Modern Proteomics: Locating the targets in drug action

Stephen Barnes, Ph.D.

Department of Pharmacology & Toxicology

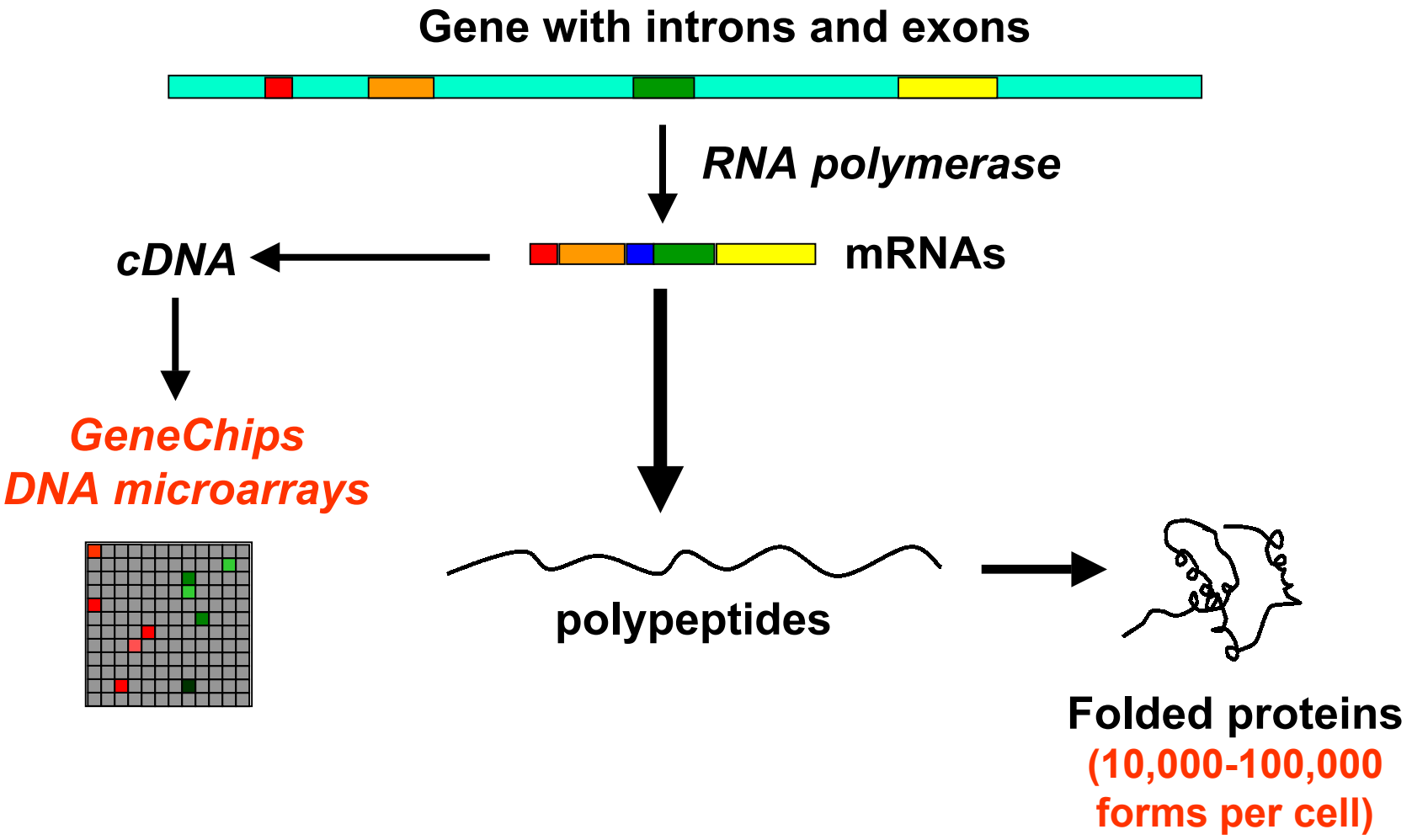
Senior Scientist,

Comprehensive Cancer Center Mass Spectrometry Shared Facility

Outline of talk

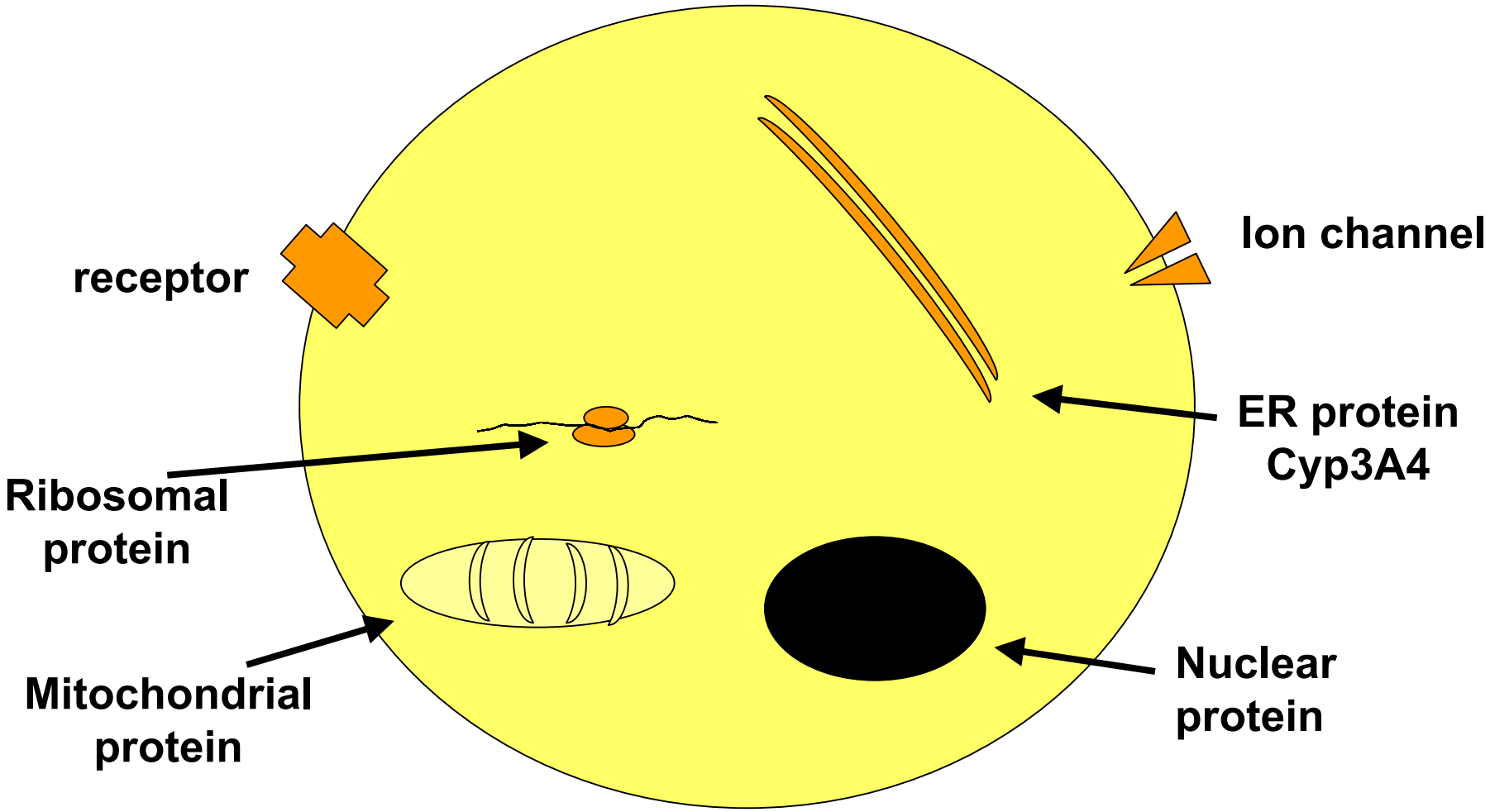
- Reminder about proteins in biochemistry
- The collapsing gene-disease-drug paradigm
- Protein networks
- What is 2D-proteomics?
- Peptide fingerprinting and protein identification
- Doing MALDI-TOF experiments and bioinformatics
- The Q-tof and peptide sequencing
- Realities, the future and closing remarks

