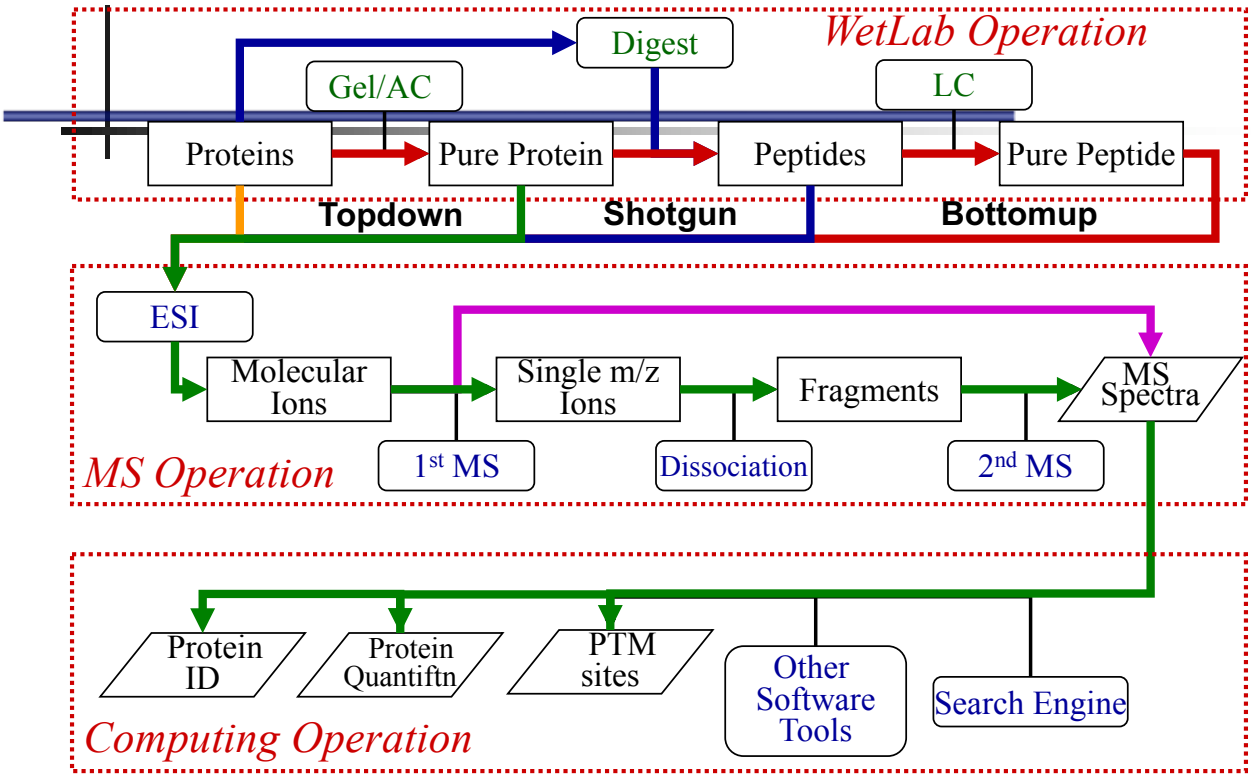
# Mass Spectrometry

Dr. S. Sadeghi



## Proteomics Approaches




Taken from: http://ms-facility.ucsf.edu/documents/PC235_2009_Lec1_MS_Intro.ppt

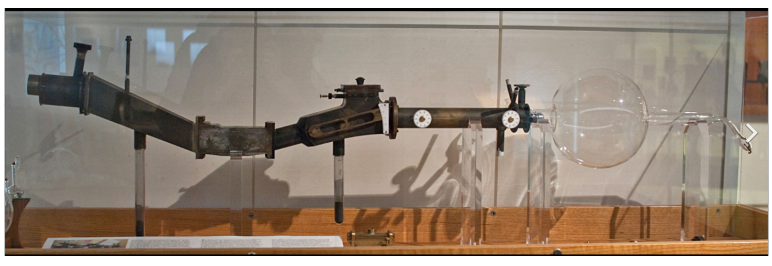
Mass Spectrometry is a technique for the detection and resolution of a sample of ions by their mass-to-charge ratio - represented by *m/z* where m is the mass in Daltons and z is the charge.



# History of MS

Sir Joseph John Thomson

- Mass Spectrometry is generally recognized to have been started with the work of Sir Joseph John Thomson.
- His work on conduction of electricity through ionized gasses lead to the Nobel Prize for Physics in 1906.
- Thomson's best known work in mass spectrometry was in demonstrating the presence of atomic isotopes of gasses - atoms of the same element with differing masses.



# Mass Definitions

Molecular masses are measured in Daltons (Da) or mass units (u).

One Dalton = 1/12 of the mass of a $^{12}C$ atom.

**Monoisotopic mass** = sum of the exact masses of the most abundant isotope of each element present, i.e. $^{1}H$=1.007825, $^{12}C$=12.000000, $^{16}O$=15.994915.

This is the most accurately defined molecular mass.

**Average mass** = sum of the abundant averaged masses ("atomic weights") of the constituent atoms of a given molecule.

The result is a weighted average over all of the naturally occurring isotopes present in the compound. This is the common chemical molecular weight that is used for stoichiometric calculations (H=1.0080, C=12.011, O=15.994). The average mass cannot be determined as accurately as the monoisotopic mass because of variations in natural isotopic abundances.

**The mass to charge ratio (m/z).** A quantity formed by dividing the mass (in u) of an ion by its charge number; unit: Thomson or Th.



# Isotopic Abundances of Common Elements

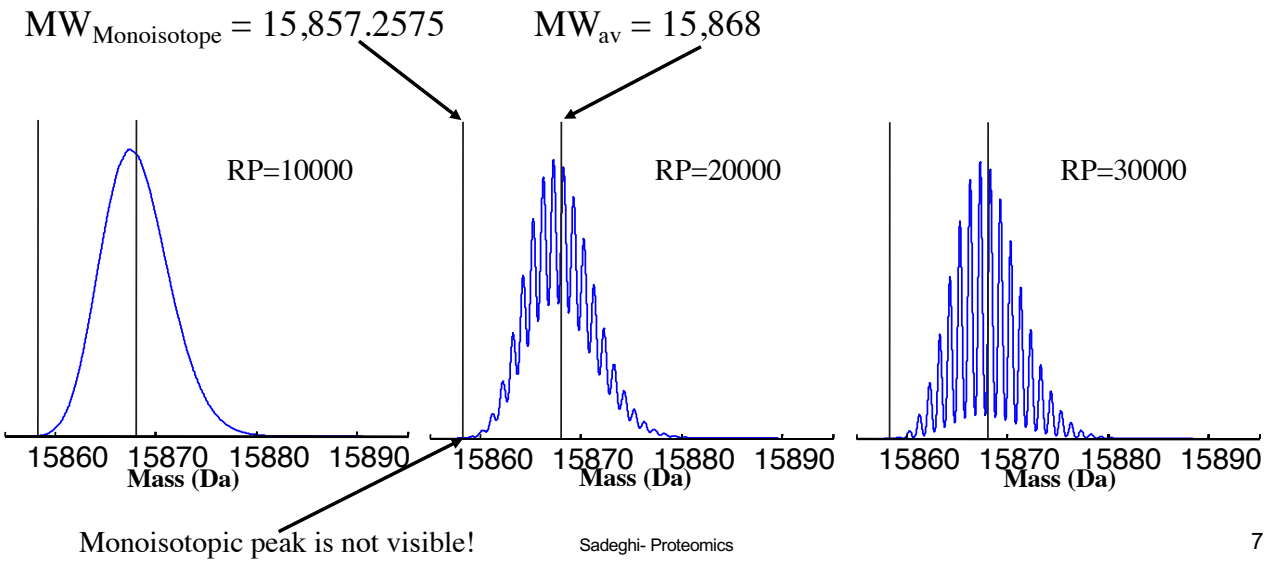
| Element | Mass | Natural Abundance |
|---|---|---|
| H | 1.0078<br>2.0141 | 99.985%<br>0.015 |
| C | 12.0000<br>13.0034 | 98.89<br>1.11 |
| N | 14.0031<br>15.0001 | 99.64<br>0.36 |
| O | 15.9949<br>16.9991<br>17.9992 | 99.76<br>0.04<br>0.20 |
| P | 30.9737 | 100 |
| S | 31.9721<br>32.9715<br>33.9679<br>35.9671 | 95.00<br>0.76<br>4.22<br>0.02 |

By coincidence, the most abundant isotope of common elements has the lowest mass.

# Protein Mass Measurement

Protein masses are normally reported as average masses

Effect of different resolving power on Hemoglobin beta chain peak, $C_{724}H_{1119}N_{195}O_{201}S_3$






# Data from MS

- Proteins in mixtures
- Quantitative analysis of protein expression
- Post-translational modification
  - Phosphorylation
- Protein interactions

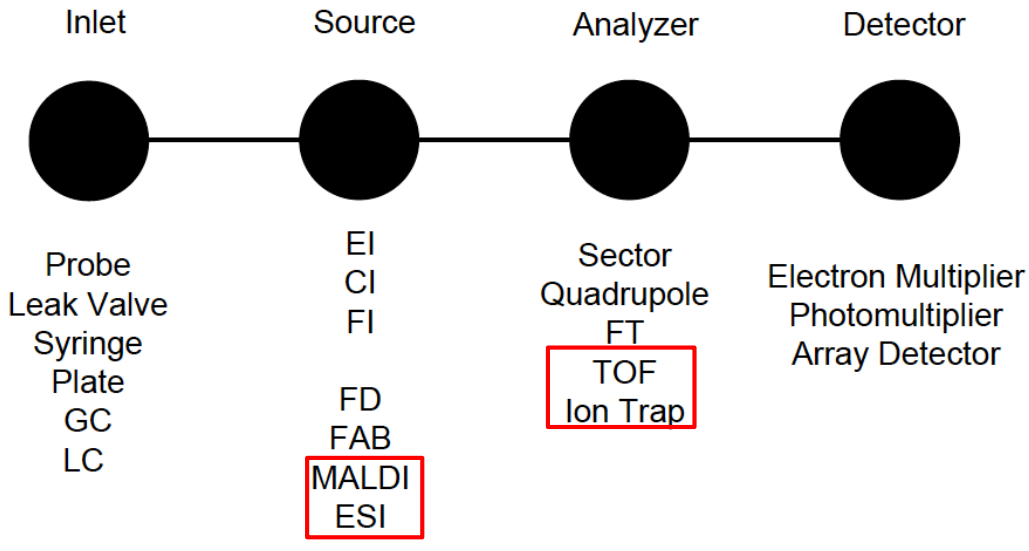
Common post-translational modifications detected by mass spectrometry.

| PTM | Residues | Chemical group | Δ mass (Da)[a] |
|---|---|---|---|
| **Phosphorylation** | Ser, Thr, Tyr | $HPO_3$ | 79.9663 |
| **N-Glycosylation** | Asn | Glycan | ≥ 132.0432[b] (−0.9840 and 2.9890)[c] |
| **O-Glycosylation** | Ser, Thr | Glycan | ≥132.0432[bc] |
| **Oxidation** | Met | O | 15.9949 |
| **Methylation** | N- C- terminus, Lys, Ser, Thr, Asn, Gln, (Iso)Asp[d] | $CH_2$ | 14.0156 |
| **Dimethylation** | Arg, Lys | $CH_2CH_2$ | 28.0313 |
| **Trimethylation** | Arg, Lys | $CH_2CH_2CH_2$ | 42.0470 |
| ***S*-Nitrosylation** | Cys | NO | 28.9902 |
| **Citrullination** | Arg | O | 0.9840 |
| **Ubiquitination** | Lys | Ubiquitin | ≥ 8564.8448 (114.0429)[e] |
| **Acetylation** | N terminus, Lys, Ser | $CH_3$ CO | 42.0106 |
| **Carbamylation** | N-terminus, Lys, Arg | $CONH_2$ | 43.0058 |
| **Biotinylation (amide bond to)** | N-terminus, Lys | Biotin | 226.0776 |

[a]Δ mass (Da) is the change in mass of the peptide and amino acid in Daltons due to the addition of a PTM.



