Introduction of Proteomics

CH 500

By Ekawit Threenet
Biochemistry

Citations and Sources

Josh Leung
Biology 1220
April 13th, 2010

- Nature Insight Review Articles: Proteomics (2003)
  http://www.nature.com/nature/insights/6928.html
- Large-scale analysis of the yeast proteome by multidimensional protein identification technology
  http://proteome.gs.washington.edu/classes/Genome490/papers/Washburn_et_al_Nat_Biotech_2001.pdf

  http://www.biotechniques.com/multimedia/archive/00001/BTN_A_000112604_O_1428a.pdf
  http://www.pnas.org/content/99/18/11564.full
  http://pcarvalho.com/patternlab/mudpitsim.shtml
  http://en.wikipedia.org/wiki/Proteomics
  http://en.wikipedia.org/wiki/Fourier_transform_ion_cyclotron_resonance
  http://en.wikipedia.org/wiki/Two-dimensional_gel_electrophoresis
  http://masspec.scripps.edu/mshistory/whatisms_details.php#Basics
  http://pcf.epfl.ch/page58412.html
- Yates Lab: Developers of MUDPIT
  http://fields.scripps.edu/?q=content/home
- Images:
  http://www.scq.ubc.ca/wp-content/uploads/2006/08/Proteomics.gif
  http://www.pnas.org/content/99/18/11564.full
  http://www.nature.com/nrd/journal/v2/n2/images/nrd1011-f5.gif
  http://www.bio.davidson.edu/Courses/genomics/2004/Farrow/FUN19%20ExPASy.jpg
  http://www.mth.kcl.ac.uk/~tcoolen/SystemsBiology/pics/proteome.jpg
  http://www.scripps.edu/newsandviews/e_20021007/MudPIT.jpg
  http://www.nature.com/nature/journal/v422/n6928/images/nature01511-f1.2.jpg

Dimitri Raptis & Alexander Koegel

www.draptis.eu/proteomics.ppt

The PROTEin complement of the genOME.

Judson Hervey

References

- Nature Insight: Proteomics. Nature 422: 191-237.
- Zhu, H. *et al.* Proteomics. Annual Review of Biochemistry 72: 783-812.
- Griffiths *et al.* Modern Genetic Analysis. Online: http://ncbi.nih.gov

Overview

- Proteomics: Study of the complete complement of proteins present in a cell or system of cells

- Includes:
  - Protein structures, quantities, functions, locations
  - Post-translational modifications
  - Study of protein interactions and complexes
  - How each of the above change with time and in

What is Proteomics?

- Defined as “the analysis of the entire protein complement in a given cell, tissue, or organism.”
- Proteomics “also assesses activities, modifications, localization, and interactions of proteins in complexes.”
- Proteomes of organisms share intrinsic differences across species and growth conditions.

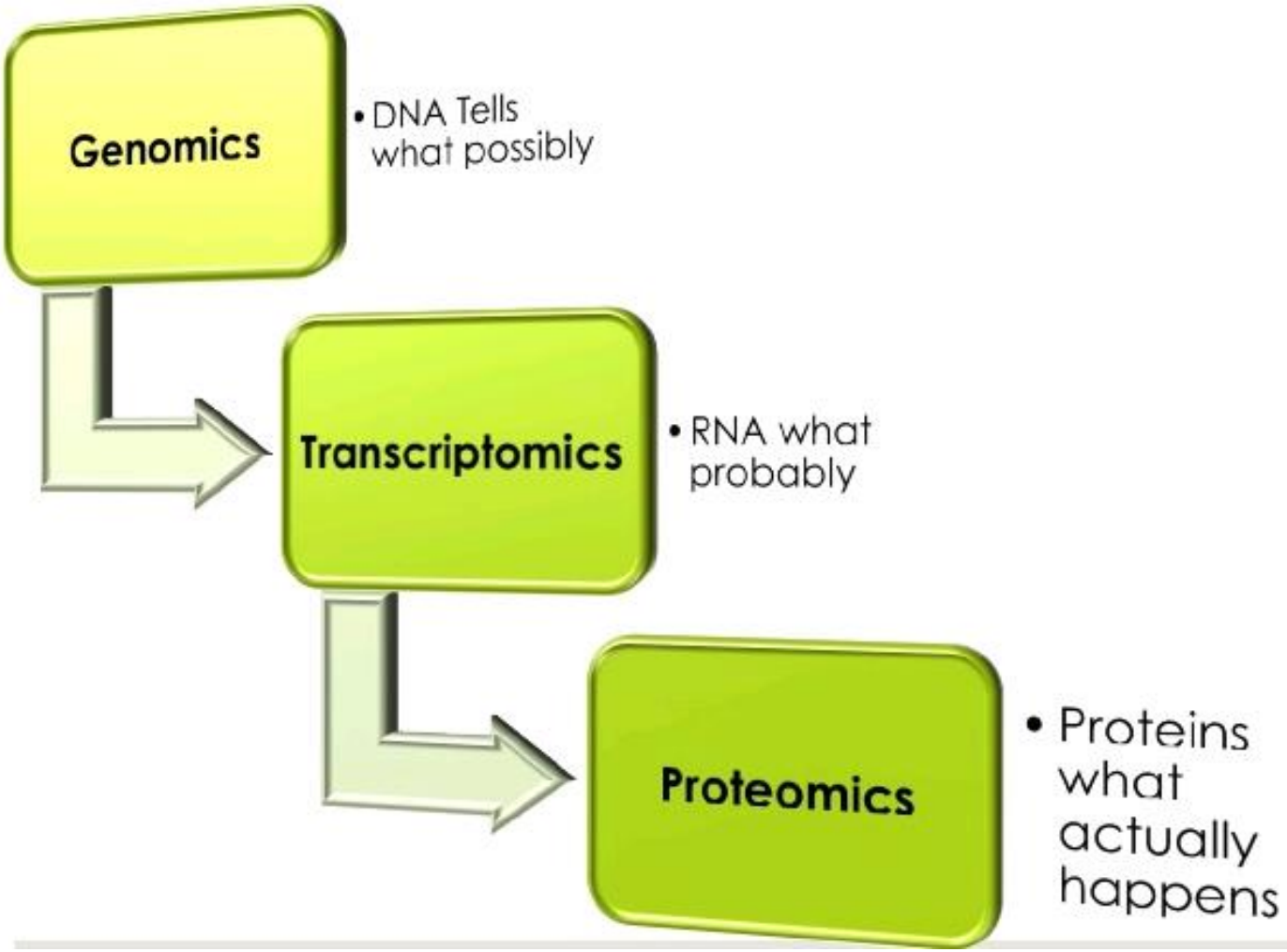
But what about the Genome?

- What does having the genome of an organism give us?
  - A great diagram, or "blueprint," of the genes within an organism.
  - Think of the genome as code that needs compiled into functional units.
  - The genome gets "compiled" into the proteome via the central dogma of biology.
  - Proteomic strategies attempt to utilize information from the genome in an attempt to conceptualize protein function.

Challenges facing Proteomic Technologies

- Limited/variable sample material
- Sample degradation (occurs rapidly, even during sample preparation)
- Vast dynamic range required
- Post-translational modifications (often skew results)
- Specificity among tissue, developmental and temporal stages
- Perturbations by environmental (disease/drugs) conditions
- Researchers have deemed sequencing the genome "easy," as PCR was able to assist in overcoming many of these issues in genomics.