Basics of biochemistry

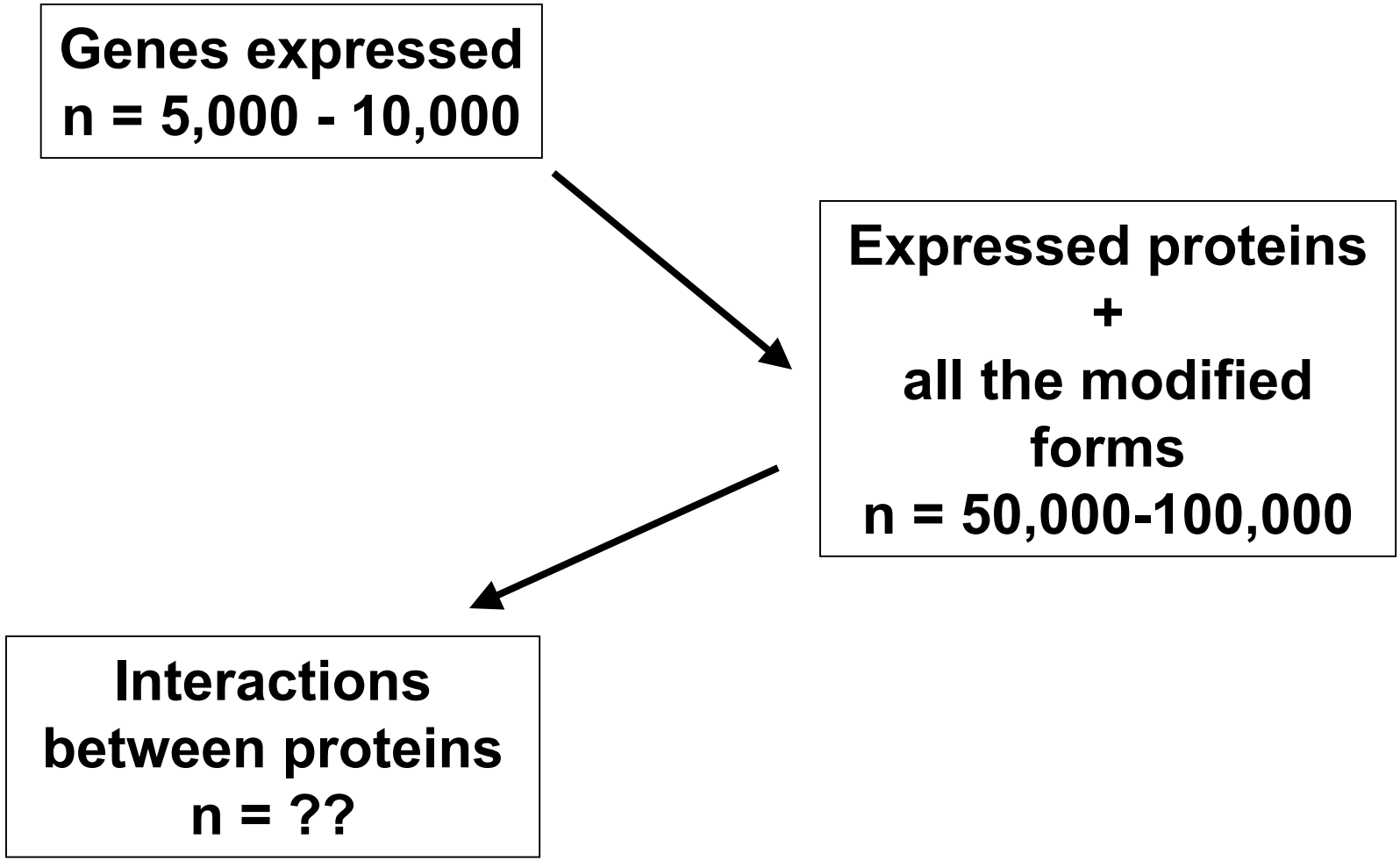


Proteins in a cell

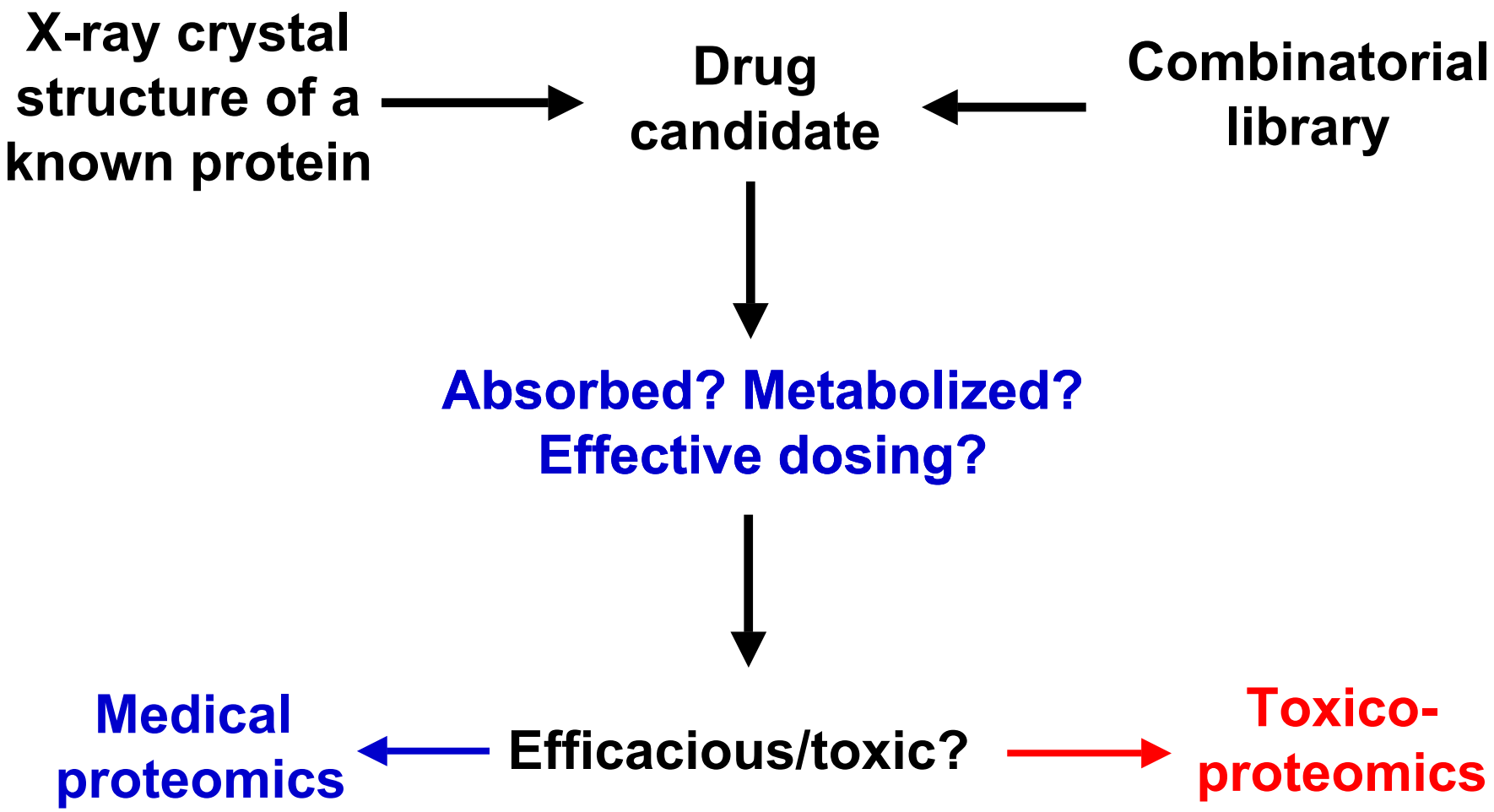
or *excuse me, I'm just a target*



The proteomics numbers challenge



Building a drug and rediscovering how it works



Collapse of the single target paradigm

Old paradigm

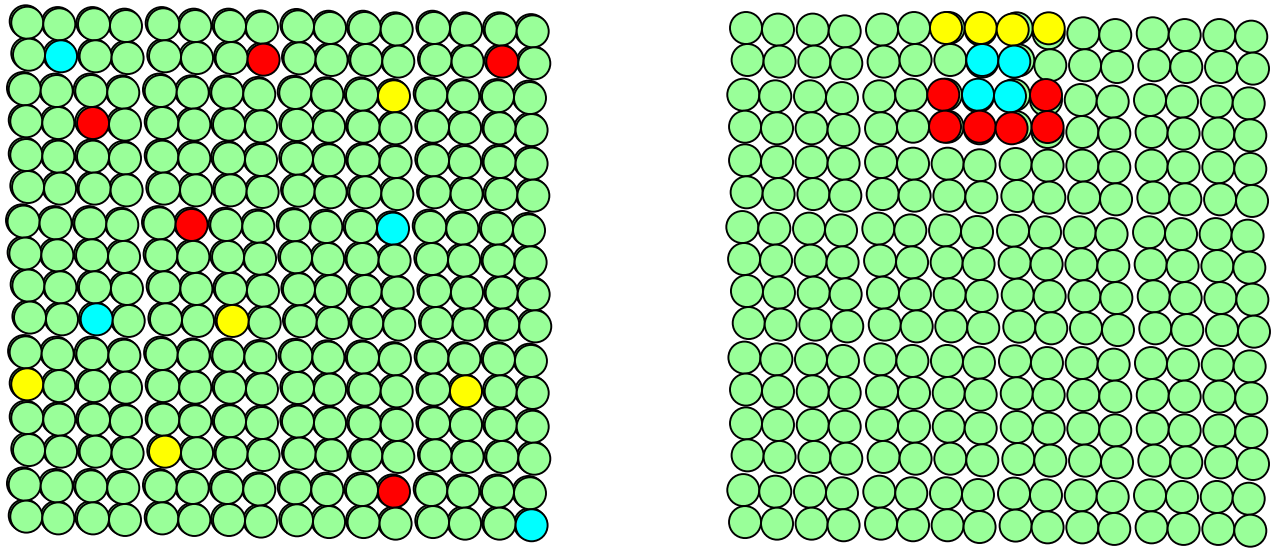
Diseases are due to single genes - by knocking out the gene, or designing specific inhibitors to its protein, disease can be cured