# Mass Spec Components



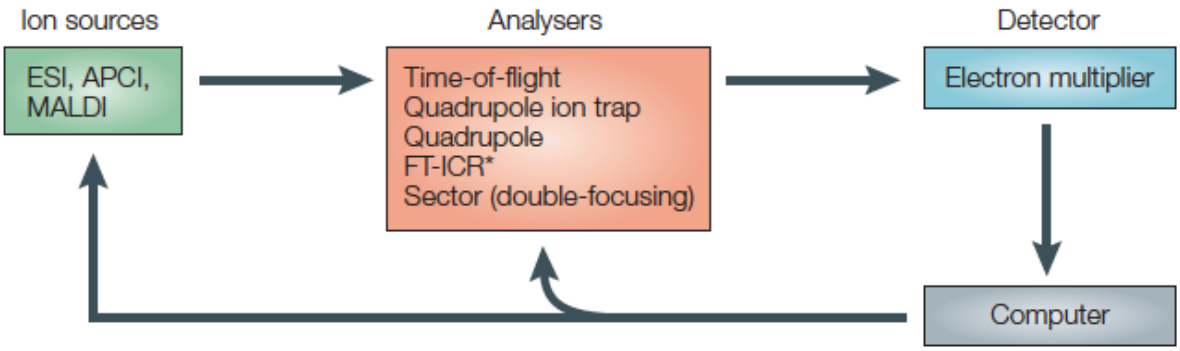


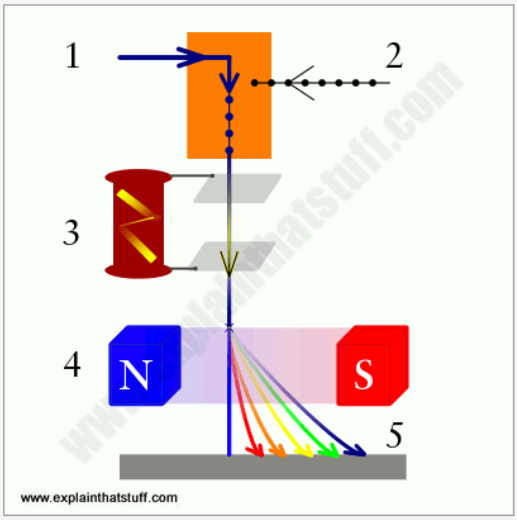

Figure 1 | **Basic components of a typical mass spectrometer used in drug discovery.** *FT-ICR does not use an electron multiplier. APCI, atmospheric-pressure chemical ionization; ESI, electrospray ionization; MALDI, matrix-assisted laser desorption/ionization; FT-ICR, Fourier-transform ion-cyclotron resonance.

Ref: Glish & Vachet. *Nature Reviews*. 2003, 2, 140.



## How does a mass spectrometer work?

There are numerous different kinds of mass spectrometers, all working in slightly different ways, but the basic process involves broadly the same stages.




1. You place the substance you want to study in a vacuum chamber inside the machine.
2. The substance is bombarded with a beam of electrons so the atoms or molecules it contains are turned into ions. This process is called **ionization**.
3. The ions shoot out from the vacuum chamber into a powerful electric field (the region that develops between two metal plates charged to high voltages), which makes them accelerate. Ions of different atoms have different amounts of electric charge, and the more highly charged ones are accelerated most, so the ions separate out according to the amount of charge they have. (This stage is a bit like the way electrons are accelerated inside an old-style, cathode-ray television.)
4. The ion beam shoots into a magnetic field (the invisible, magnetically active region between the poles of a magnet). When moving particles with an electric charge enter a magnetic field, they *bend* into an arc, with lighter particles (and more positively charged ones) bending more than heavier ones (and more negatively charged ones). The ions split into a spectrum, with each different type of ion bent a different amount according to its mass and its electrical charge.
5. A computerized, electrical detector records a spectrum pattern showing how many ions arrive for each mass/charge. This can be used to identify the atoms or molecules in the original sample. In early spectrometers, photographic detectors were used instead, producing a chart of peaked lines called a **mass spectrograph**. In modern spectrometers, you slowly vary the magnetic field so each separate ion beam hits the detector in turn.

http://www.explainthatstuff.com/how-mass-spectrometers-work.html

Measures a mass to charge ratio or $m/z$


- MALDI
- Electrospray

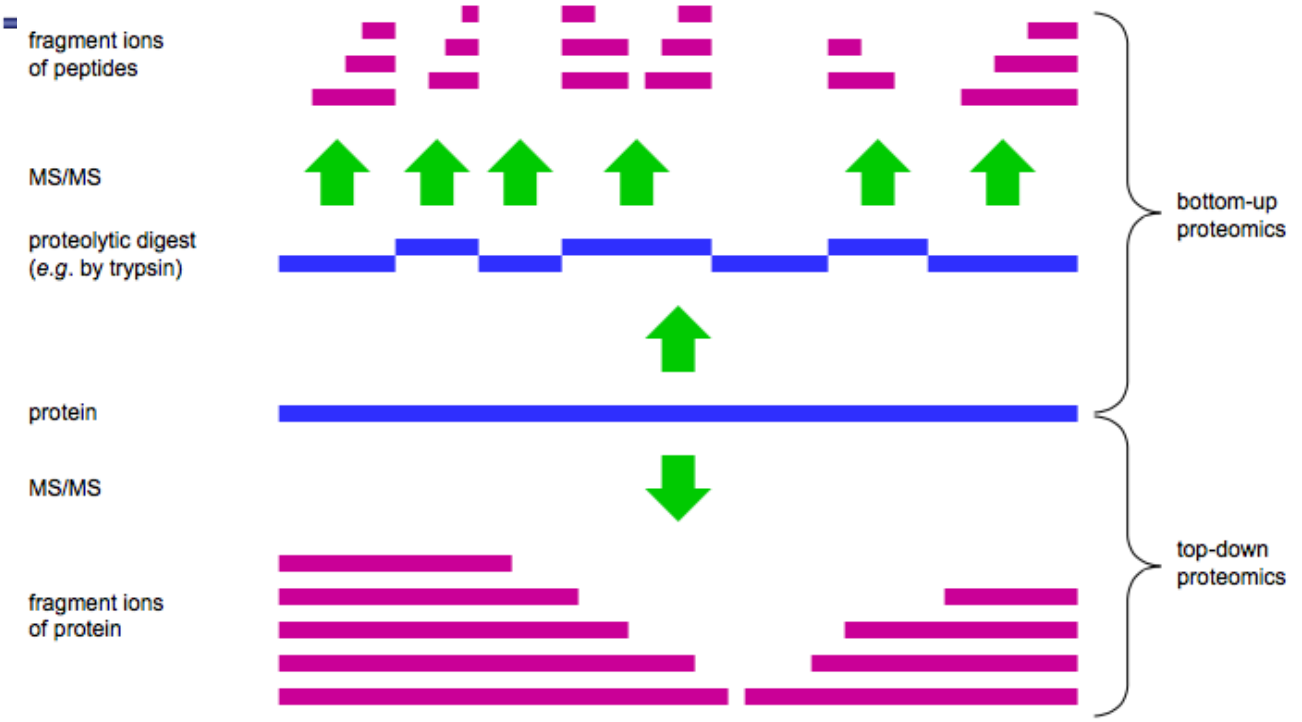
quadrupole
ion trap
time of flight

mass analysis



- Classical biochemistry techniques and 2DGE are, in general, 'top-down proteomics' – identify and quantify whole proteins.
- Most modern proteomic MS is 'bottom-up'

# Bottom-up vs. top-down



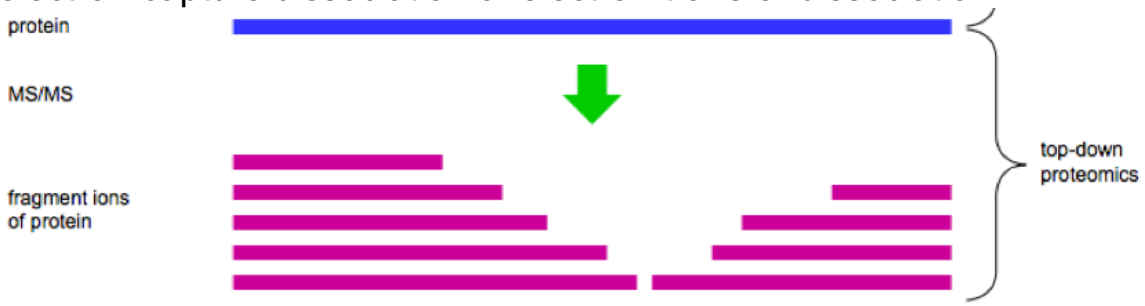



# Bottom-up proteomics

- Bottom-up proteomics is a common method to identify proteins and characterize their aa sequences and PTM by proteolytic digestion of proteins prior to analysis by mass spectrometry.
- The proteins may first be purified by a method such as gel electrophoresis resulting in one or a few proteins in each proteolytic digest. Alternatively, the crude protein extract is digested directly, followed by one or more dimensions of separation of the peptides by liquid chromatography coupled to mass spectrometry, a technique known as **shotgun proteomics**.
- By comparing the masses of the proteolytic peptides or their tandem mass spectra with those predicted from a sequence database or annotated peptide spectral in a peptide spectral library, peptides can be identified and multiple peptide identifications assembled into a protein identification

# Top-down proteomics

- Top-down proteomics is a method of protein identification that uses an ion trapping mass spectrometer to store an isolated protein ion for mass measurement and tandem mass spectrometry analysis.
- The name is derived from the similar approach to DNA sequencing.
- Proteins are typically ionized by electrospray ionization and trapped in a quadrupole ion trap mass spectrometer.
- Fragmentation for tandem mass spectrometry is accomplished by electron-capture dissociation or electron-transfer dissociation.






# Shotgun proteomics

- refers to the use of bottom-up proteomics techniques in identifying proteins in complex mixtures using a combination of HPLC combined with MS.
- The name is derived from the rapidly expanding, quasi-random firing pattern of a shotgun.
- The most common method of shotgun proteomics starts with:
  - the proteins in the mixture being digested and
  - the resulting peptides separated by liquid chromatography
  - Tandem mass spectrometry is then used to identify the peptides.

# Shotgun proteomics

- Shotgun proteomics arose from the difficulties of using previous technologies to separate complex mixtures.
- In 1975, 2D-PAGE was described by O'Farrell and Klose with the ability to resolve complex protein mixtures
- The development of matrix-assisted laser desorption ionization (MALDI), electrospray ionization (ESI), and database searching continued to grow the field of proteomics.
- Above methods difficulty identifying and separating low-abundance proteins and membrane proteins.
- Shotgun proteomics could resolve even these proteins



Mass analyzers used in shotgun proteomics and their commonly-achieved analytical metrics for peptide analysis.

| Analyzer | Instruments | Type | Resolution | Mass accuracy | Dynamic Range |
|---|---|---|---|---|---|
| **Quadrupole** | QQQ<br>QToF | Beam | 1 – 2 K | ~ 1 ‰[a] | 5 – 6 |
| **Ion trap** | LIT | Trapping (Electric field) | 1 – 2 K | ~ 1 ‰[a] | 3 – 4 |
| **ToF** | QToF | Beam | 10 – 50 K | 5 – 10 ppm | 4 |
| **Orbitrap** | FT | Trapping (Electric field) | 7.5 – 240 K | 500 ppb – 10 ppm | 4 |
| **ICR** | FT | Trapping (Magnetic and axial DC fields) | 100 – 500 K | ~100 ppb | 3 |

[a]parts-per-thousand.

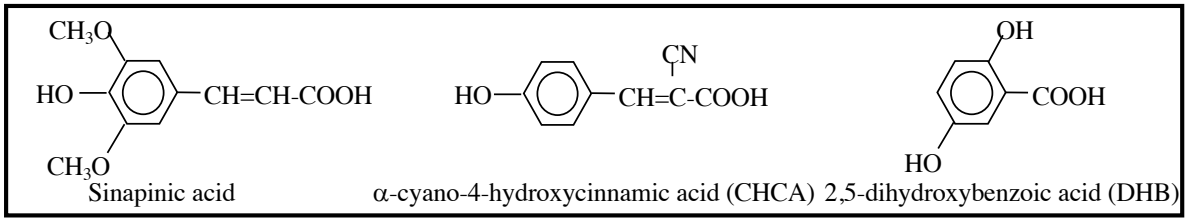
# Ion sources

## *MALDI & ESI*

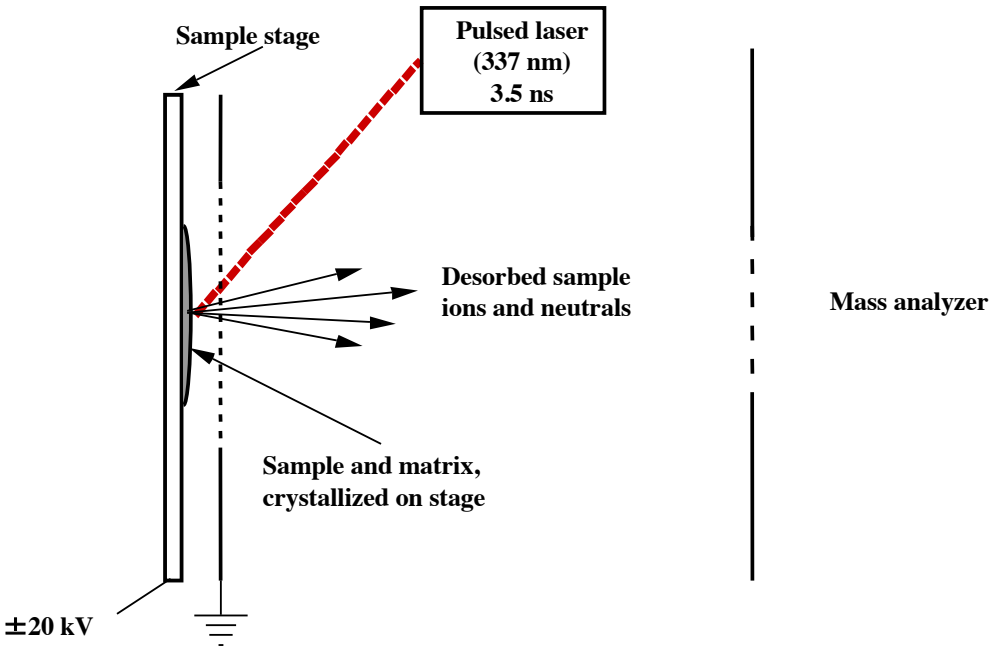


## Matrix-Assisted Laser Desorption/Ionization (MALDI)

- Analyte is dissolved in solution with excess matrix (>$10^4$).
- Sample/matrix mixture is dried on a target and placed in the MS vacuum.
- Requirements for a satisfactory matrix, it must:
  - co-crystallize with typical analyte molecules
  - absorb radiation at the wavelength of the laser (usually 337 nm)
  - transfer protons to the analyte (it should be acidic)
    - Typical matrices for UV MALDI are aromatic carboxylic acids.