Why proteomics?



Why Proteomics?

- Proteomics grew out of Genomics
  - Level of mRNA not necessarily equal to level of protein product
  - Some transcripts give rise to multiple products (alternative splicing, etc)
- Proteins are the actuators:
  - Direct measurement of protein = best measurement
  - Many undergo post-translational modifications

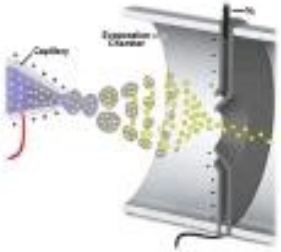
Multidisciplinary

Sample preparation

Cells, tissue

Separation

HPLC

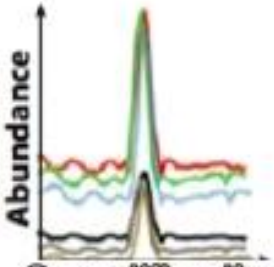


Ionisation

MALDI, ESI

Identification

TOF, Q, IT



Quantification

Algorithms

Bioanalytics

Bioinformatics

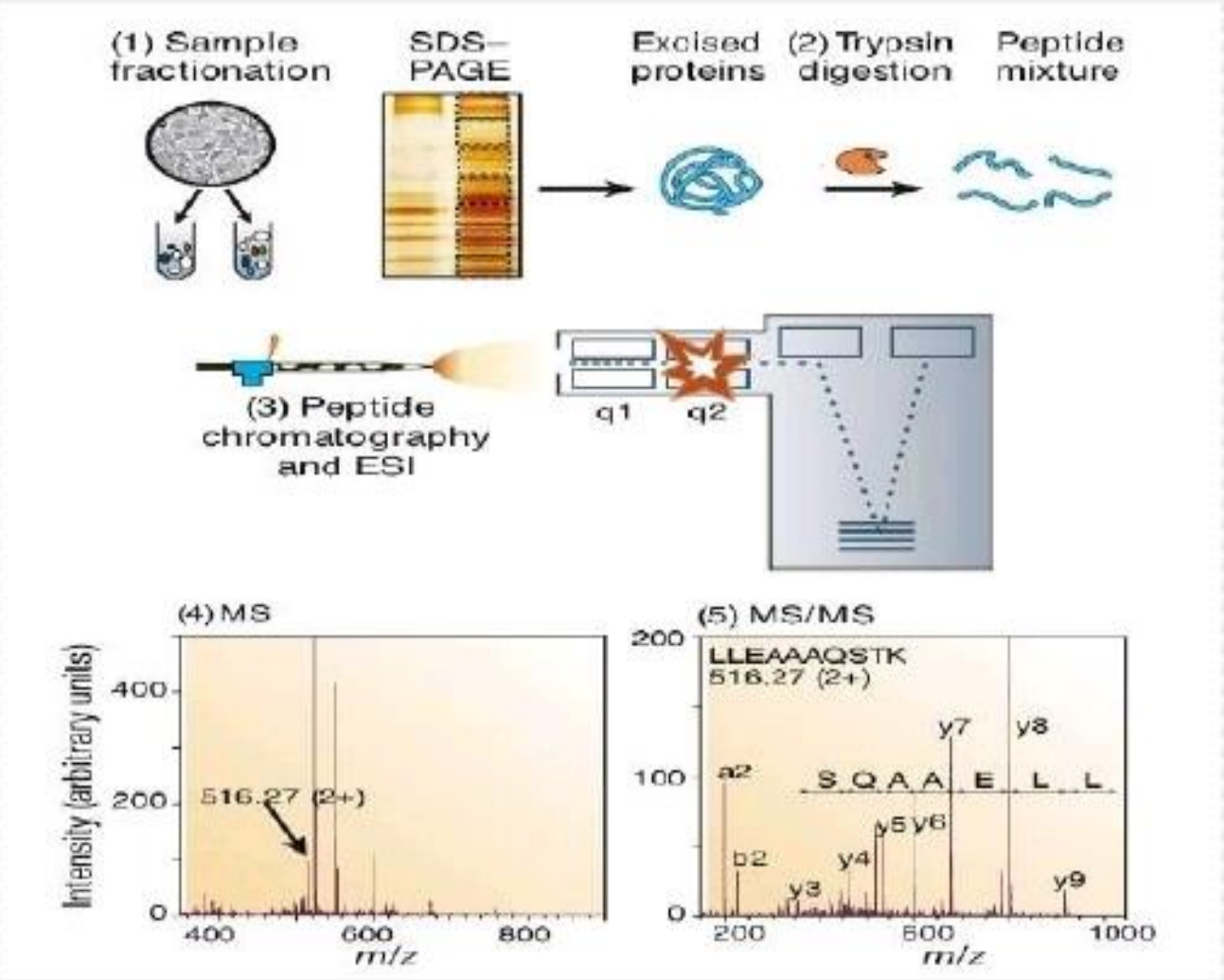
General Approaches

- Mass Spectrometry = Main Approach
  - MUDPIT
  - MS/MS
- Other Methods:
  - Protein, Chemical, Antibody Arrays
  - GFP + FRET, or other fluorescence based approaches