But the gene KO mouse didn't notice the loss of the gene

New paradigm

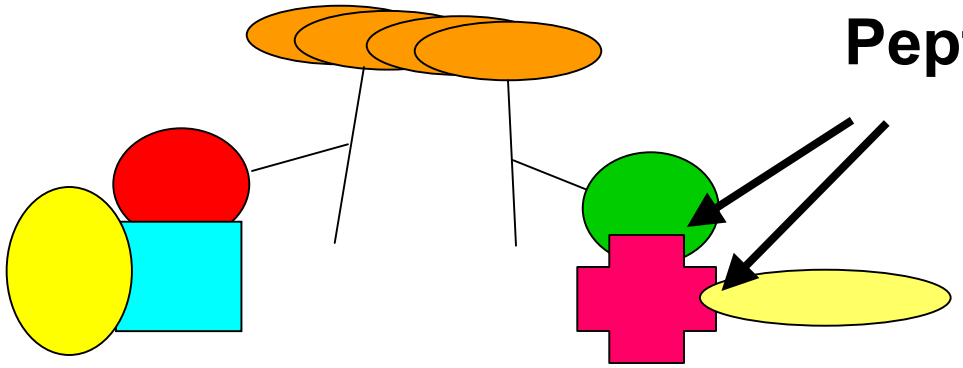
We have to understand gene and protein networks - proteins don't act alone - effective systems have built in redundancy

Proteins aren't random in cells



So, who's binding to whom?

Proteins (and spies) don't act alone



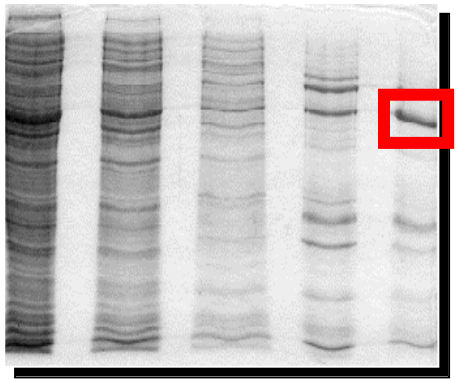
Signal transduction complex lying in anticipation