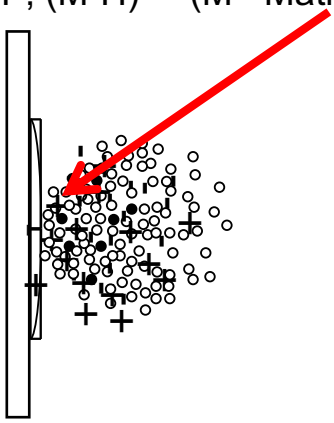
# Matrix-assisted laser desorption ionization (MALDI)






# MALDI Ionization Mechanism

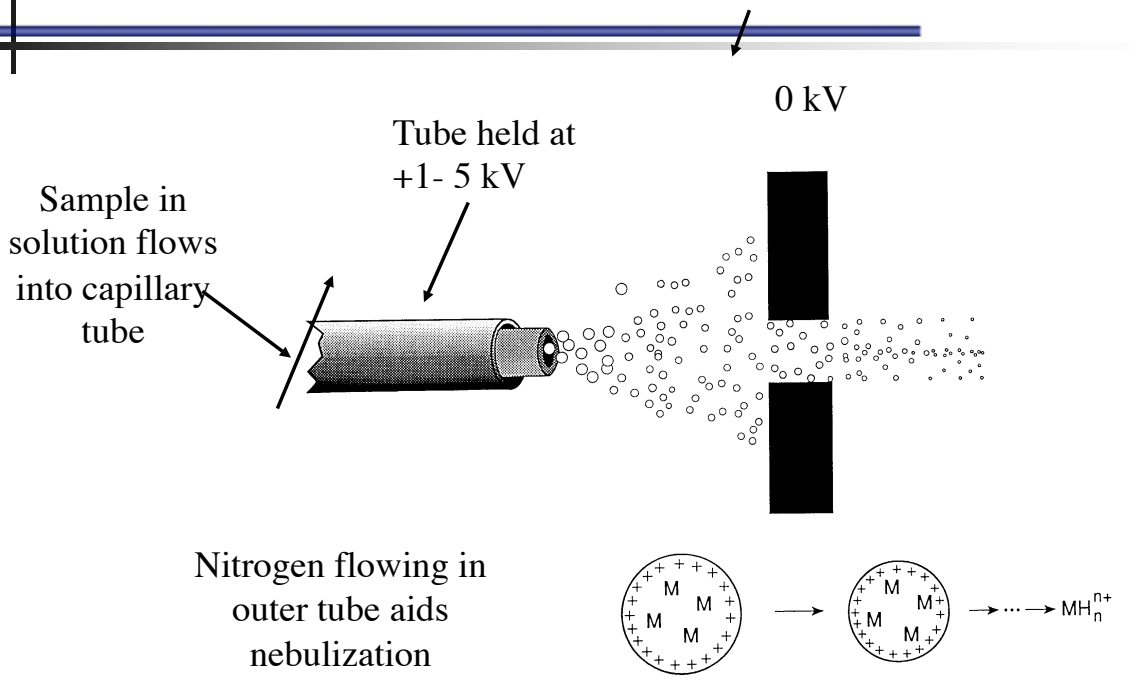
1. Laser pulse produces matrix neutrals, + and - ions, and sample neutrals:
   M --> M*, $MH^+$, $(M-H)^-$  (M= Matrix)



2. Sample molecules are ionized by gas-phase proton transfer:

   $MH^+ + A \rightarrow AH^+ + M$  (A=Analyte)
   $(M-H)^- + A \rightarrow (A-H)^- + M$

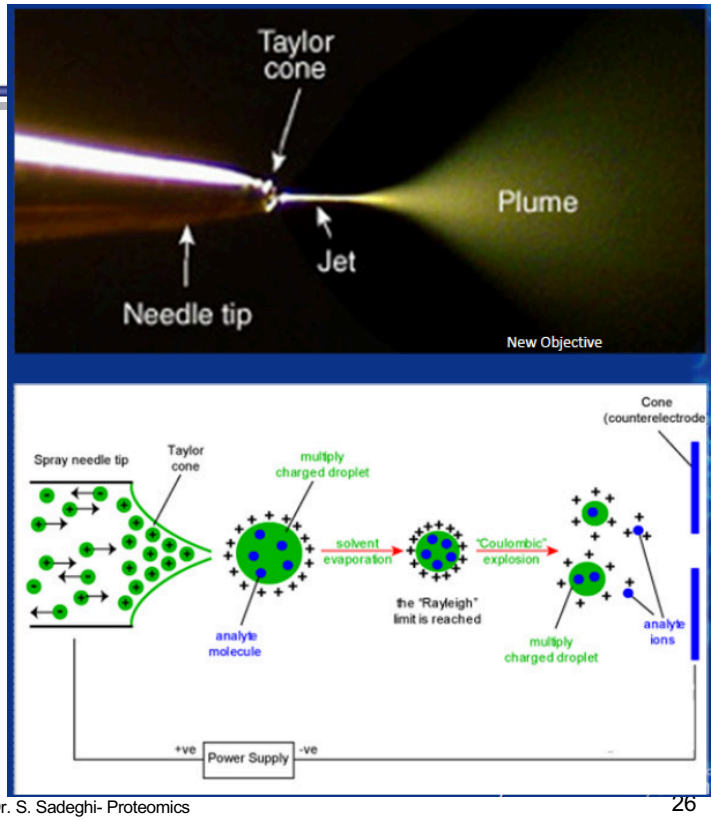
# Electrospray Ionization (ESI)



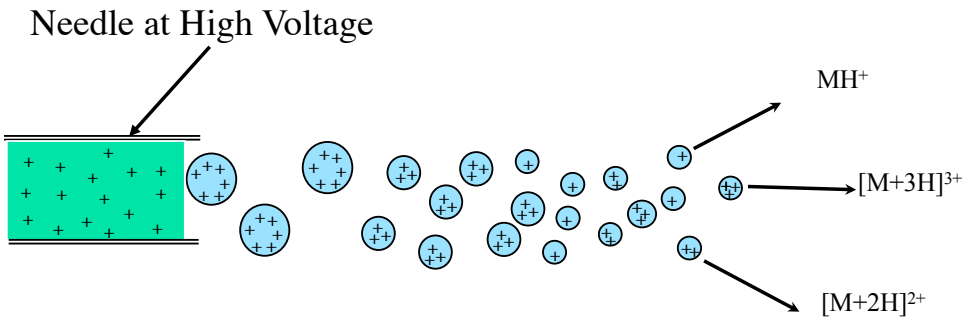



# Electrospray ionisation ESI

- High voltage placed on a fused silica column causes a spray of charged droplets which evaporate leaving charged peptides



# Electrospray Ion Formation




Droplets formed in electric field have excess positive ions.

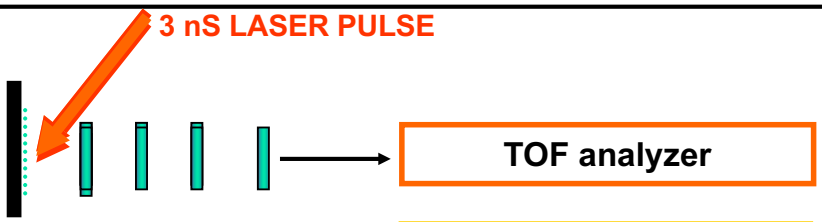
Evaporation of neutrals concentrates charge.

Droplets break into smaller droplets.

Eventually one molecule + *n* protons is left.



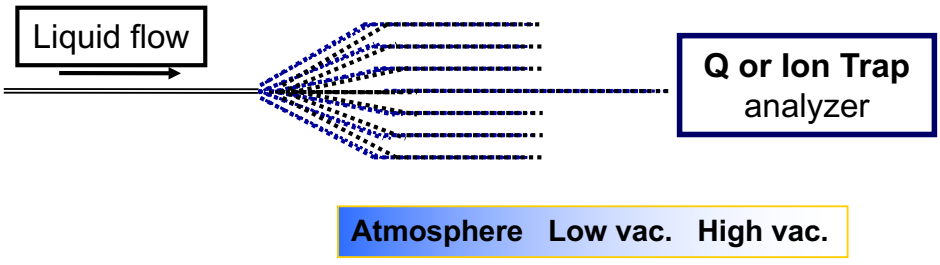
## MALDI




Sample (solid) on target at high voltage/ high vacuum

MALDI is a solid-state technique that gives ions in pulses, best suited to time-of-flight MS.

## ESI




ESI is a solution technique that gives a continuous stream of ions, best for quadrupoles, ion traps, etc.

# Ionization Methods for Biomolecule Analysis

| Electrospray | MALDI |
|---|---|
| •Online LC/MS possible | •Very long sample lifetime; repeated measurements possible |
| •Poor for mixtures without LC | •Good for mixtures |
| •Quantitation possible | •Matrix peaks can interfere at MW <600 |
| •Good for MW <600 | •Salt tolerant |
| •Generate highly charged ions | •Low maintenance |
| | •Generate ions with few charges |

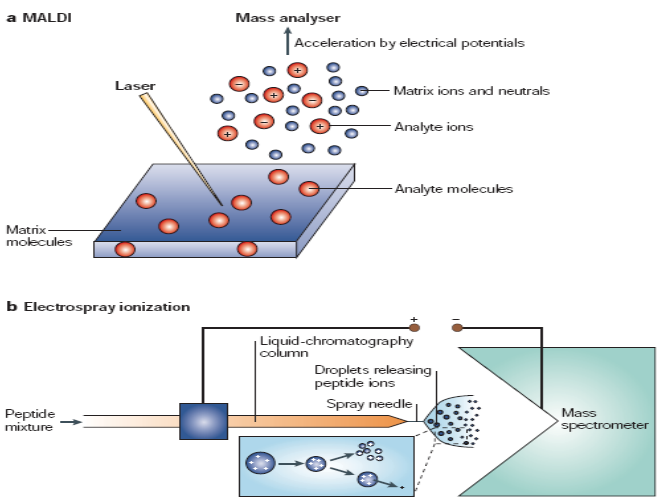


# ….MALDI or Electrospray ?

MALDI is limited to solid state, ESI to liquid

ESI is better for the analysis of complex mixture as it is directly interfaced to a separation techniques (i.e. HPLC)

MALDI is more "flexible" (MW from 200 to 400,000 Da)



# ESI vs MALDI

Ref: Glish & Vachet. *Nature Reviews*. 2003, 2, 140.

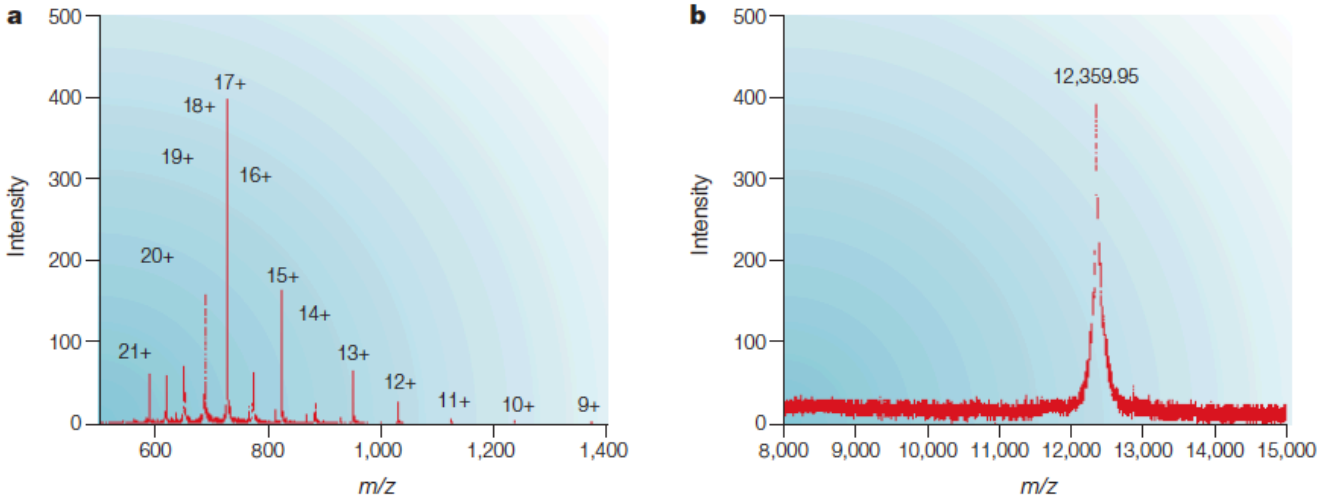



Figure 2 | **A comparison of the mass spectra for cytochrome *c* generated using electrospray ionization and matrix-assisted laser desorption/ionization. a** | Electrospray ionization (ESI) mass spectrum of cytochrome *c*: multiple peaks are observed due to the different charge states that arise from varying degrees of protonation. **b** | Matrix-assisted laser desoprtion/ionization (MALDI) mass spectrum of cytochrome *c*: only a single peak is observed for the analyte because ionization in MALDI generally occurs by the addition of a single proton. Note the different mass-to-charge (*m/z*) scales.