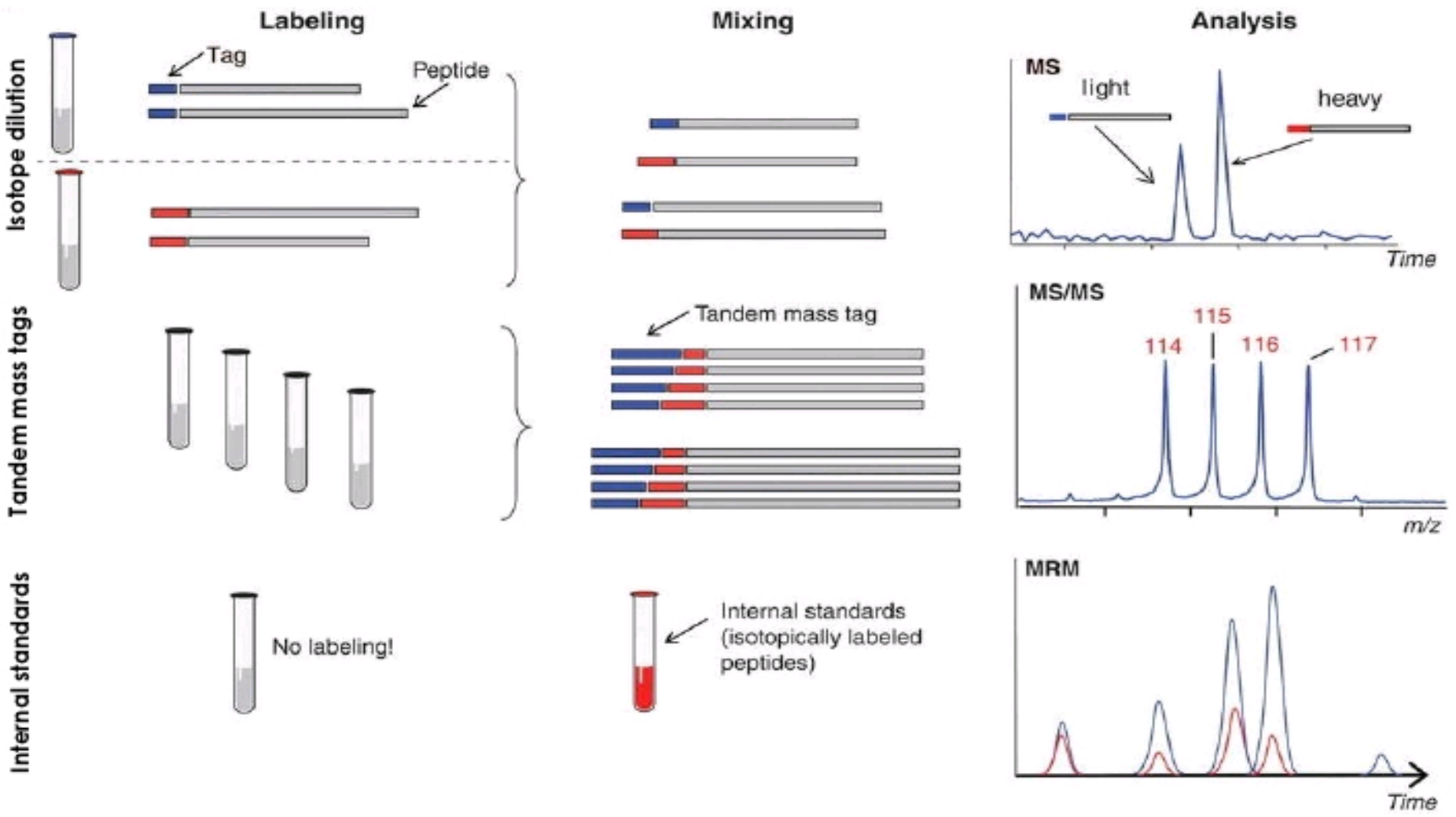
General Approaches: Mass Spec

- Step 1: Isolate cell or other protein source
- Step 2: Lyse cells and isolate proteins
- Step 3: Break up proteins into smaller (but still relatively large) amino acid chains
- Step 4: Separate chains (2D gel, gas or liquid chromatography)
- Step 5: Analyze separated protein parts by mass spectrometry

Typical MS experiment:



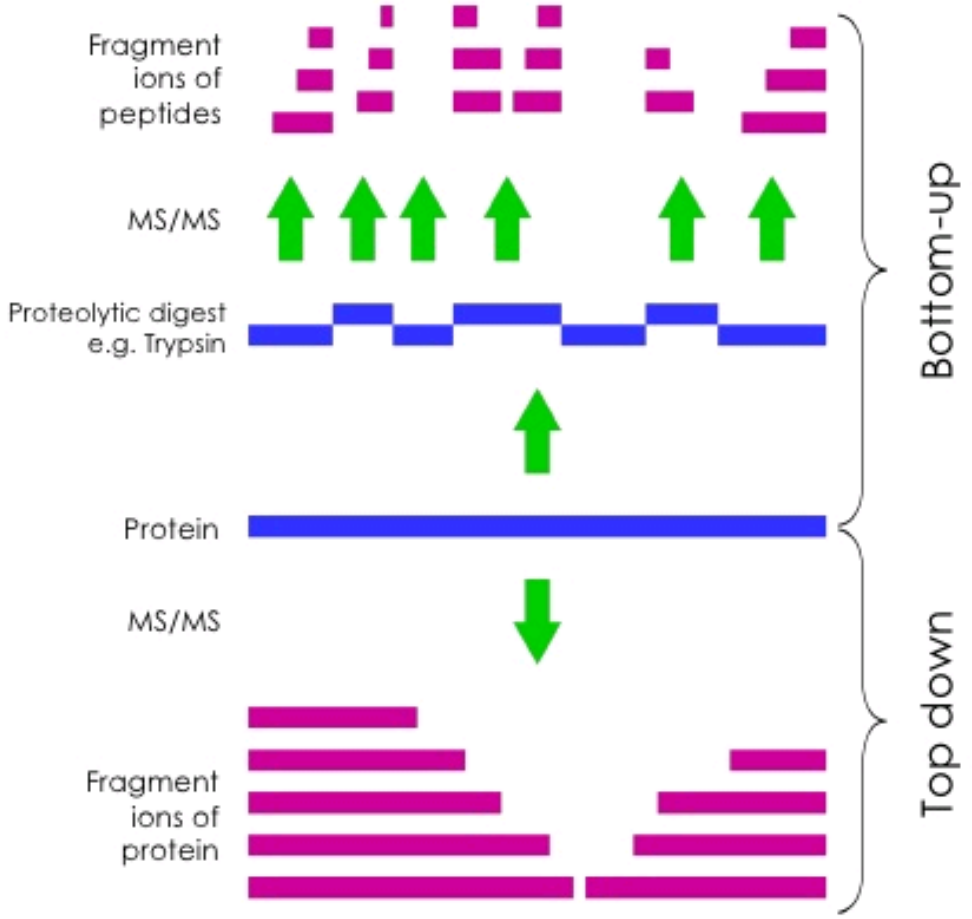
Quantification strategies



Domon B, Aebersold R. Science. 2006

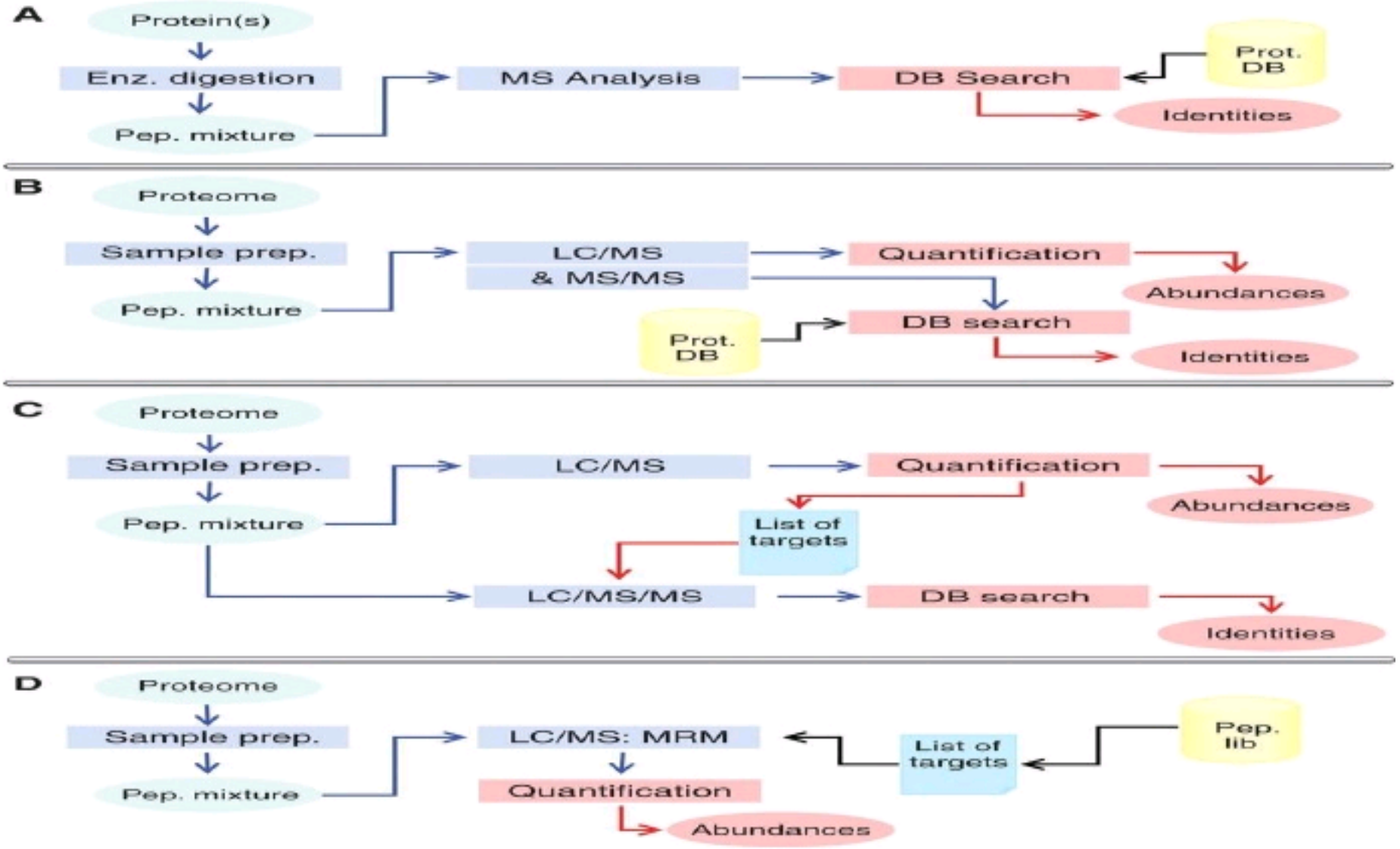
Top down or bottom up?