http://www.bind.ca

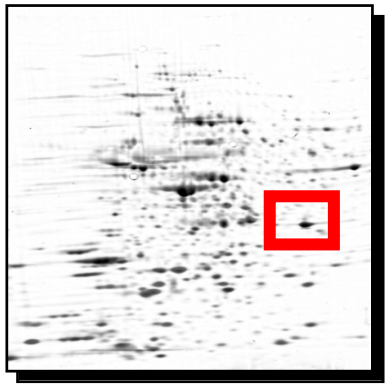
How to discover protein brotherhoods

Old method:
Yeast 2-hybrid screen

New method:
Recover protein complexes



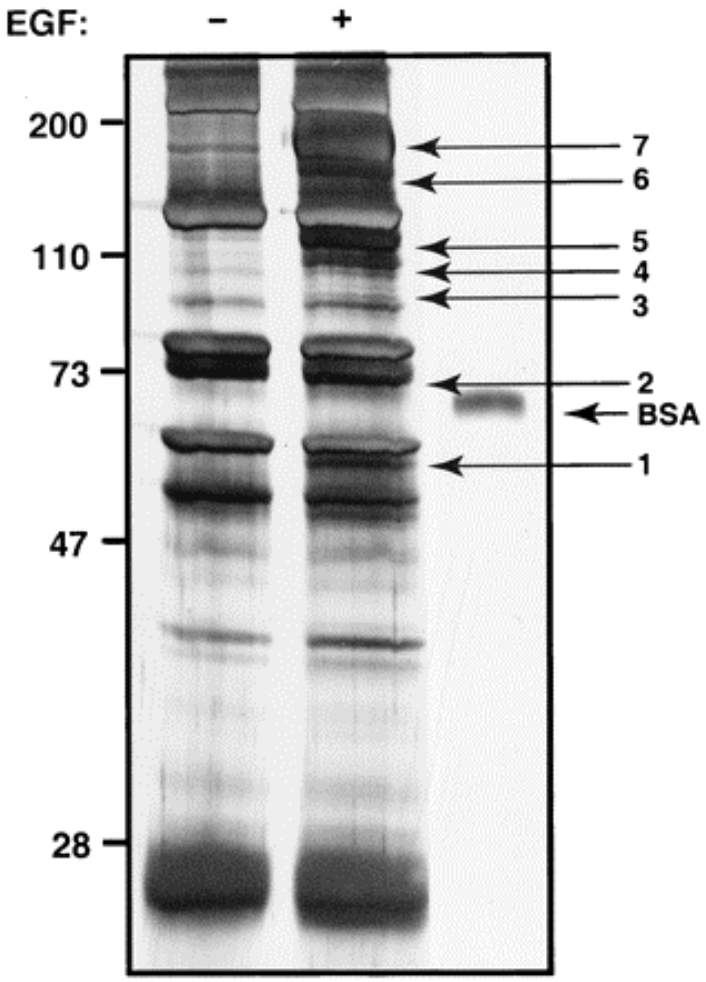
SDS-PAGE



IEF/SDS-PAGE

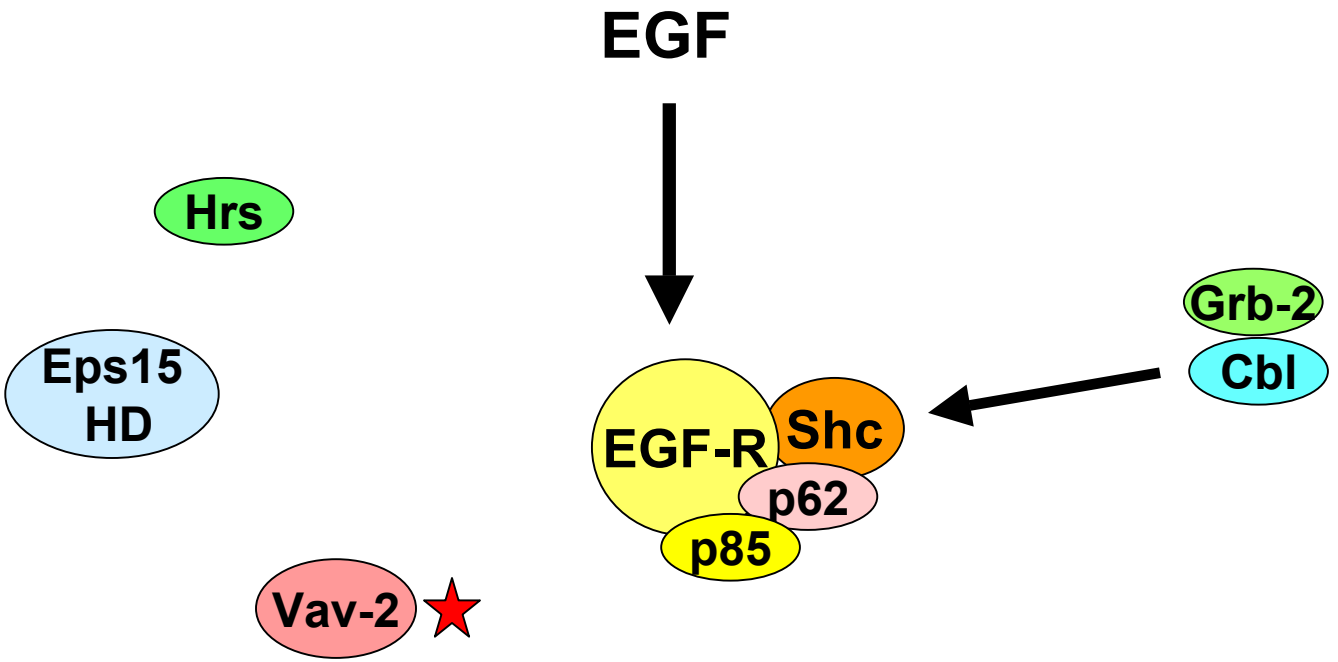
Affinity isolation of EGF-responsive proteins

Pandey et al., PNAS 97: 179-184 (2000)



EGF-induced tyrosine phosphorylation in HeLa cells. ***Serum-deprived HeLa S3 cells (5 x 10^9) were either left untreated or treated with 1 μg/ml EGF for 5 min. Cleared cell lysates were immunoprecipitated with a mixture of monoclonal anti-phosphotyrosine antibodies, washed, and resolved by SDS/PAGE. The gel was then silver-stained. Numbers indicate the positions of the bands that were excised for enzymatic digestion by trypsin and subsequent mass spectrometric analysis.***

EGF-stimulated, tyrosine-phosphorylated proteins identified by mass spec



See protein interactions at www.bind.ca

Limitations of proteomics

- **Unlike its mRNA counterpart, proteomics doesn't have a PCR equivalent**
- **It's limited by Avogadro's number - 1 fmol is 6 x 10^8 molecular ions**
- **If a cell has a 100 copies of a protein, then at a minimum you need to have 6 x 10^6 cells - in reality, you need much more**
- **Fourier Transform-MS using trapped ions may be the solution**
- **Beware of Western blots**

What are the proteomics methods?