# Mass analysis in time-of-flight TOF

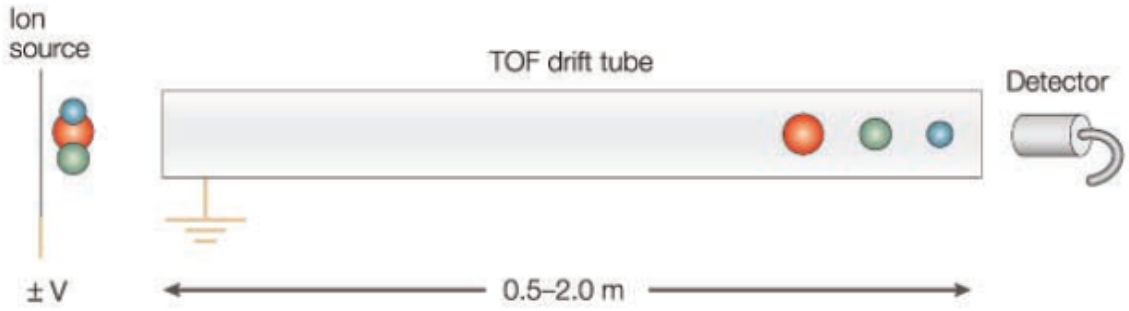
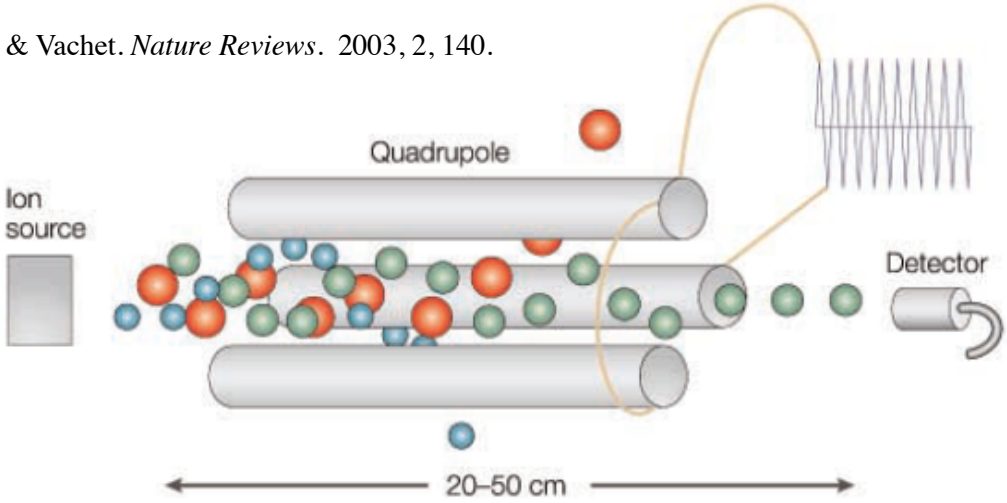



Figure 3 | **Pictorial diagrams of the common beam mass analysers viewed from above. a** | Mass analysis in time-of-flight (TOF) spectrometry is achieved because ions of different mass-to-charge (*m/z*) values have different velocities and therefore reach the detector at different times.

Ref: Glish & Vachet. *Nature Reviews*. 2003, 2, 140.

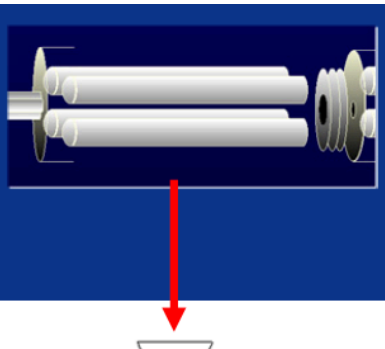
# Mass analysis in quadrupole

Ref: Glish & Vachet. *Nature Reviews*. 2003, 2, 140.

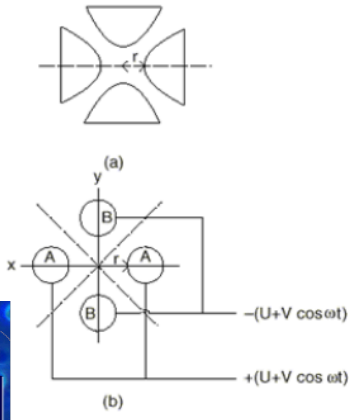


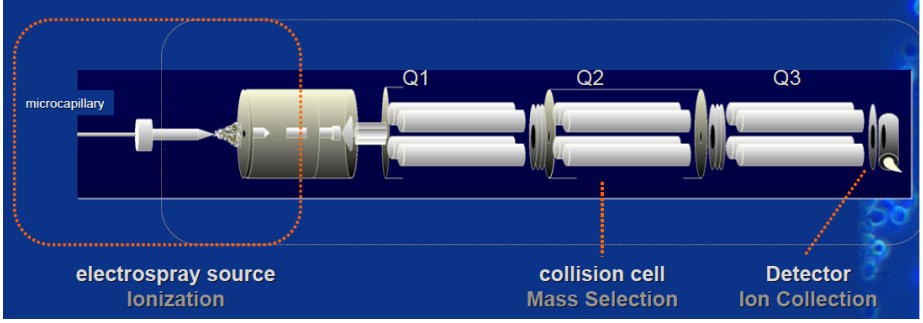

In a quadrupole mass analyses, the correct magnitude of the radio frequency and direct current voltages applied to the rods allows ions of a single *m/z* to maintain stable trajectories from the ion source to the detector whereas ions with different *m/z* values are unable to do so.





- In a quadrupole mass spectrometer four (quad) parallel rods (poles) are arranged equidistantly from a central (imaginary) axis.
- Charged ions are injected along the central axis of the quadrupole assembly.
- Static and alternating (radio frequency) electric potentials are applied to opposite pairs of rods, creating a fluctuating electric field.



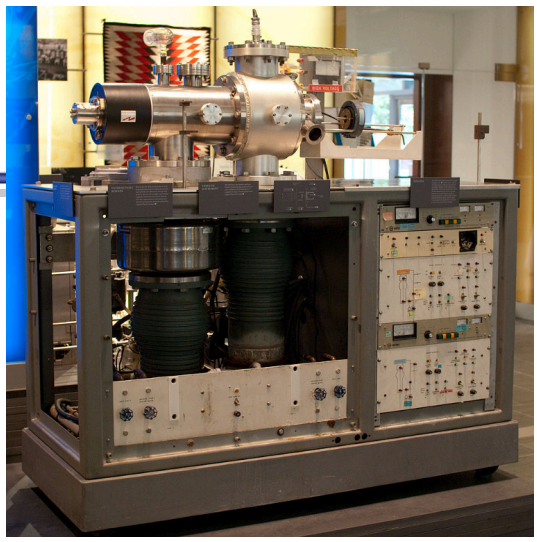




Triple Quadrupole MS

# The Nobel Prize in Chemistry 2002

John B. Fenn

Koichi Tanaka



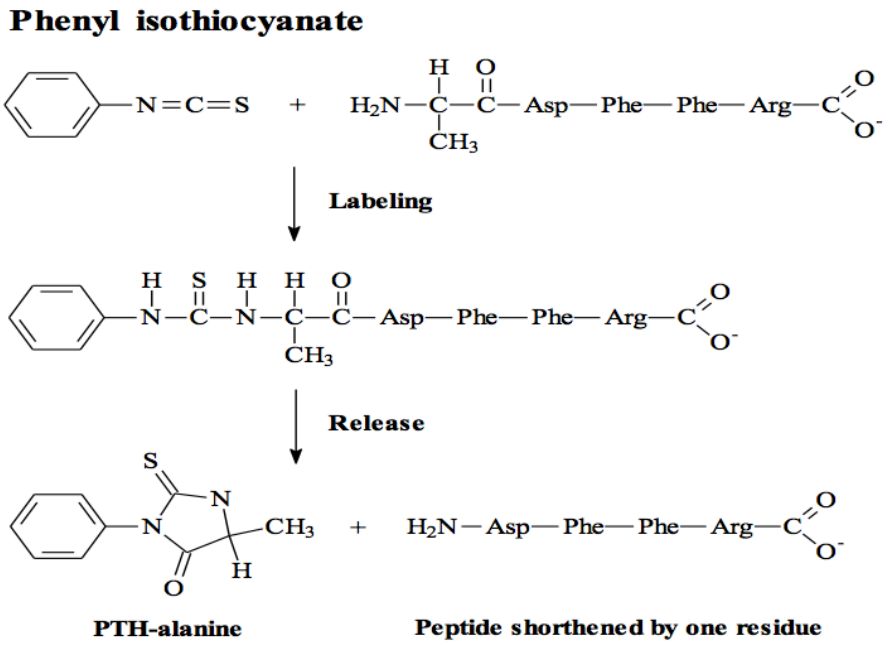
The Nobel Prize in Chemistry 2002 was awarded "*for the development of methods for identification and structure analyses of biological macromolecules*" with one half jointly to John B. Fenn and Koichi Tanaka "*for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules*"



# Peptide Analysis

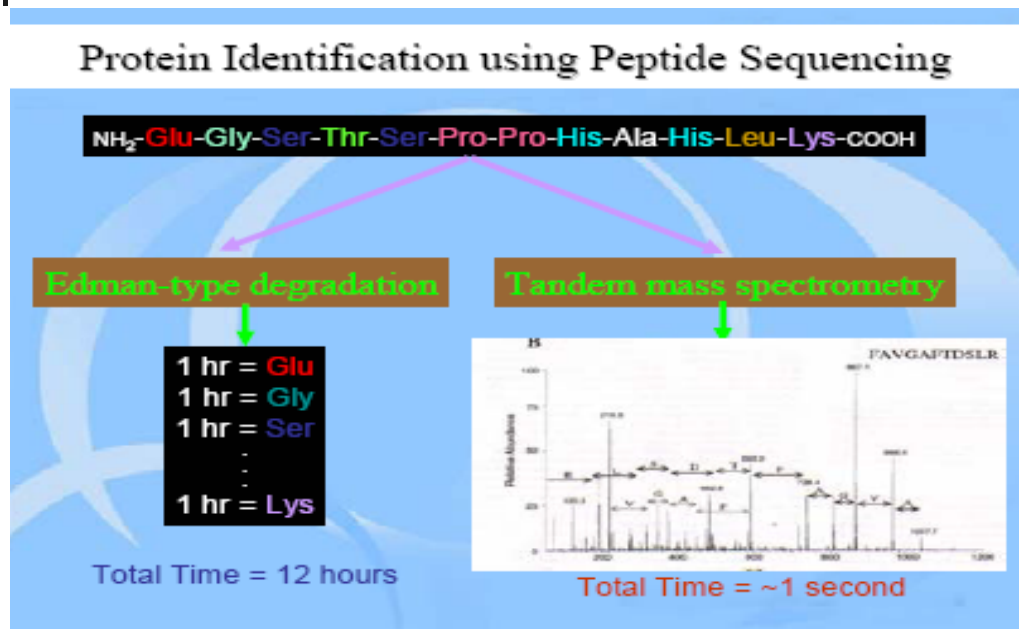
- Edman Degradation
- MS
  - More sensitive
  - Can fragment peptides faster
  - Does not require proteins or peptides to be purified to homogeneity
  - Has no problem identifying blocked or modified proteins