- Bottom-up
  - Most common
  - Starting with proteolytic fragments
  - Piecing the protein back together
    - de novo repeat detection
- Top down
  - Tandem MS of whole protein ions
    - Pulling them apart
  - Electron capture dissociation
  - Extensive sequence information



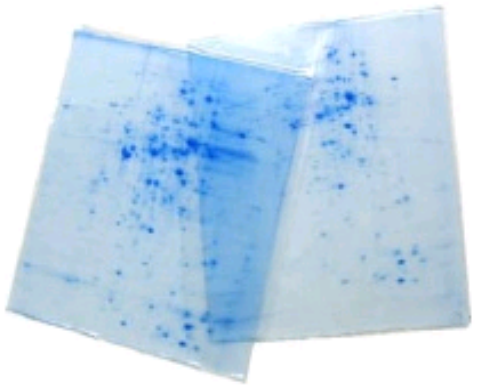
"Protein mass spectrometry" *Wikipedia, The Free Encyclopedia*. Wikimedia Foundation, Inc.

Proteomic strategies



Separation

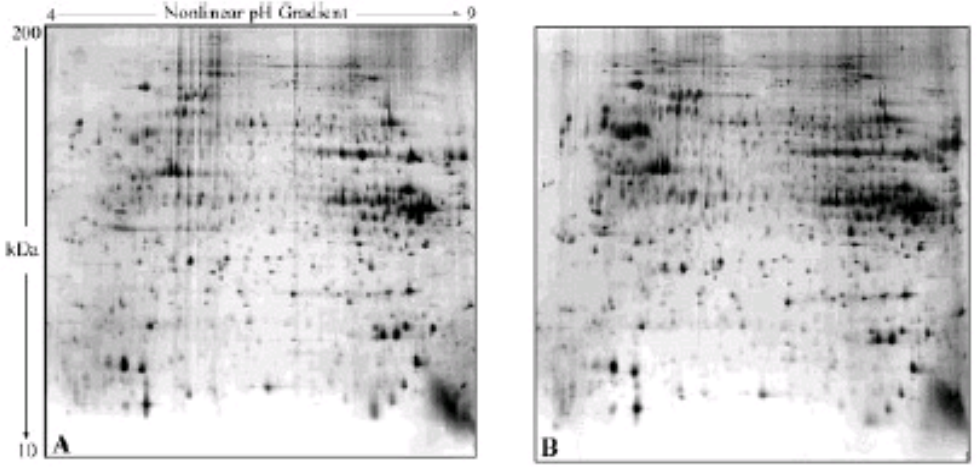
- Specific protein biophysical parameters
  - Isoelectric point
  - Molecular weight
  - Affinity
- Chromatographic methods
  - HPLC
  - 2D-HPLC
  - ProteinChips
- Electrophoretic methods
  - SDS-PAGE
  - 2-D E
- Reverse phase (RPLC) – Hydrophobicity



Yates JR, et al. Annu Rev Biomed Eng. 2009

2D Gel electrophoresis

- 1D: isoelectric focussing (IEF) separation by IP
- 2D: dimension: SDS-PAGE separation by MW
  - staining > 1000 proteins /gel
- molecular analysis by
  - MS
  - HPLC
  - Westernblot
- Pitfalls
  - very basic / acidic;
  - large / small;
  - hydrophobic;
  - low-abundance proteins



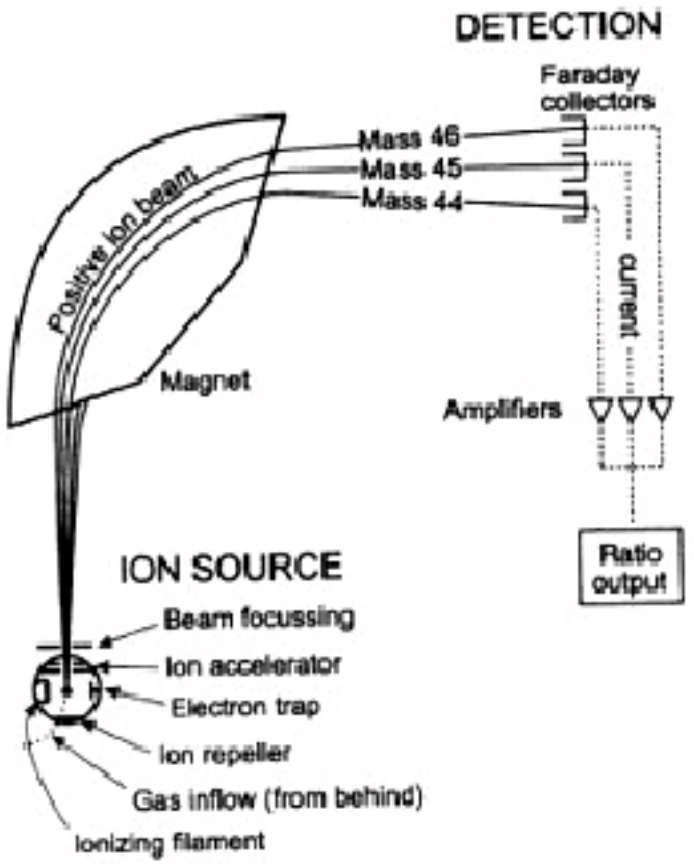
Fontana et al. Proteomics 2004

Mass Spectrometry

- Mass Spectrometry is another tool to analyze the proteome.
- In general a Mass Spectrometer consists of:
  - Ion Source
  - Mass Analyzer
  - Detector
- Mass Spectrometers are used to quantify the mass-to-charge (m/z) ratios of substances.
- From this quantification, a mass is determined, proteins are identified, and further analysis is performed.