I. **2-dimensional electrophoresis (2-DE) & mass spectrometry**

II. **Liquid chromatography & mass spectrometry (LC-MS or LC/LC MS/MS)**

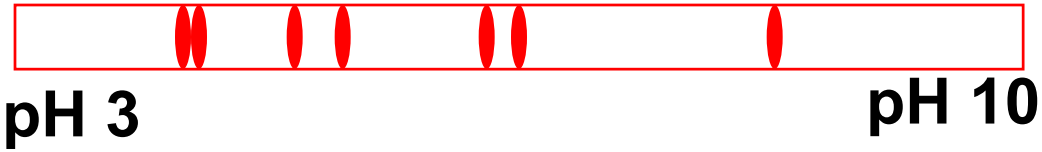
III. **"Chip" technology:**

2-D array of recombinant polypeptides or antibodies on a single microscope slide; the entire chip is probed with a labelled "ligand" (protein, lipid, drug); SELDI-TOF analysis

What 2-D electrophoresis involves:

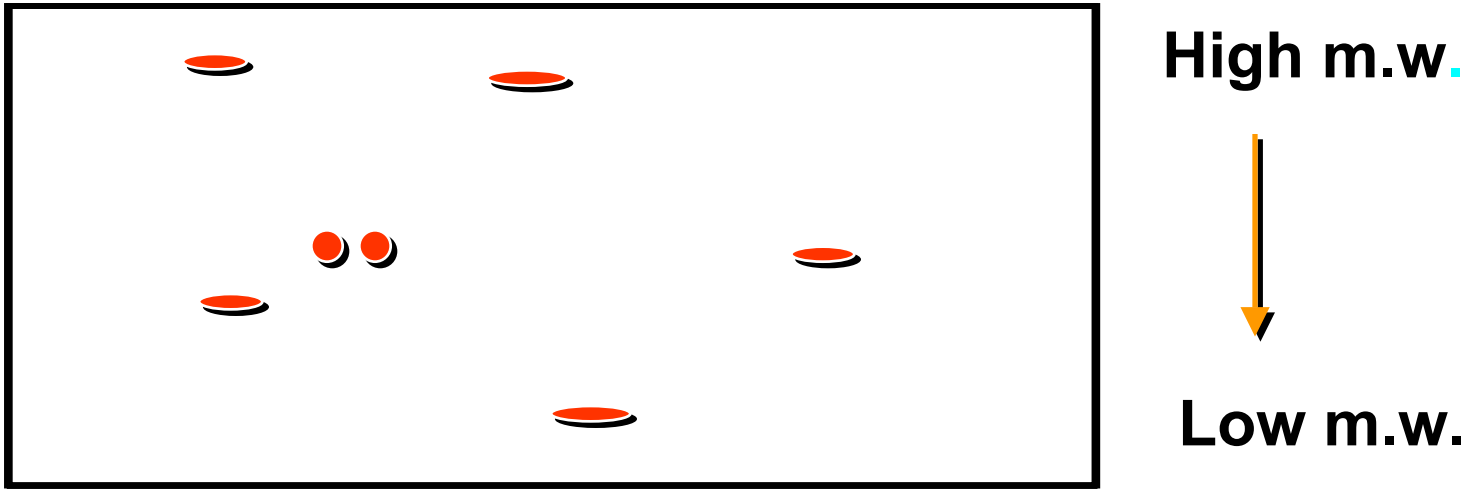
- 1st dimension: **Isoelectric focussing**

(separation according to charge)



- 2nd dimension: **(SDS)-PAGE**

(separation according to size)

