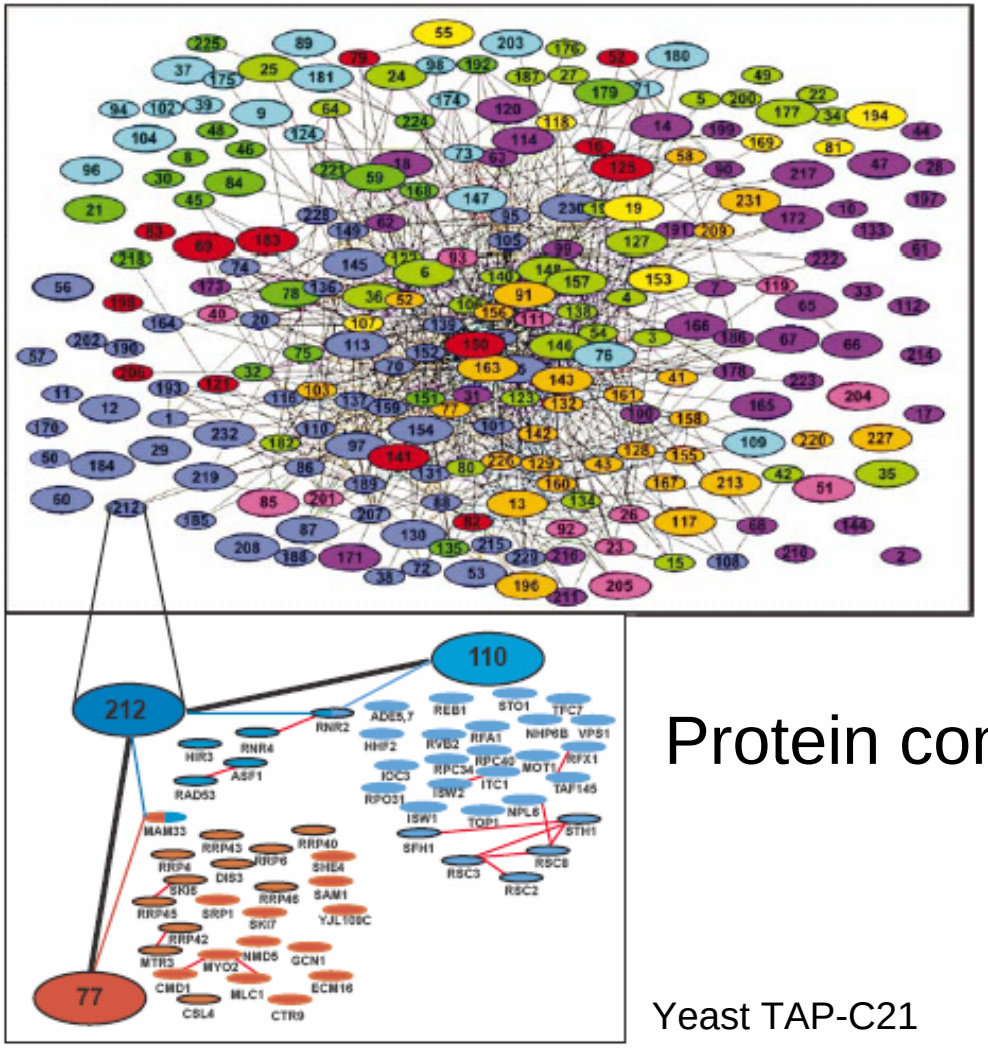


# Primary validation of complex composition by `reverse' purifcation



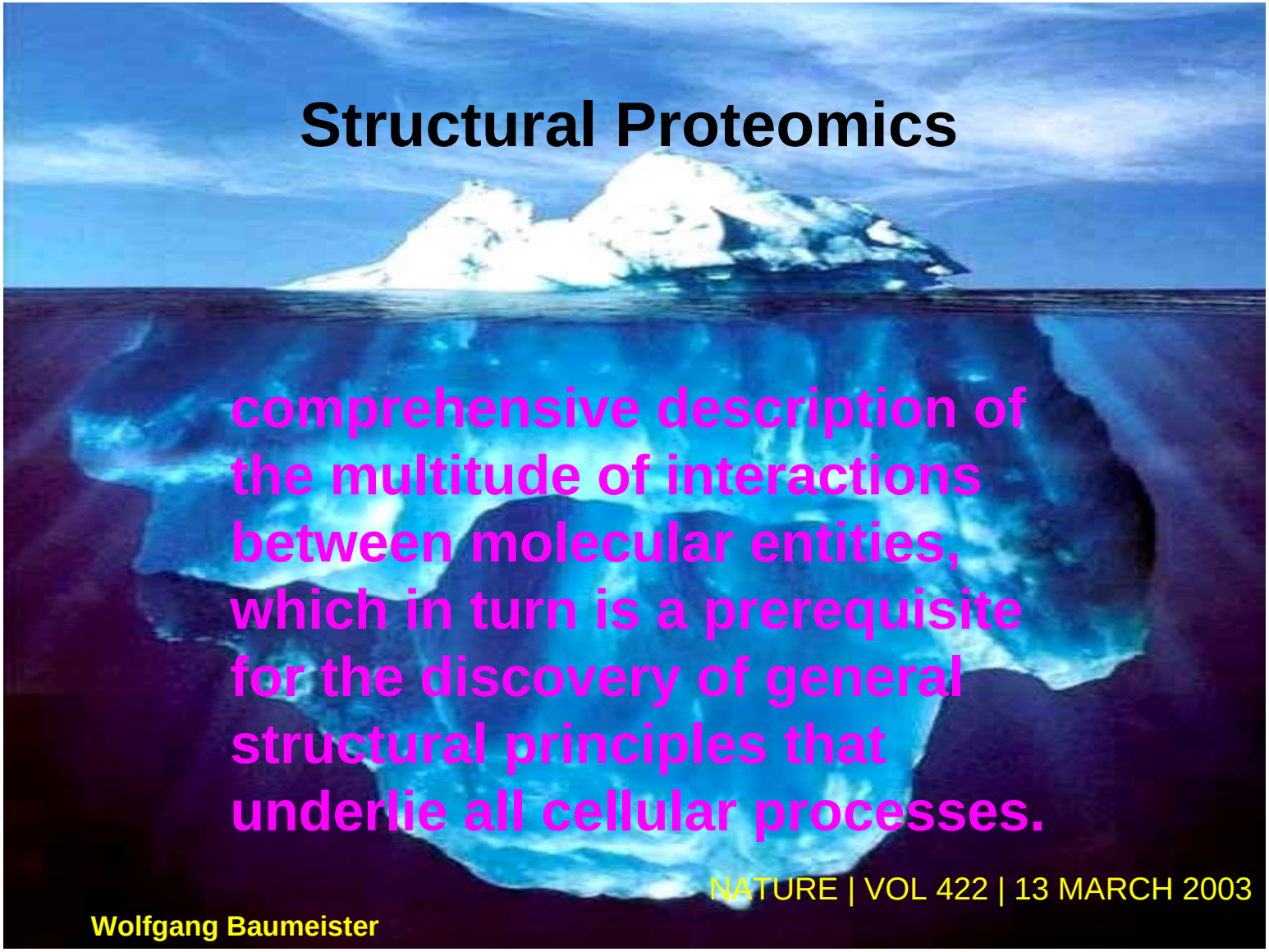

The bands of the tagged proteins are indicated by arrowheads.

New components




# Protein complex network

Yeast TAP-C21


Structural Proteomics
comprehensive description of the multitude of interactions between molecular entities, which in turn is a prerequisite for the discovery of general structural principles that underlie all cellular processes.
NATURE | VOL 422 | 13 MARCH 2003
Wolfgang Baumeister