Mass Spectrometry

- 3 Major Steps
  - Sample is ionized
  - Individual molecules are separated according to their mass/charge ratio
  - Molecules at each quantized mass/charge ratio are detected



Mass Spectrometry

- Multiple ionization sources and mass analyzers

- Mass analyzers include ion trap, time-of-flight (TOF), quadrupole, and Fourier transform ion cyclotron (FT-MS) analyzers

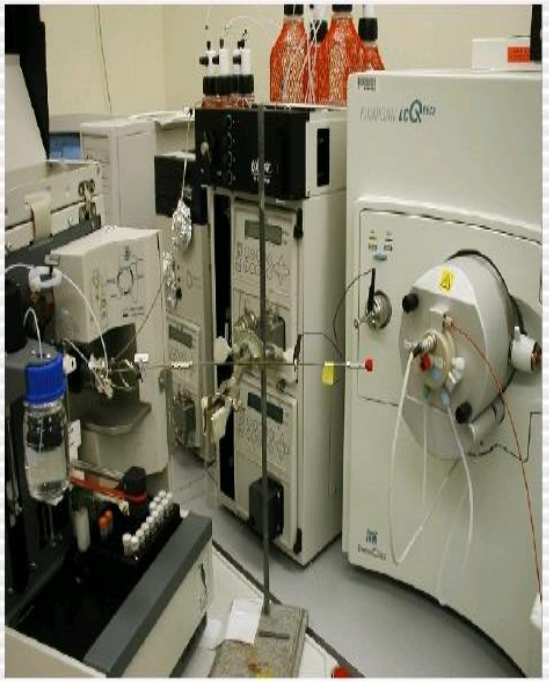
Ionization Source	Acronym	Event
Electrospray Ionization	ESI	evaporation of charged droplets
Nanoelectrospray Ionization	nano ESI	evaporation of charged droplets
Atmospheric Pressure Chemical Ionization	APCI	corona discharge and proton transfer
Matrix-assisted Laser Desorption/Ionization	MALDI	photon absorption/proton transfer
Desorption/Ionization on Silicon	DIOS	photon absorption/proton transfer
Fast Atom/Ion Bombardment	FAB	ion desorption/proton transfer
Electron Ionization	EI	electron beam/electron transfer
Chemical Ionization	CI	proton transfer

Mass Spectrometry

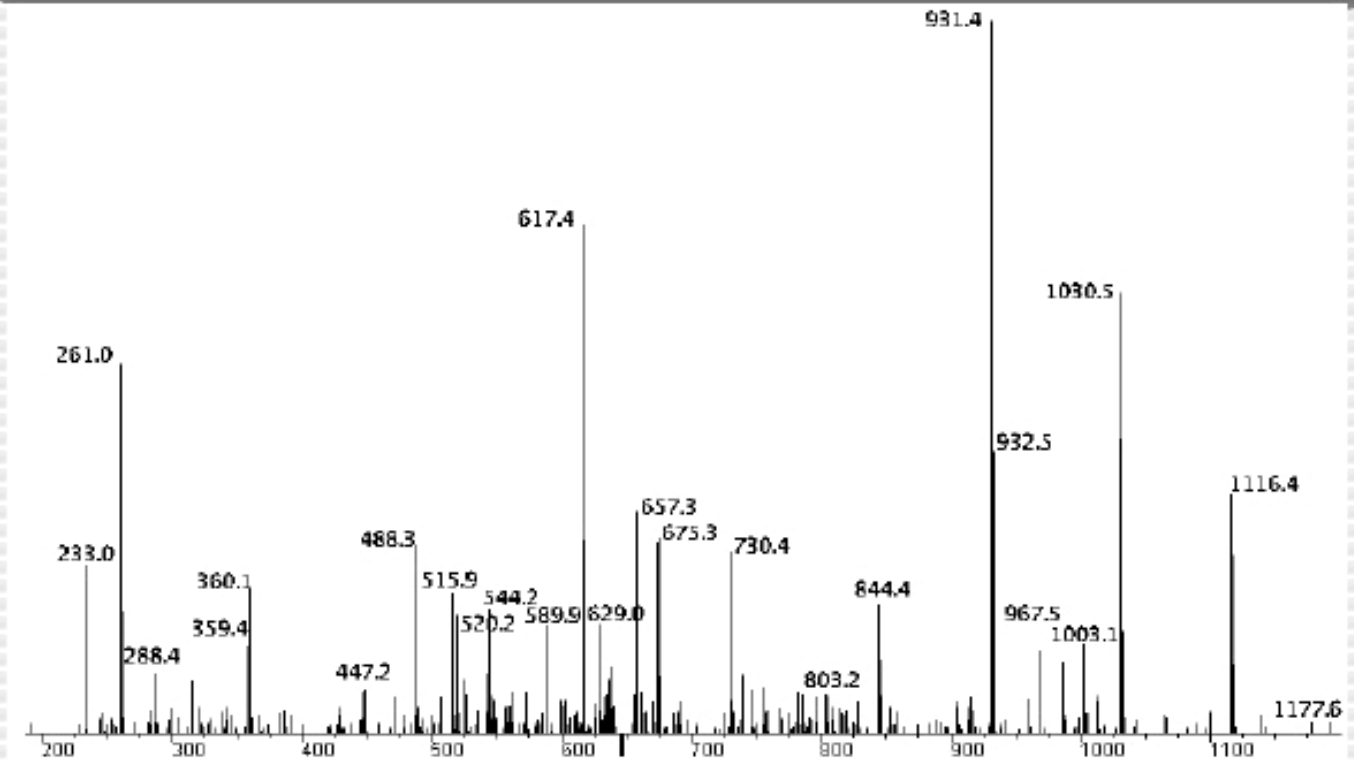
A FT-ICR mass spectrometer (**Fourier transform ion cyclotron resonance mass spectrometry**)

-Extremely High Resolving Power (discrimination of molecules with very similar charge:mass ratio)

LCQ Mass Spectrometer



Example MS/MS Spectrum



This spectrum shows the fragmentation of a peptide, which is used to determine the sequence of the peptide, via a search algorithm.