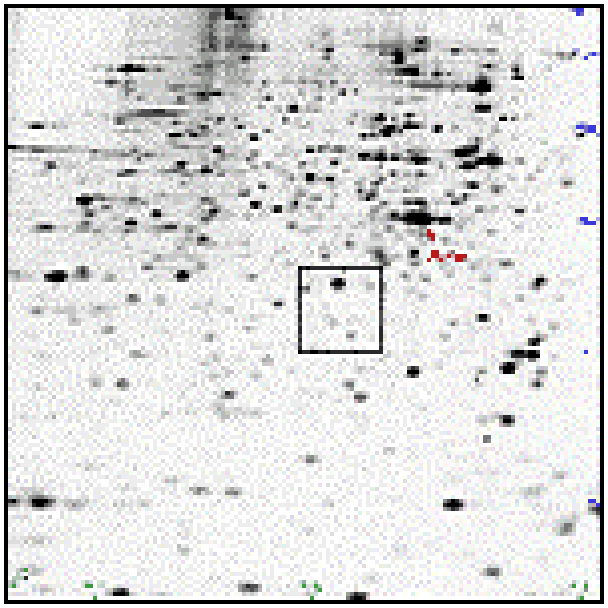


Separating proteins in two dimensions

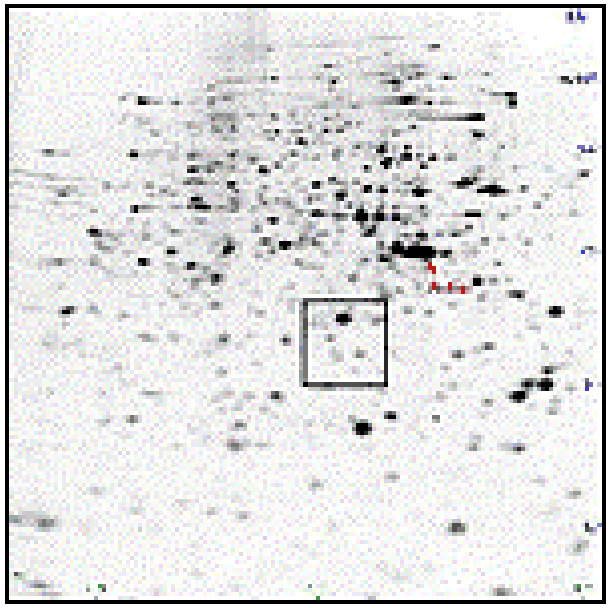
SDS-PAGE

IEF

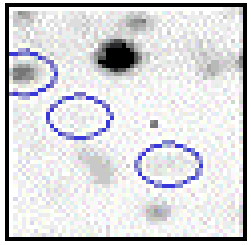
Experimental Control:
Full Image



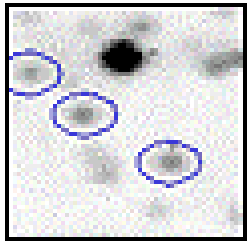
Experimental Result:
Full Image



Experimental
Control:
Region Outlined



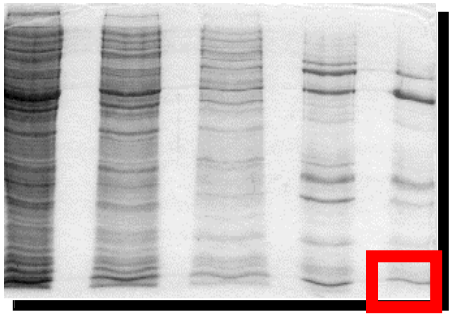
Experimental
Result:
Region Outlined

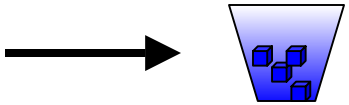


Find this and other 2-D gels at http:www.expasy.org

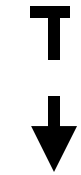
Basic protein mass spectrometry:

From a band on a gel to a "tryptic peptide fingerprint"