# Edman degradation






# Edman Degradation vs. MS/MS



# Experimental steps

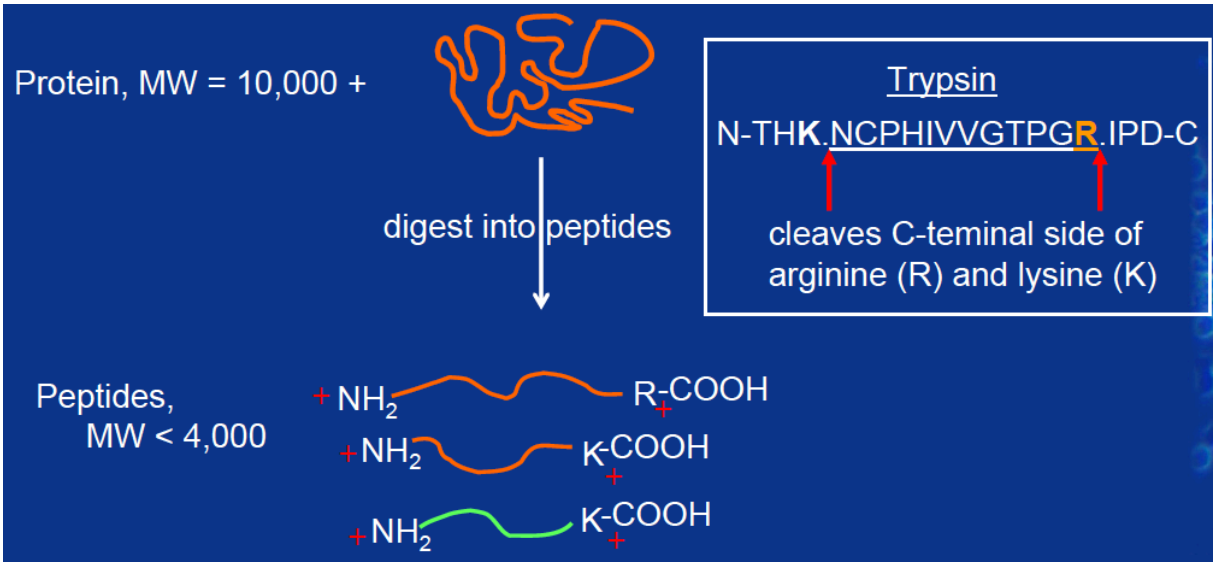
1. Proteins digested with an enzyme to produce peptides
2. Peptides charged (ionized) and separated according to their different *m/z* ratios
3. Each peptide fragmented into ions and *m/z* values of fragment ions are measured

- **Steps 2 and 3 performed within a tandem mass spectrometer.**



# Breaking Protein into Peptides and Peptides into Fragment Ions

- Proteases e.g. trypsin, break protein into *peptides*.
- A Tandem Mass Spectrometer further breaks the peptides down into *fragment ions* and measures the mass of each ion.
- MS accelerates the fragmented ions; heavier ions accelerate slower than lighter ones.
- Mass Spectrometer measure *mass/charge* ratio of an ion.





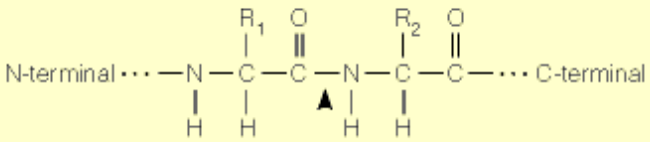

Common proteases used for shotgun proteomics.

| Protease | Cleavage Specificity[a] | Common proteomic usage |
|---|---|---|
| **Trypsin** | -K,R-↑-Z- not -K,R-↑-P- | General protein digestion |
| **Endoproteinase Lys-C** | -K-↑-Z- | Alternative to trypsin for increased peptide length; multiple protease digestion; $^{18}O$ labeling |
| **Chymotrypsin** | -W,F,Y-↑-Z- and -L,M,A,D,E-↑-Z- at a slower rate | Multiple protease digestion |
| **Subtilisin** | Broad specificity to native and denatured proteins | Multiple protease digestion |
| **Elastase** | -B-↑-Z- | Multiple protease digestion |
| **Endoproteinase Lys-N** | -Z-↑-K- | Increase peptide length; create higher charge state for ETD |
| **Endoproteinase Glu-C** | -E-↑-Z- and 3000 times slower at -D-↑-Z- | Multiple protease digestion; $^{18}O$ labeling |
| **Endoproteinase Arg-C** | -R-↑-Z- | Multiple protease digestion |
| **Endoproteinase Asp-N** | -Z-↑-D- and -Z-↑-cysteic acid- but not -Z-↑-C- | Multiple protease digestion |
| **Proteinase K** | -X-↑-Y- | Non-specific digestion of membrane-bound proteins |
| **OmpT** | -K,R-↑-K,R- | Increased peptide length for middle-down proteomics |

[a]B – uncharged, non-aromatic amino acids (i.e. A, V, L, I, G, S); X – aliphatic, aromatic, or hydrophobic amino acids; and Z – any amino acid.

Ref: Zhang et al. Chem Rev. 2013 Apr 10; 113(4): 2343–2394.

# Enzymatic degradation: Proteases:




| Enzyme | Preferred Site[a] | Source |
|---|---|---|
| Trypsin | $R_1$ = Lys, Arg | From digestive systems of animals, many other sources |
| Chymotrypsin | $R_1$ = Tyr, Trp, Phe, Leu | Same as trypsin |
| Thrombin | $R_1$ = Arg | From blood; involved in coagulation |
| V-8 protease | $R_1$ = Asp, Glu | From *Staphylococcus aureus* |
| Prolyl endopeptidase | $R_1$ = Pro | Lamb kidney, other tissues |
| Subtilisin | Very little specificity | From various bacilli |
| Carboxypeptidase A | $R_2$ = C-terminal amino acid | From digestive systems of animals |
| Thermolysin | $R_2$ = Leu, Val, Ile, Met | From *Bacillus thermoproteolyticus* |

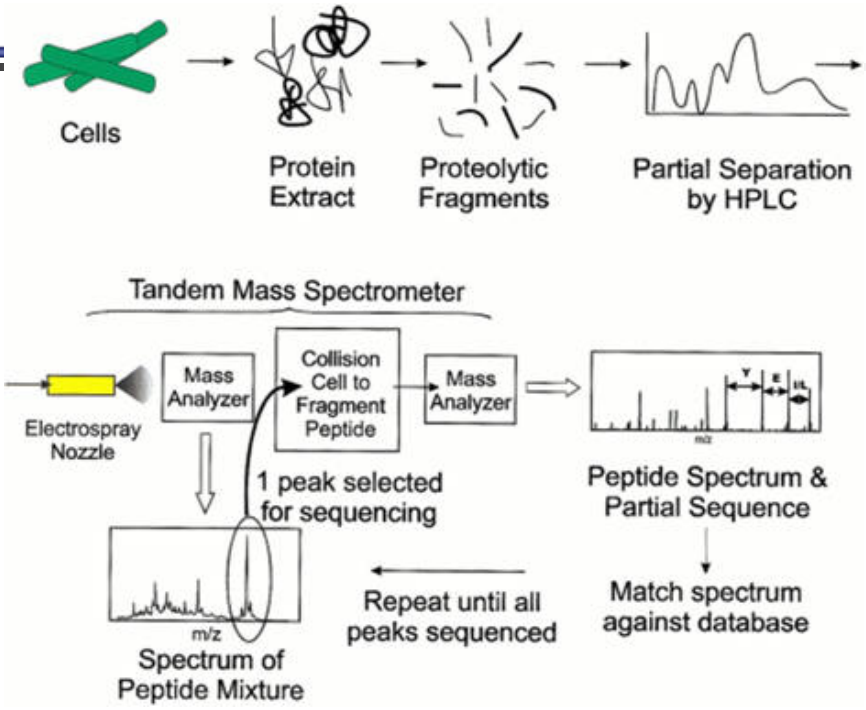
[a]The residues indicated are those next to which cleavage is most likely. Note that in some cases preference is determined by the residue on the N-terminal side of the cleaved bond ($R_1$) and sometimes by the residue to the C-terminal side ($R_2$). Generally, proteases do not cleave where proline is on the other side of the bond. Even prolyl endopeptidase will not cleave if $R_2$ = Pro.



# The overlap of the sequences of fragments allows to determine the protein sequence

| | |
|---|---|
| Chimotripsina | $H_3\overset{+}{N}$—Leu—Asn—Asp—Phe |
| Bromuro di cianogeno | $H_3\overset{+}{N}$—Leu—Asn—Asp—Phe—His—Met |
| Chimotripsina | His—Met—Thr—Met—Ala—Trp |
| Bromuro di cianogeno | Thr—Met |
| Bromuro di cianogeno | Ala—Trp—Val—Lys—$COO^-$ |
| Chimotripsina | Val—Lys—$COO^-$ |
| Sequenza complessiva | $H_3\overset{+}{N}$—Leu—Asn—Asp—Phe—His—Met—Thr—Met—Ala—Trp—Val—Lys—$COO^-$ |

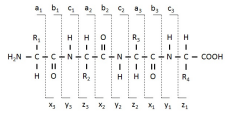
# Tandem Mass Spectrometry



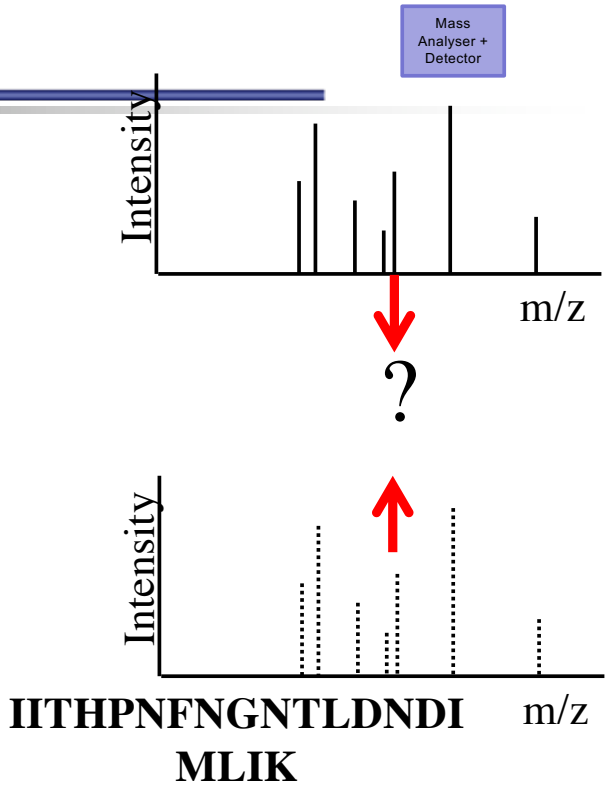



# Tandem Mass Spectrum

- Tandem Mass Spectrometry (MS/MS):
- mainly generates partial N- and C-terminal peptides
- Chemical noise often complicates the spectrum.
- Represented in 2-D: mass/charge axis vs. intensity axis



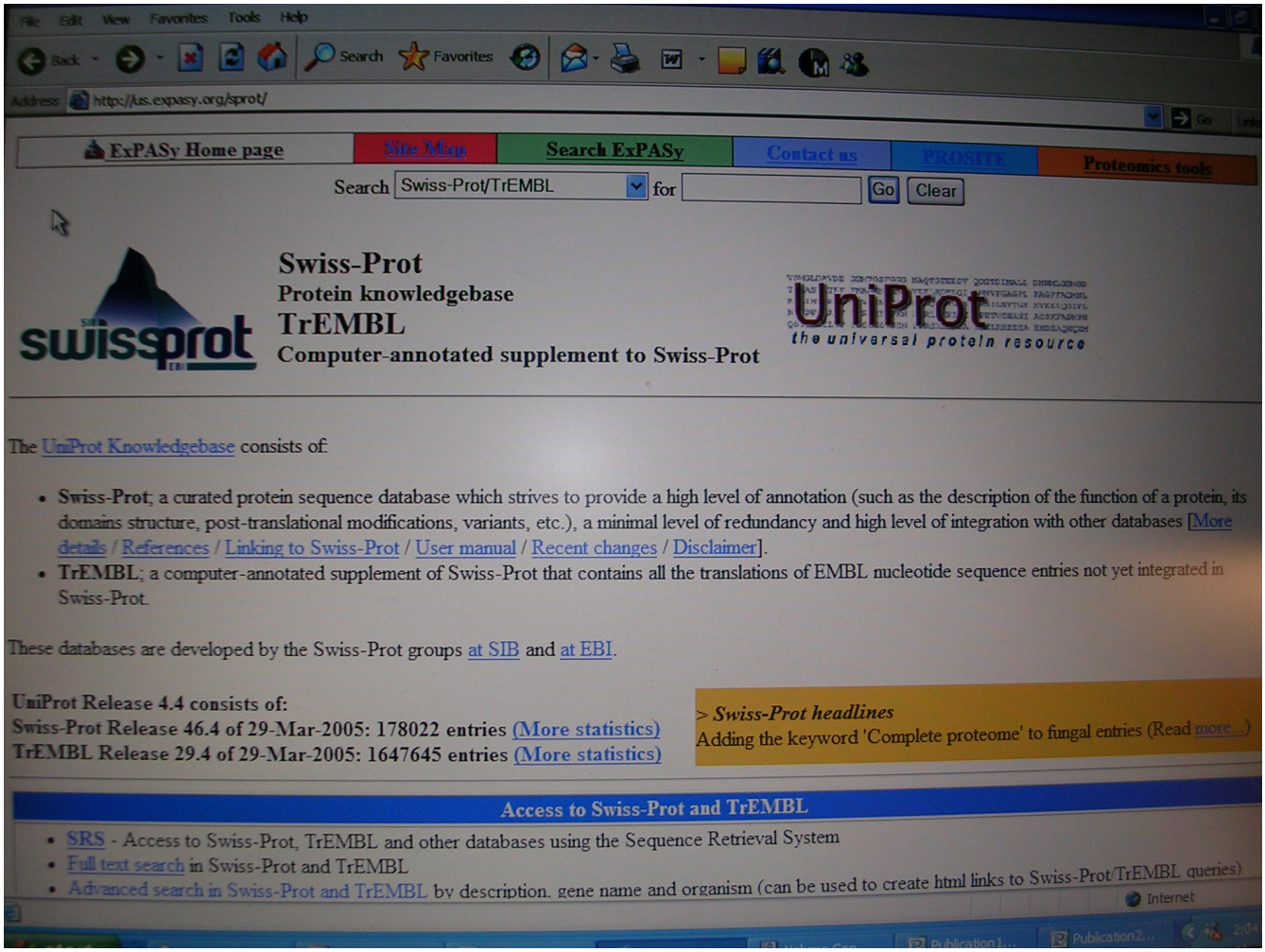
- Search fragment spectrum against a database of protein sequences. For each sequence, digest into peptides, generate an expected fragment ion spectrum, and match to observed spectrum