"Mass Spec" Analyses can be run in Tandem

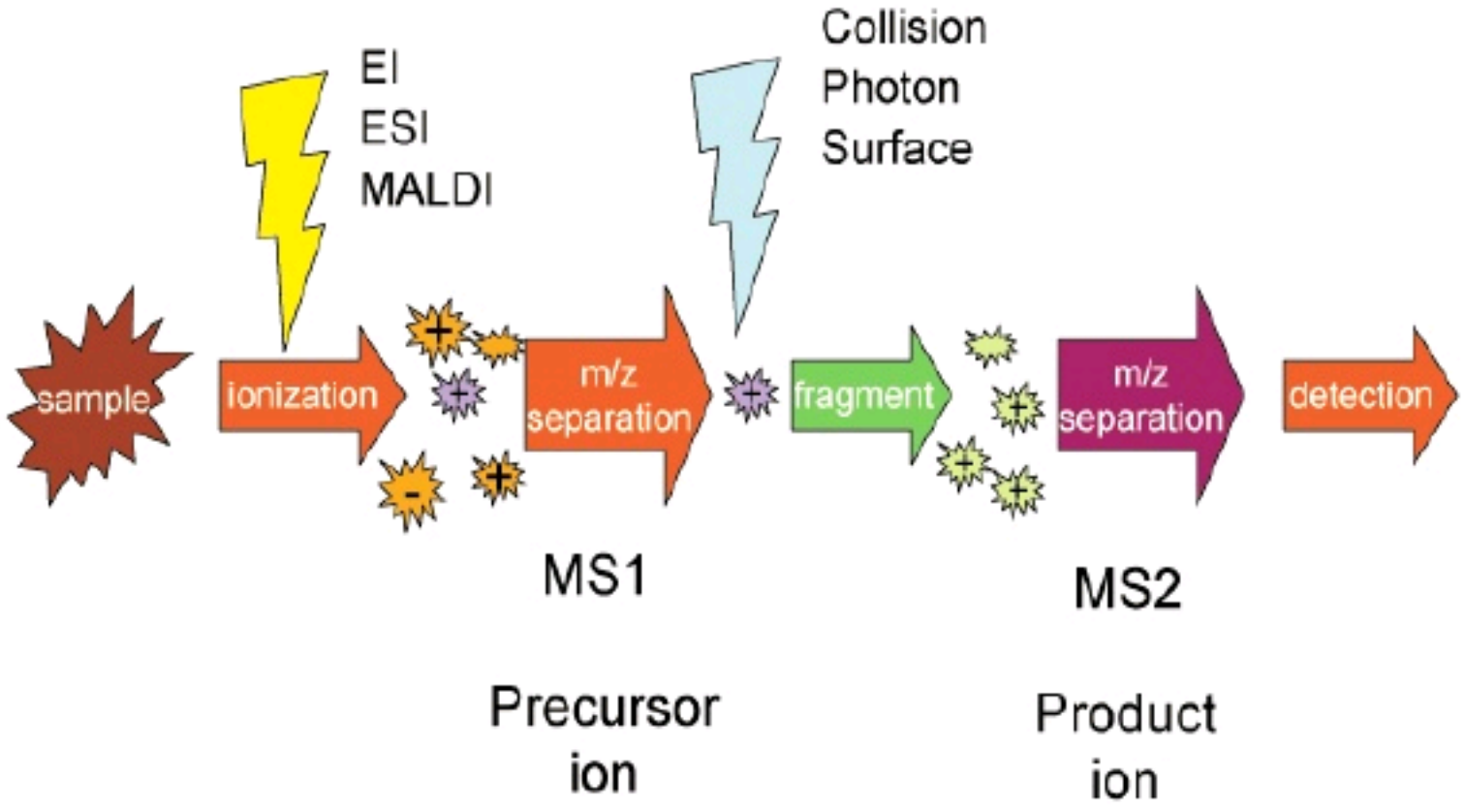
- MS/MS refers to two MS experiments performed "in tandem."
- Among other things, MS/MS allows for the determination of sequence information, usually in the form of peptides (small parts of a protein).
- This information is used by algorithms to identify a protein on the basis of mass of a constituent peptide.

Tandem Mass Spectroscopy (MS/MS)

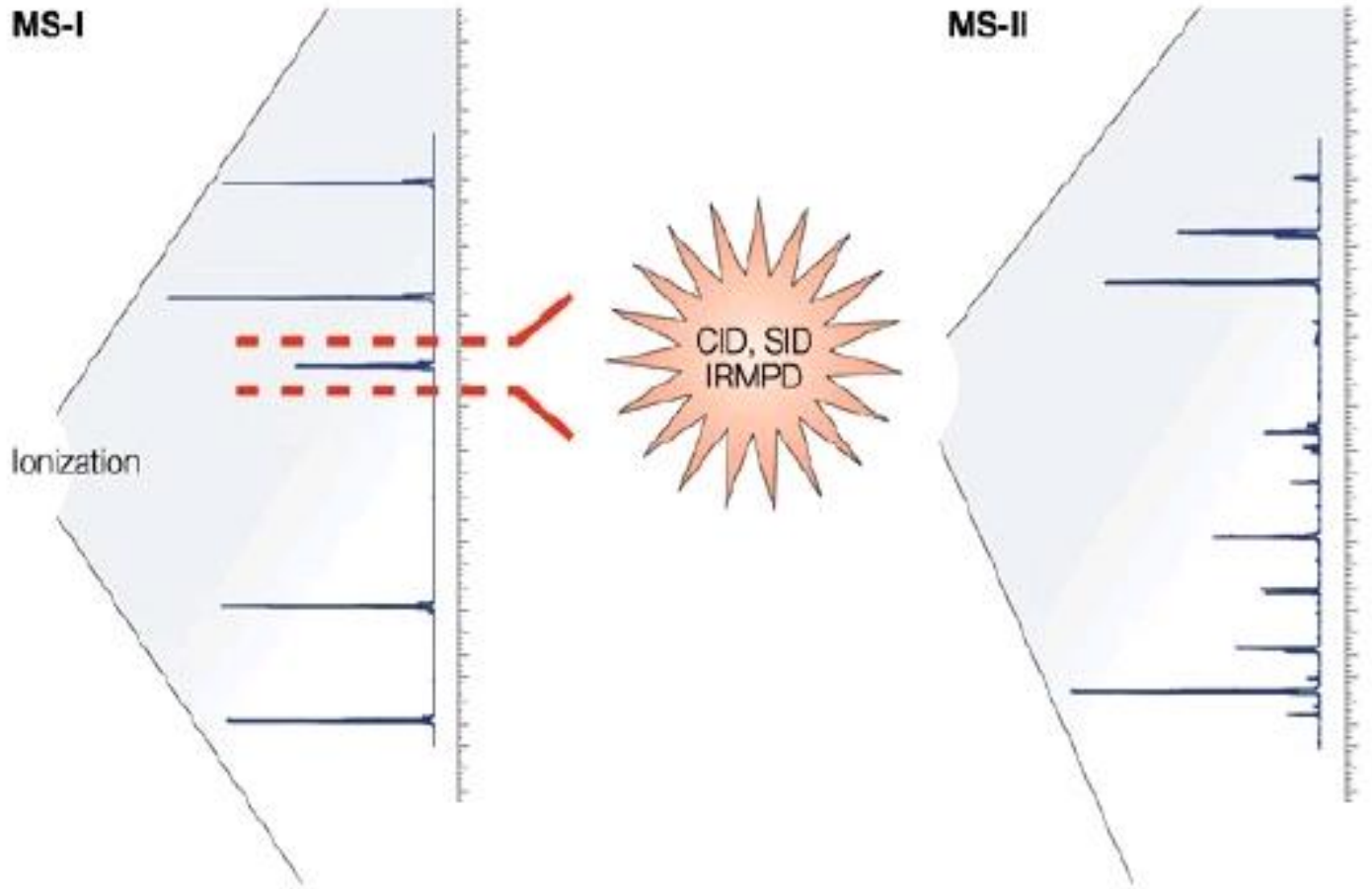
- Simple Example:
  - Two Mass Specs, (MS1, MS2)
  - A specific peak (corresponding to a specific peptide chain is identified and fragmented to form ions
  - The ions are analyzed by MS2, and identified as amino acids
  - This way each selected piece of the whole protein can be broken up and analyzed

Tandem Mass Spectroscopy (MS/MS)

- Example MS/MS flow chart



Tandem Mass Spectrometry (MS/MS)



CID, collision-induced dissociation
IRMPD, infrared multi-photon photodissociation
SID, surface-induced dissociation

Nature Reviews | Drug Discovery

MudPIT:

- Multidimensional Protein Identification Technology

- Combination of:
  - multidimensional liquid chromatography
  - tandem mass spectrometry
  - database-searching algorithms
    - Relies heavily on data-base search engines and other bioinformatics tools to interpret/generate the data

High-Throughput Proteomics: MudPIT

Plasmodium falciparum
[Sporozoites, Trophozoites, Merozoites, Gametocytes]