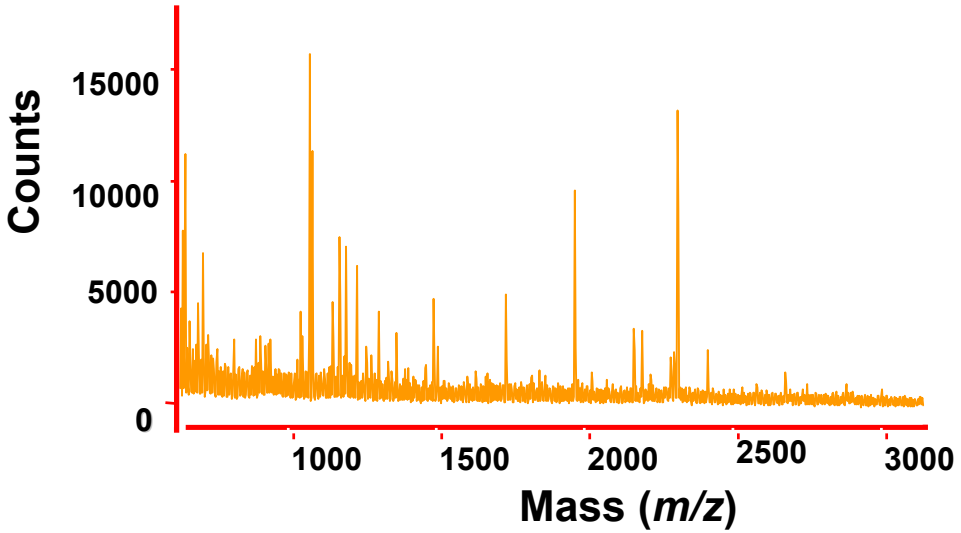




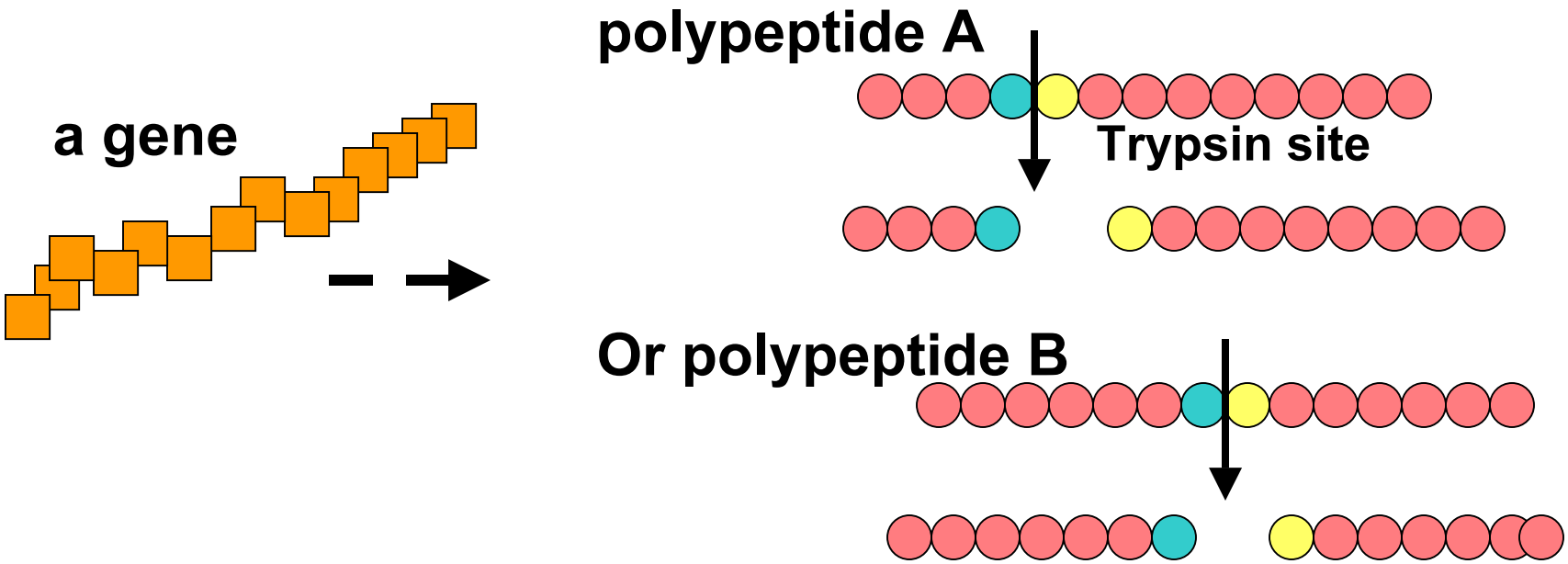
trypsin



MALDI-TOF mass spectrometry



Proteolysis generates sets of peptides