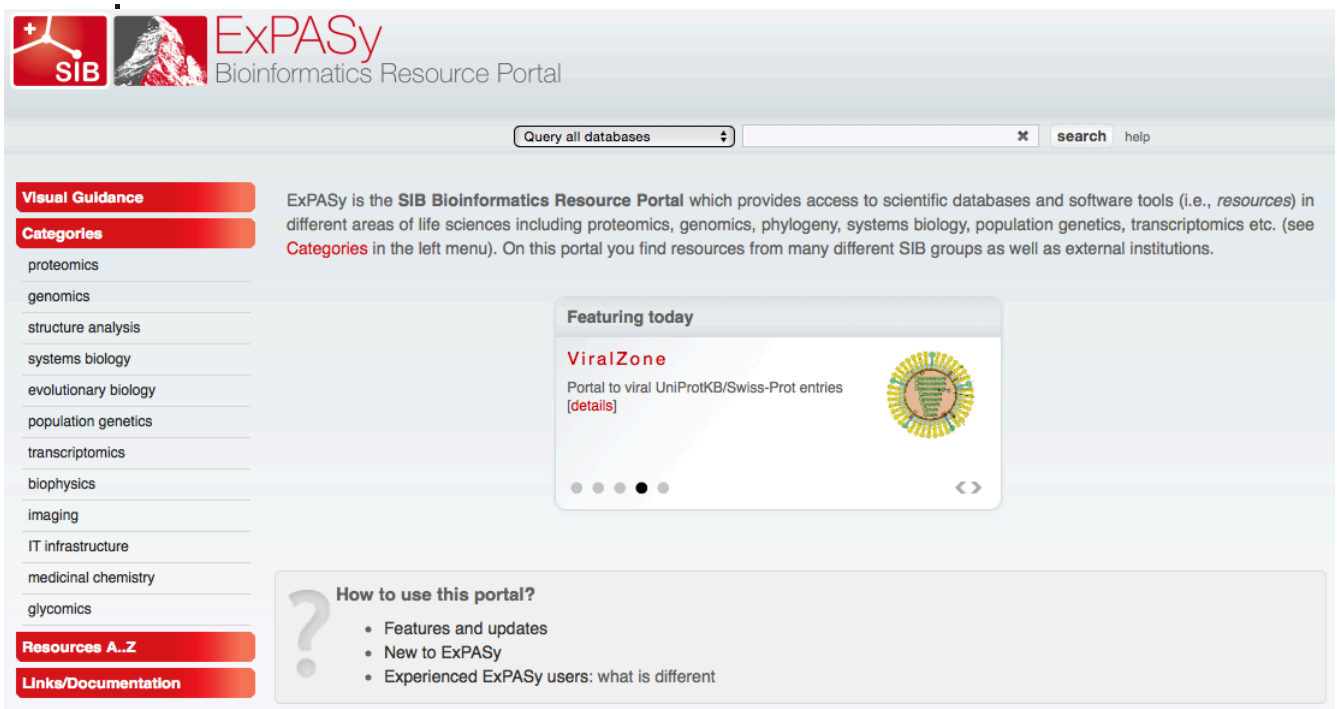


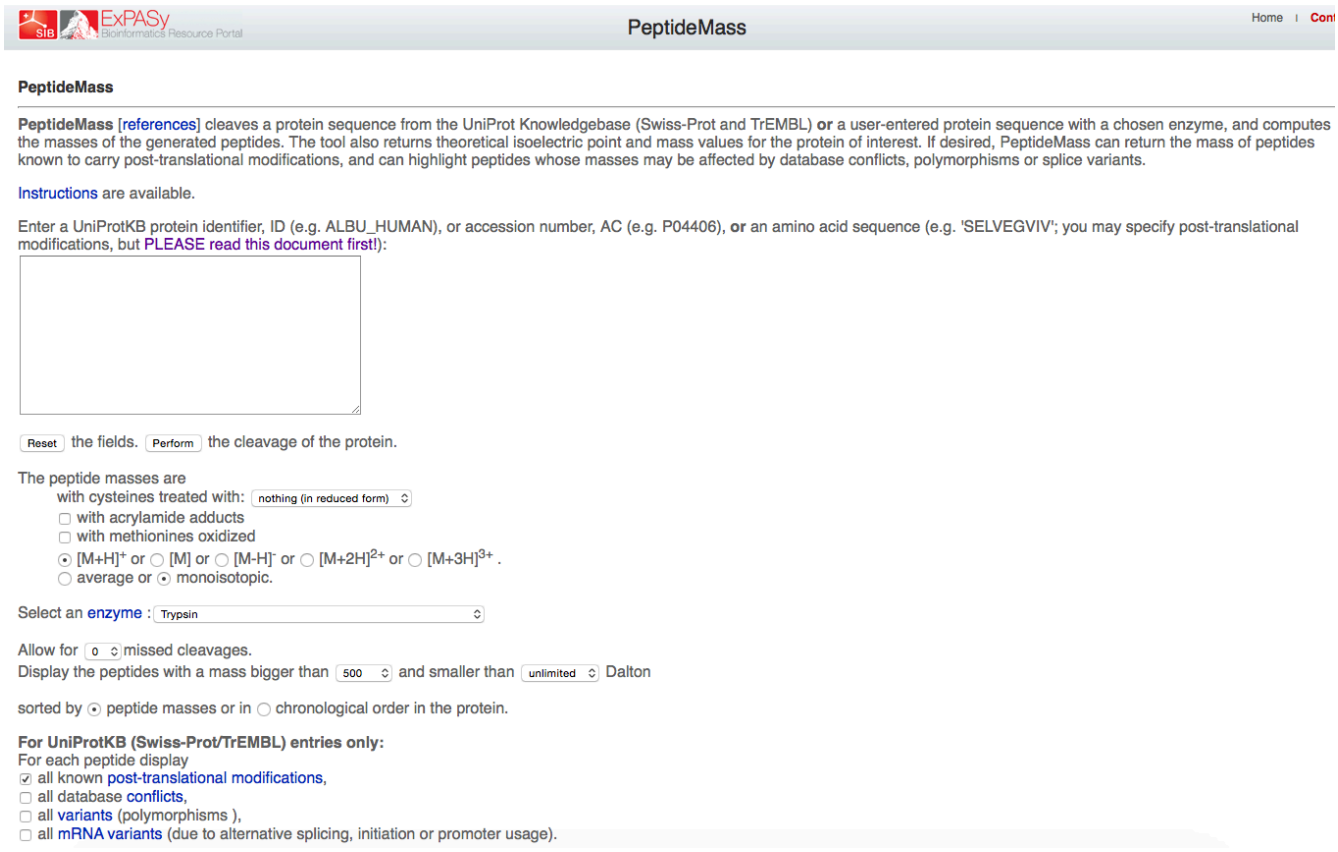



# What you need for peptide mass mapping

- Peptide mass spectrum
- Protein Database: GenBank, Swiss-Prot, dbEST
- There are multiple commonly used MS/MS fragment spectra search engines, including:
  - Mascot
  - Sequest
  - OMSSA
  - X!Tandem
  - MS Amanda
  - Andromeda
  - ProteinPilot


ExPASy Home page
Site Map
Search ExPASy
Contact us
PROSITE
Proteomics tools
Search Swiss-Prot/TrEMBL for Go Clear
Swiss-Prot
Protein knowledgebase
TrEMBL
Computer-annotated supplement to Swiss-Prot
UniProt
the universal protein resource
The UniProt Knowledgebase consists of:
Swiss-Prot; a curated protein sequence database which strives to provide a high level of annotation (such as the description of the function of a protein, its domains structure, post-translational modifications, variants, etc.), a minimal level of redundancy and high level of integration with other databases [More details / References / Linking to Swiss-Prot / User manual / Recent changes / Disclaimer].
TrEMBL; a computer-annotated supplement of Swiss-Prot that contains all the translations of EMBL nucleotide sequence entries not yet integrated in Swiss-Prot.
These databases are developed by the Swiss-Prot groups at SIB and at EBI.
UniProt Release 4.4 consists of:
Swiss-Prot Release 46.4 of 29-Mar-2005: 178022 entries (More statistics)
TrEMBL Release 29.4 of 29-Mar-2005: 1647645 entries (More statistics)
> Swiss-Prot headlines
Adding the keyword 'Complete proteome' to fungal entries (Read more...)
Access to Swiss-Prot and TrEMBL
SRS - Access to Swiss-Prot, TrEMBL and other databases using the Sequence Retrieval System
Full text search in Swiss-Prot and TrEMBL
Advanced search in Swiss-Prot and TrEMBL by description, gene name and organism (can be used to create html links to Swiss-Prot/TrEMBL queries)


SIB
ExPASy
Bioinformatics Resource Portal
Query all databases
search
help
Visual Guidance
Categories
proteomics
genomics
structure analysis
systems biology
evolutionary biology
population genetics
transcriptomics
biophysics
imaging
IT infrastructure
medicinal chemistry
glycomics
Resources A..Z
Links/Documentation
ExPASy is the SIB Bioinformatics Resource Portal which provides access to scientific databases and software tools (i.e., resources) in different areas of life sciences including proteomics, genomics, phylogeny, systems biology, population genetics, transcriptomics etc. (see Categories in the left menu). On this portal you find resources from many different SIB groups as well as external institutions.
Featuring today
ViralZone
Portal to viral UniProtKB/Swiss-Prot entries
[details]
How to use this portal?
Features and updates
New to ExPASy
Experienced ExPASy users: what is different






# Mascot
# identifies proteins by interpreting MS data

- The prevailing experimental method for protein identification is a bottom-up approach, where a protein sample is typically digested with Trypsin to form smaller peptides.
- Proteins are too big, peptides usually fall within the limited mass range that a typical mass spectrometer can measure.
- Mass spectrometers measure the MW of peptides in a sample.
- Mascot then compares these molecular weights against a database of known peptides.
- The program cleaves every protein in the specified search database in silico according to specific rules depending on the cleavage enzyme used for digestion and calculates the theoretical mass for each peptide.
- Mascot then computes a score based on the probability that the peptides from a sample match those in the selected protein database. The more peptides Mascot identifies from a particular protein, the higher the Mascot score for that protein.

Home | Mascot database search | Products | Technical support | Training | News | Blog | Newsletter | Contact

Access Mascot Server | Database search help

Mascot database search > Access Mascot Server

# Access Mascot Server

You are welcome to submit searches to this free Mascot Server. Searches of MS/MS data are limited to 1200 spectra and some functions, such as no enzyme searches, are unavailable. Automated searching of batches of files is not permitted. If you want to automate search submission, perform large searches, search additional sequence databases, or customise the modifications, quantitation methods, etc., you'll need to license your own, in-house copy of Mascot Server.

## Peptide Mass Fingerprint

The experimental data are a list of peptide mass values from the digestion of a protein by a specific enzyme such as trypsin.

Perform search | Example of results report | Tutorial

## Sequence Query

One or more peptide mass values associated with information such as partial or ambiguous sequence strings, amino acid composition information, MS/MS fragment ion masses, etc. A super-set of a sequence tag query.

Perform search | Example of results report | More information

## MS/MS Ions Search

Identification based on raw MS/MS data from one or more peptides.

Perform search | Example of results report | Tutorial

### More info

- Mascot overview
- Search parameter reference
- Data file format
- Results report overview

# Database search for protein identification

HOME
MASCOT
HELP
WHAT'S NEW
PRODUCTS
SUPPORT
SITE SEARCH
LINKS
EMPLOYMENT
CONTACT US

## Mascot: Peptide Mass Fingerprint

Your name | Email

Search title

Database: NCBInr

Taxonomy: All entries

Enzyme: Trypsin | Allow up to 1 missed cleavages

Fixed modifications: AB_old_ICATd0 (C), AB_old_ICATd8 (C), Acetyl (K), Acetyl (N-term), Amide (C-term)

Variable modifications: AB_old_ICATd0 (C), AB_old_ICATd8 (C), Acetyl (K), Acetyl (N-term), Amide (C-term)

Protein mass: kDa | Peptide tol. ± 1.0 Da

Mass values: MH+ / Mr | Monoisotopic / Average

Data file: Browse...