Lysis

Proteins

Digestion

Peptide Mixture

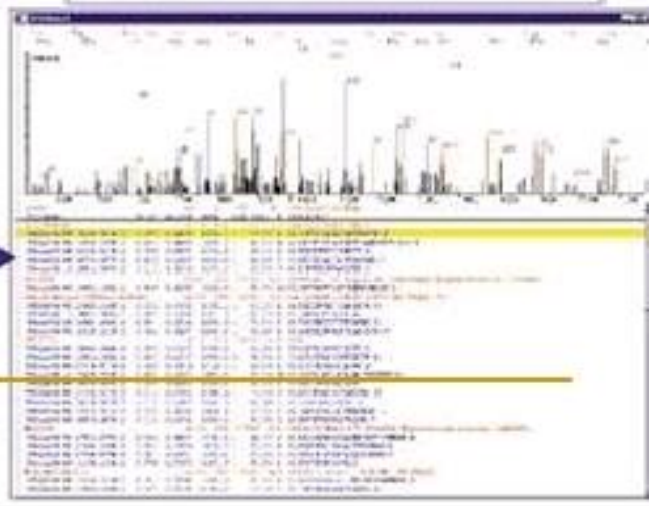
Tandem Mass Spectrometer

2D Chromatography

RP

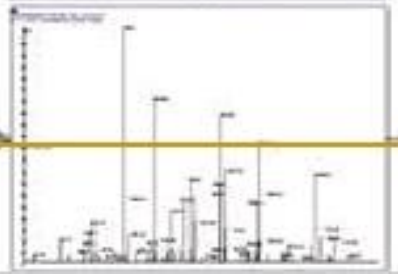
SCX

> 1,000 Proteins Identified



SEQUEST®
DTASelect & Contrast

MS/MS Spectrum



MudPIT Strengths

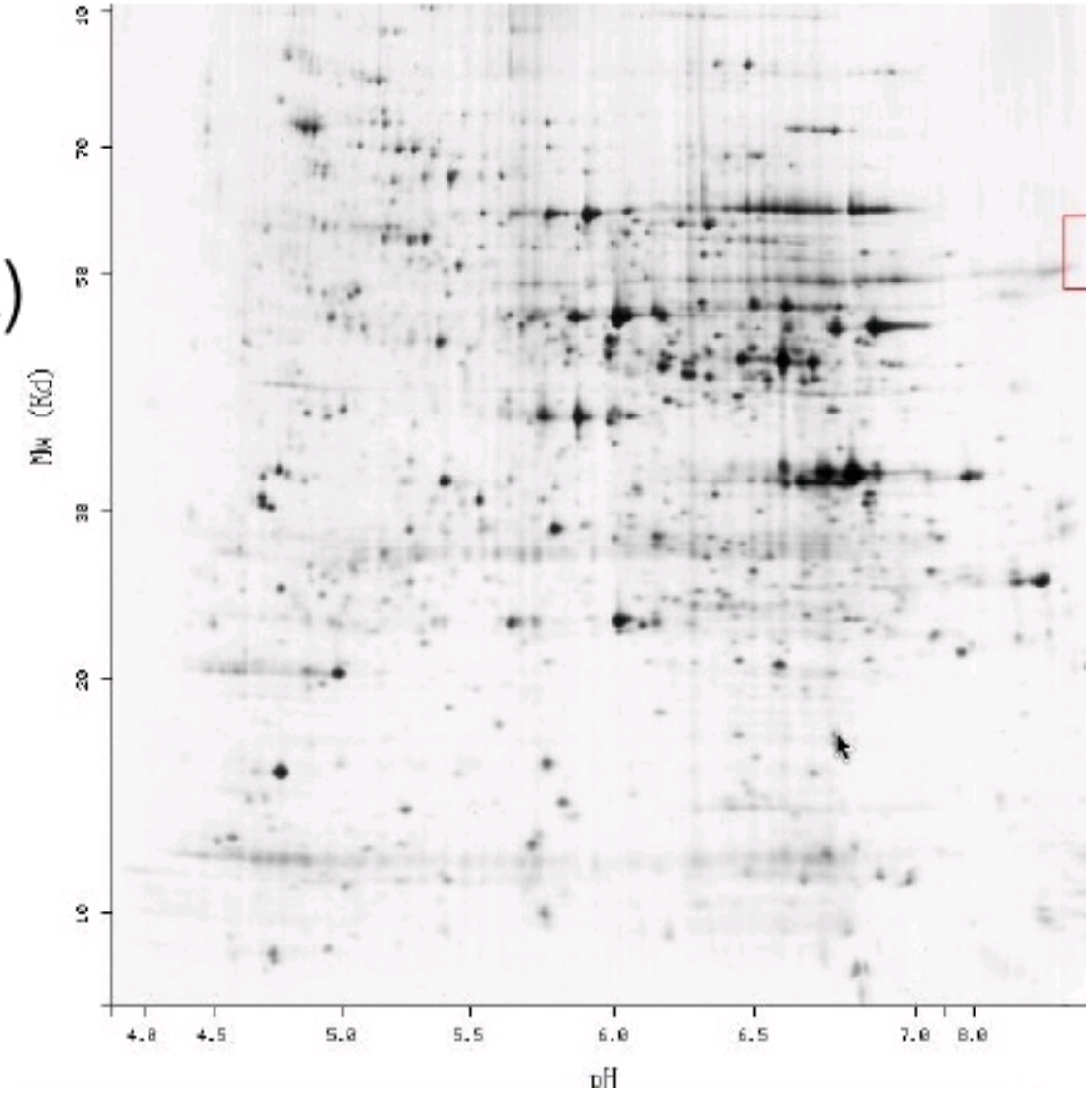
- High throughput
- Able to differentiate thousands of different proteins
- Statistical methods allow rapid identification of peptide chains when compared to existing databases
- No need to separate the proteins from a gel before running the Mass Spec (LC vs Gel)

Other Separation Techniques: 2D Gels

- Proteins are first separated according to isoelectric point
  - pH gradient is applied (usually horizontally)
  - Each protein is charged except at it's isoelectric point
- Proteins are then denatured in sodium dodecyl sulfate (SDS)
  - Unfolds them into straight molecules
  - Binds SDS molecules roughly proportional to the length of the denatured protein
  - Electric current then separates the proteins according to mass, similar to a regular agarose gel