Query: NB Contents of this field are ignored if a data file is specified.

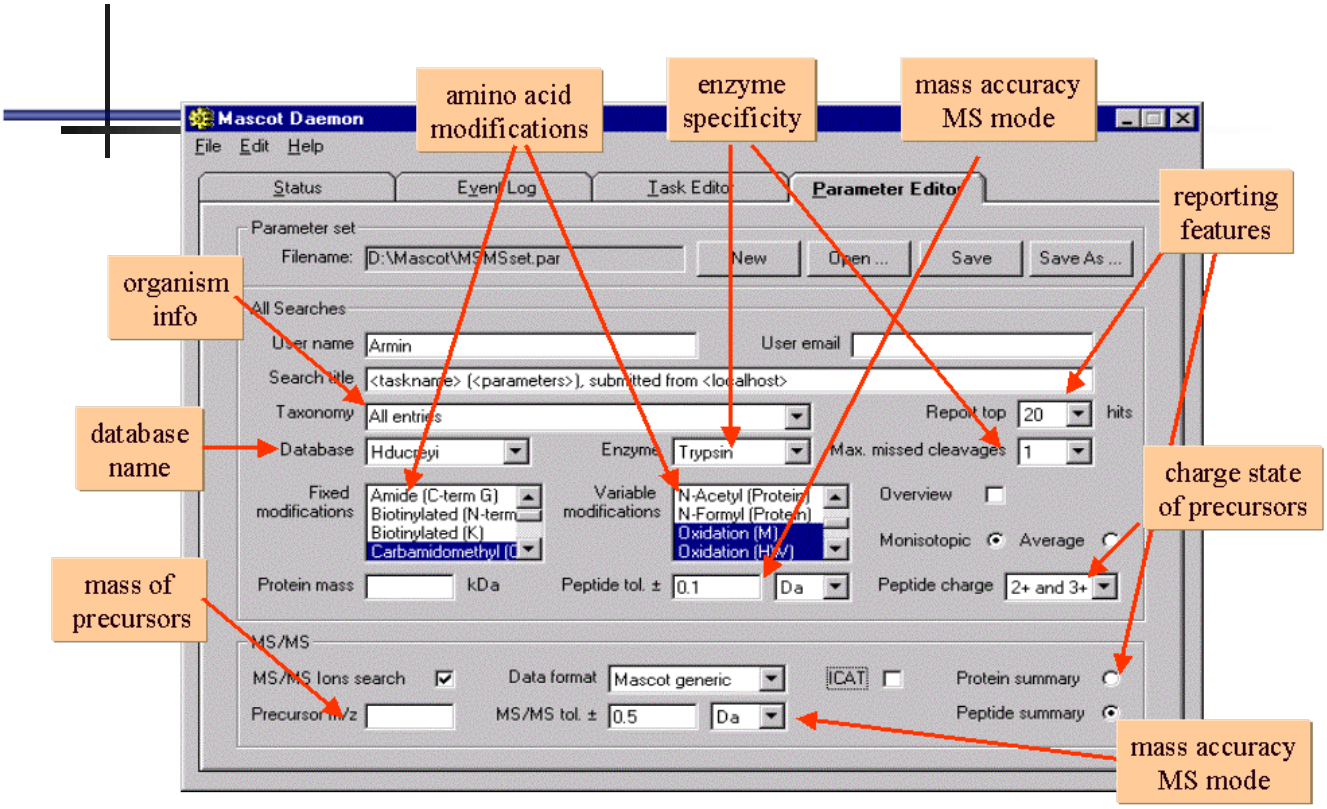
Overview | Report top 20 hits

Start Search ... | Reset Form



Internet

# Mascot






Ion mode and adducts:

| positive | negative | neutral |
|---|---|---|
| ◉ $[M+H]^+$ | ○ $[M-H]^-$ | ○ [M] |
| ○ $Na^+$ or ○ $K^+$ | ○ acetate or ○ trifluoroacetic acid | |
| ○ other: mass: | ○ other: mass: | |

| ◉ N-linked oligosaccharides | OR | ○ O-linked oligosaccharides |
|---|---|---|
| Form of N-linked oligosaccharide: Glycopeptides (motif N-X-S/T/C (X not P) will be used) | | Form of O-linked oligosaccharide: Glycopeptides (only those containing S or T will be used) |

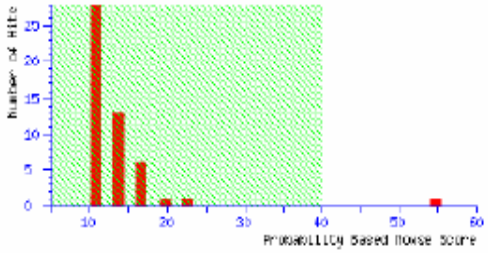
if 'Glycopeptides', please specify:

A protein sequence or a Swiss-Prot/TrEMBL ID or AC:

Enzyme: Trypsin
max. 0 missed cleavage sites (MC).
Cysteines treated with: nothing (in reduced form)
☐ acrylamide adducts on cysteines ☐ methionines oxidized

Alternatively, you can also enter a set of unmodified peptide masses ([M]). These masses must be average or monoisotopic in agreement with that specified above for experimental masses:

# Peptide sequencing using MASCOT



Peptide Summary Report

Switch to Protein Summary Report

To create a bookmark for this report, right click this link: Peptide Summary Report (../data/20021008/FoteIca.dat)

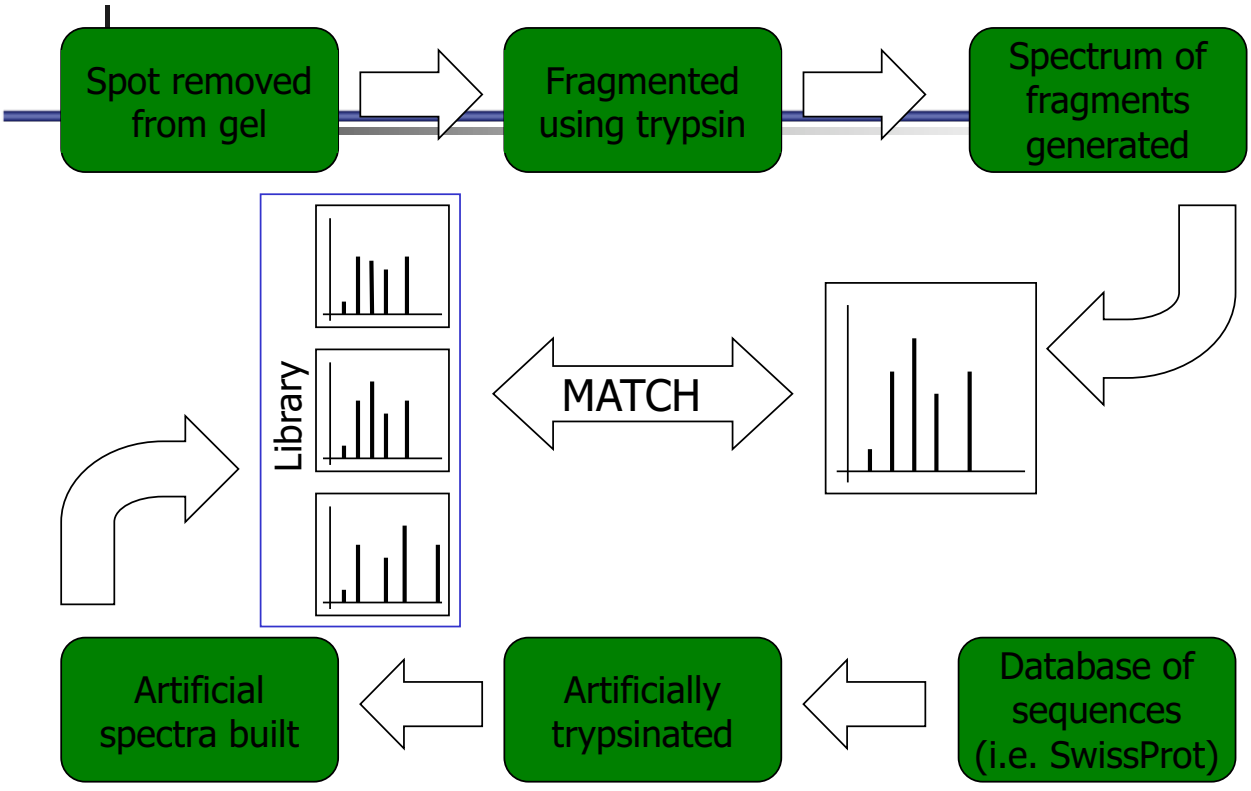
Select All | Select None | Search Selected | ☐ Error tolerant

1. gi|16924319 Mass: 40477 Total score: 55 Peptides matched: 1
(BC017450) Unknown (protein for IMAGE:3538275) [Homo sapiens]
☐ Check to include this hit in error tolerant search

| Query | Observed | Mr(expt) | Mr(calc) | Delta | Miss | Score | Rank | Peptide |
|---|---|---|---|---|---|---|---|---|
| ☑ 14 | 895.70 | 1789.39 | 1789.88 | -0.50 | 0 | 55 | 1 | SYELPDGQVITIGNER |

Proteins matching the same set of peptides:
gi|4501887 Mass: 41766 Total score: 55 Peptides matched: 1
(NM_001614) actin, gamma 1 propeptide; cytoskeletal gamma-actin; actin, cytoplasmic 2 [Homo sapiens]
gi|16359158 Mass: 41736 Total score: 55 Peptides matched: 1
(BC016045) actin, beta [Homo sapiens]
gi|4885049 Mass: 41992 Total score: 55 Peptides matched: 1
(NM_005159) actin, alpha, cardiac muscle precursor [Homo sapiens]
gi|14714562 Mass: 18762 Total score: 55 Peptides matched: 1

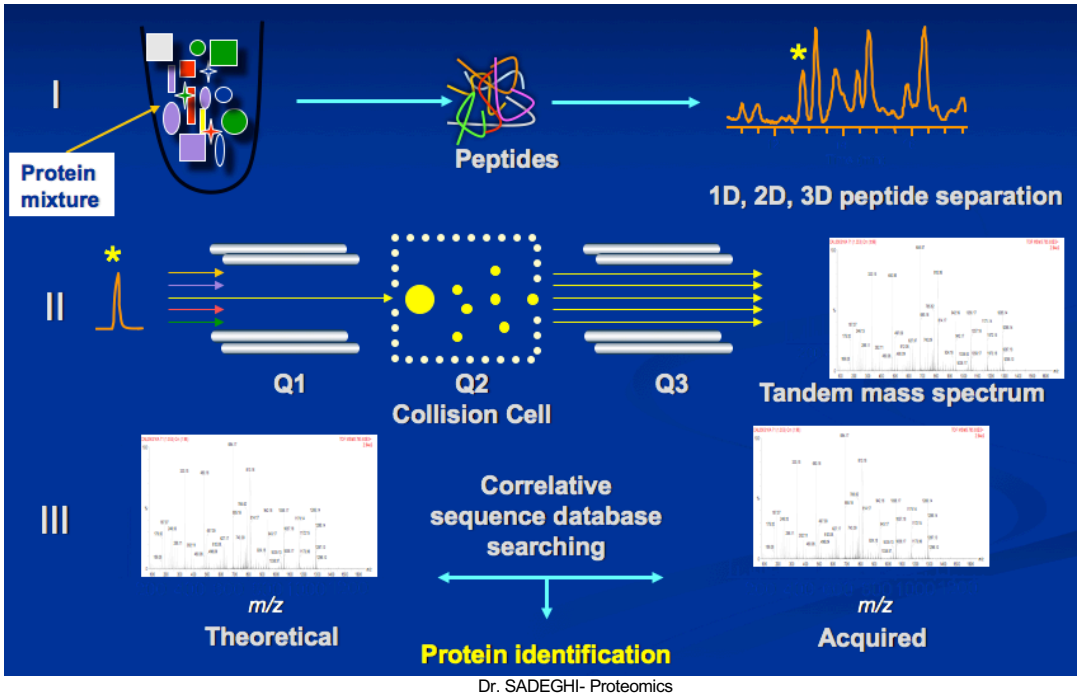
# Protein Identification by MS



# Applications



# 1. Protein identification



# 2. Protein-protein interactions
## (a) Y2H system

- Biological processes are carried out by interactions between many biomolecules.
- There are diverse types of interactions, such as
  - protein-DNA or RNA,
  - protein-protein interactions (are challenging to study due to their high diversity)
- The classic approach to studying protein-protein interactions is yeast two-hybrid (Y2H) system which was introduced more than 20 years ago.

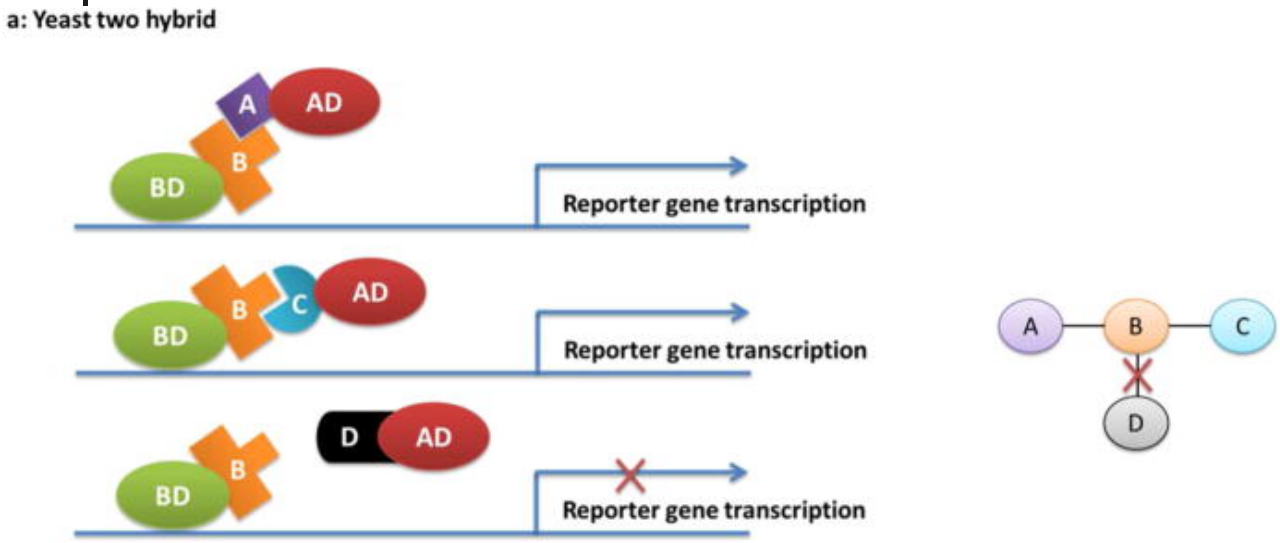
Ref: Zhang et al. Chem Rev. 2013 Apr 10; 113(4): 2343–2394.



# 2. Protein-protein interactions
## (a) Y2H system

- In a Y2H experiment, a transcript factor is split into two subunits, one is the *binding domain (BD)* and the other one is the *activating domain (AD)*.
- The engineered bait protein is fused to the BD and the second protein (prey) is fused to the AD. If the bait and prey proteins interact with each other, the transcript factor can be activated, starting the transcription of reporter gene.
- Y2H was designed to investigate direct binary interactions.
- The initial Y2H experiments focused on the interactions between a limited number of proteins. However, after genome scaled resources of open reading frames (ORFeomes) became available, comprehensive network maps have been drawn for various model organisms by large scale Y2H, including Saccharomyces cerevisiae, Drosophila melanogaster, Caenorhabditis elegans, and even humans.

# 2. Protein-protein interactions
## (a) Y2H system



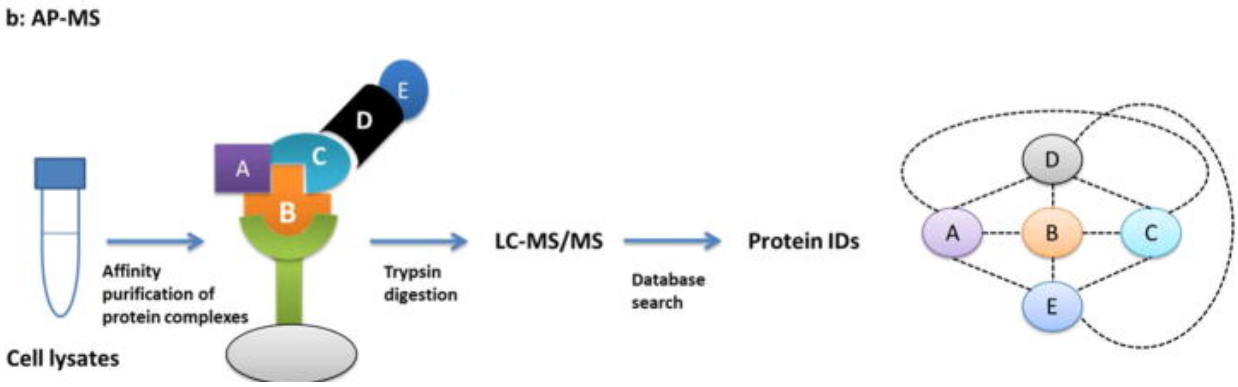

Ref: Zhang et al. Chem Rev. 2013 Apr 10; 113(4): 2343–2394.



# 2. Protein-protein interactions
## (b) AP-MS

- An approach complementary to Y2H is based on affinity purification and mass spectrometry (AP-MS)
- In an AP-MS experiment the target protein of interest, together with its interacting partners, is purified from a protein mixture.
- The purified protein complex is then subjected to shotgun proteomic identification and quantification

# 2. Protein-protein interactions
## (b) AP-MS




b: AP-MS is used to identify the whole protein complex. All the interactors binding to protein B, including both direct and indirect binders, are identified by shotgun proteomics.

Ref: Zhang et al. Chem Rev. 2013 Apr 10; 113(4): 2343–2394.

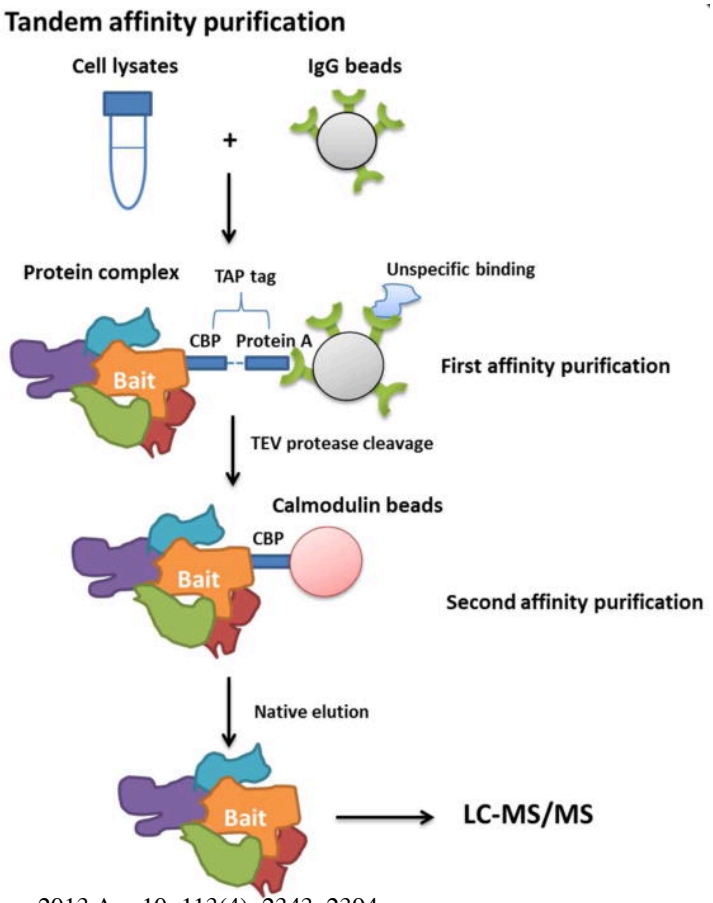


# 2. Protein-protein interactions
## (b) AP-MS

- Ideally, interactome experiments based on AP-MS should employ high quality monoclonal antibodies against the bait proteins.
- this is sometimes difficult due to a lack of good antibodies.
- Commonly, an engineered protein with an affinity has been used in AP-MS based protein interactome studies.
- Instead of using a specific antibody against target protein, the *affinity tag system* employs a uniform tag specific purification which can be used for many bait proteins.

- A widely used tag system is tandem affinity purification (TAP) which consists of :
  - calmodulin-binding peptide (CBP) and
  - protein A of Staphylococcus aureus (ProtA),
  - linked by a tobacco etch virus (TEV) cleavage site
- The TAP fused protein and its interactors are first pulled down by the distal ProtA affinity tag, and then released by cutting at the TEV cleavage site.
- The bait protein complex is subsequently subjected to the second purification step, which binds the CBP.






- Scheme of tandem affinity purification (TAP). Two steps of purification significantly remove the unspecific binding proteins.
- The advantage of TAP compared with the normal single-step procedure is reduced background protein levels.
  - all the proteins non-specifically binding to the affinity beads are excluded after the second purification.

Ref: Zhang et al. Chem Rev. 2013 Apr 10; 113(4): 2343–2394.

# 3. New Tumour markers

- Until 2010, strategy for tumour marker search= 2-DE
  - Comparing proteome between healthy + tumour tissue
  - Removal of spots + identification by MALDI
- Nowadays, LC-MS/MS:
  - Liquid chromatography coupled to tandem MS
  - Allows for separation before MS
  - Can be done with complex digested protein mixtures (shotgun)



- PSME3 (proteasome activator complex subunit 3)
  - Intracellular CRC-associated protein
  - Discovered by LC-MS/MS
  - Up-regulation of this protein would have been missed by image analysis
    - This protein was masked by another co-migrating high abundant protein (annexin A4)

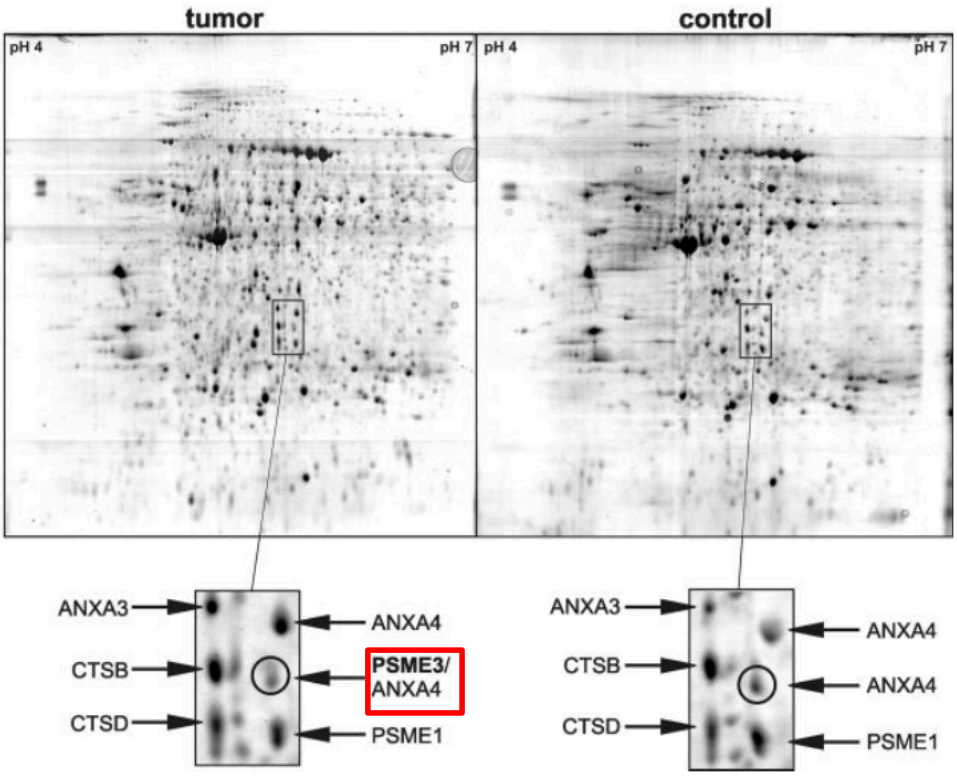



FIG. 1. **Representative 2-D gel of human primary colorectal cancer.** Tissue lysates of malignant (*tumor*) and healthy (*control*) stripped mucosa 1.5 mg of protein were subjected to 2-DE with a first dimension pH gradient of 4–7. Gels were stained with Coomassie Blue. Spots were excised, processed and analyzed by peptide mass fingerprinting. Protein identification of selected spots is indicated. *ANXA3*, an- nexin A3; *CTSB*, cathepsin B; *CTSD*, ca- thepsin D.

REF: Roessler et al. Mol Cell Proteomics. 2006;5(11):2092-101.



# Learning Outcome

- Different components of MS
  - Ionisation methods
    - MALDI
    - ESI
- Applications of MS
  - Protein identification
  - Yeast-two-hybrid
  - New tumour markers

# References

R. Aebersold and D. Goodlett. “Mass Spectrometry in Proteomics” *Chem. Rev.,* 2001, 101, 269.

G.L. Glish and R.W. Vachet. “The basics of mass spectrometry in the twenty-first century”. *Nature Reviews*. 2003, 2, 140.

T. S. Nuhse, A. Stensballe, O. Jensen, and S. Peck. “Large-scale Analysis of *in Vivo* Phosphorylated Membrane Proteins by Immobilized Metal Ion Affinity Chromatography and Mass Spectrometry” *Molecular & Cellular Proteomics, 2003, 2.11,* 1234.

Roessler et al. Mol Cell Proteomics. 2006 Nov;5(11):2092-101.

H. Steen and M. Mann. “The ABC’s (and XYZ’s) of Peptide Sequencing” Molecular Cell Biology, *Nature Reviews*. 2004, 5, 699.

Zhang et al. Chem Rev. 2013 Apr 10; 113(4): 2343–2394