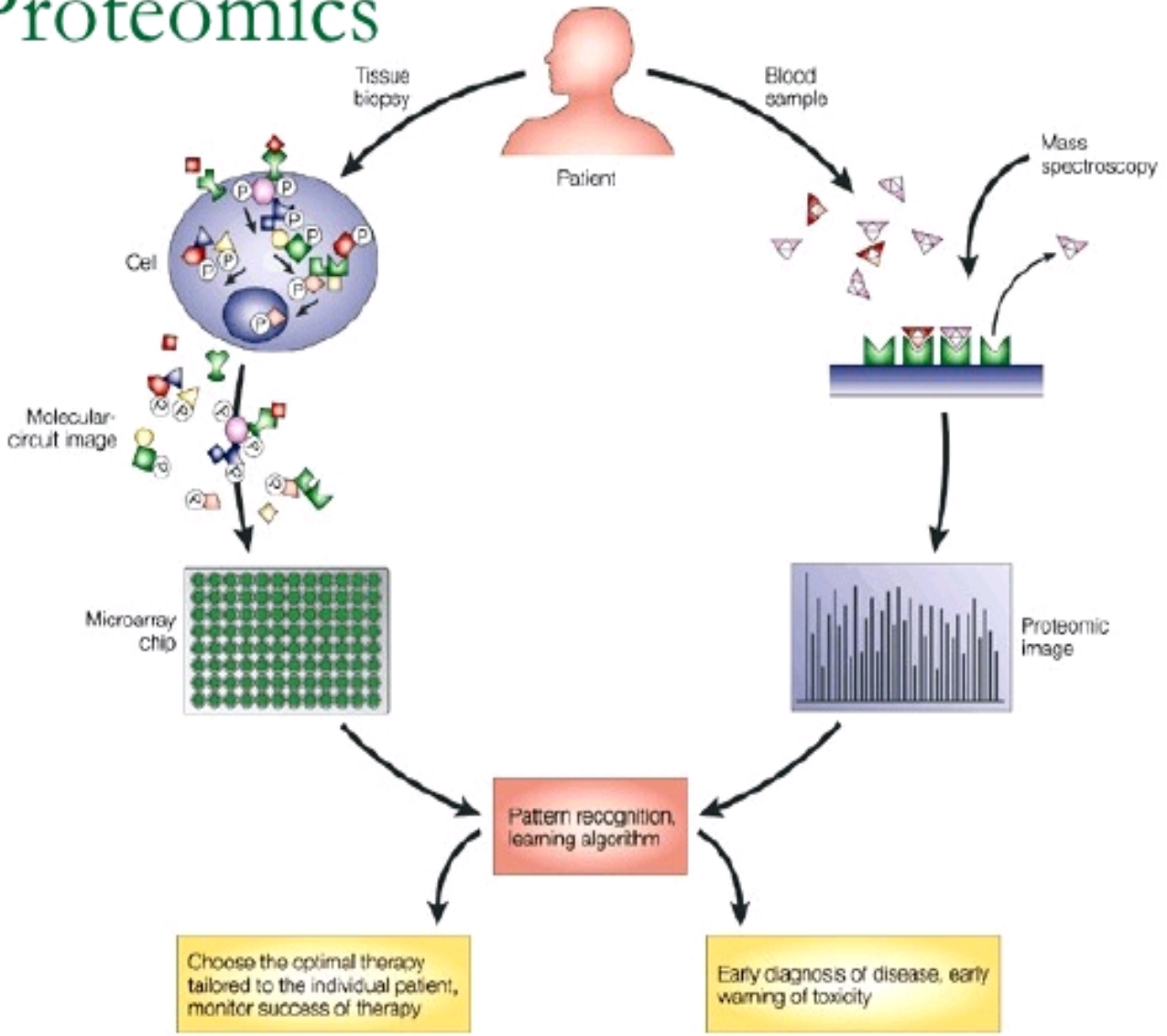
Sample 2D Gel

- X-axis = pH (Isoelectric Point)
- Y-axis = Kilo Daltons (Mass)



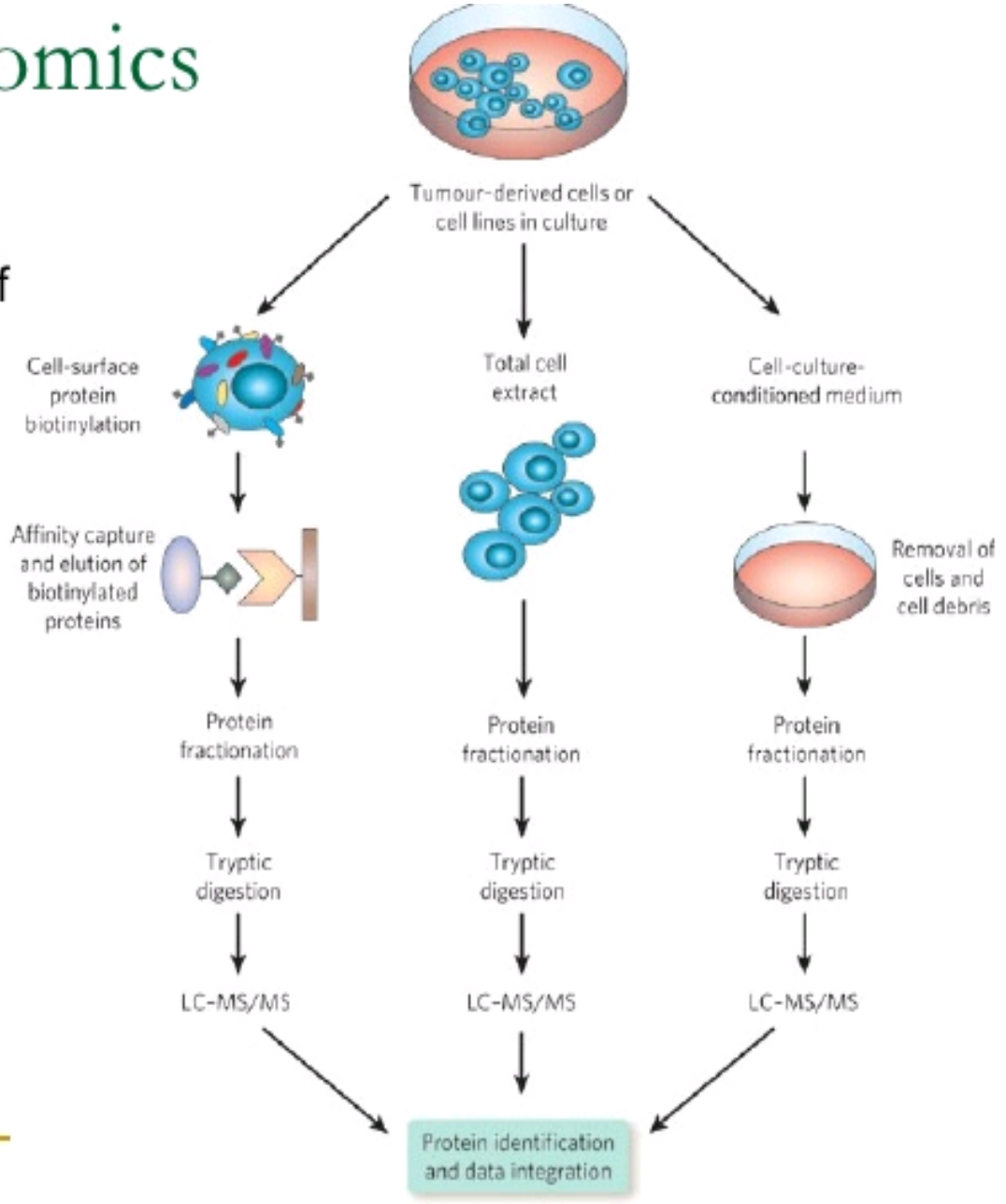
Clinical Proteomics

- Proteome Profiling Technologies



Clinical Proteomics

- Analyze the proteome of both diseased and healthy cells
- Find changes in:
  - Cell or tissues
  - Subcellular structures
  - Protein complexes
  - Biological fluids



Clinical Proteomics: Goals

- Develop new biomarkers for disease diagnosis and early detection
- Identify new targets for drugs
- Better evaluate the therapeutic effect of possible drugs

Summary

- Proteomics = study of full complement of proteins, including modifications
- Most current approaches employ advanced mass spectrometry techniques to separate and identify amino-acid chains
- Huge potential in clinical applications, as well as basic research