that are a “fingerprint” for that polypeptide

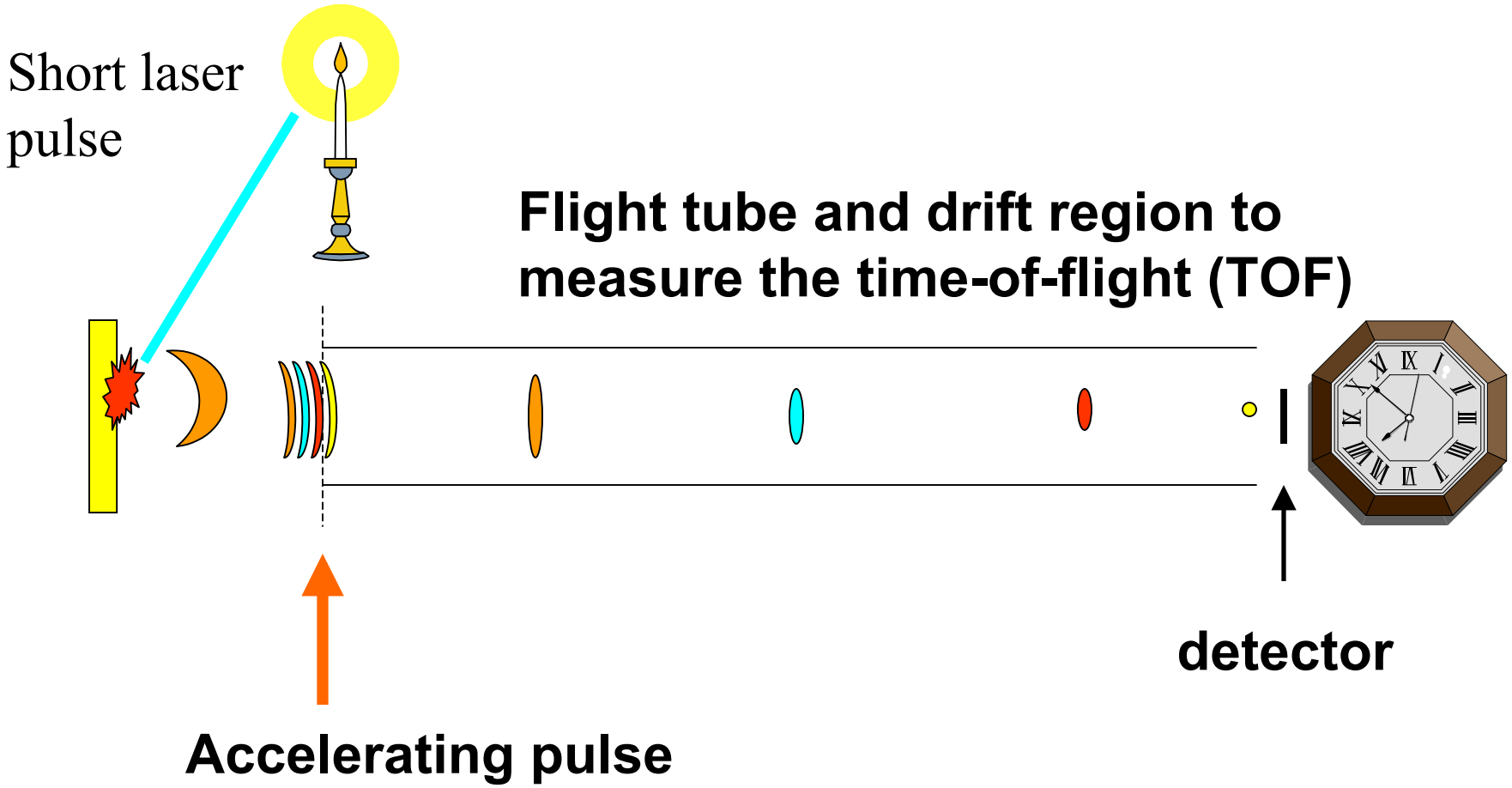


So, polypeptides of identical mass can have nonidentical tryptic “fingerprints.”

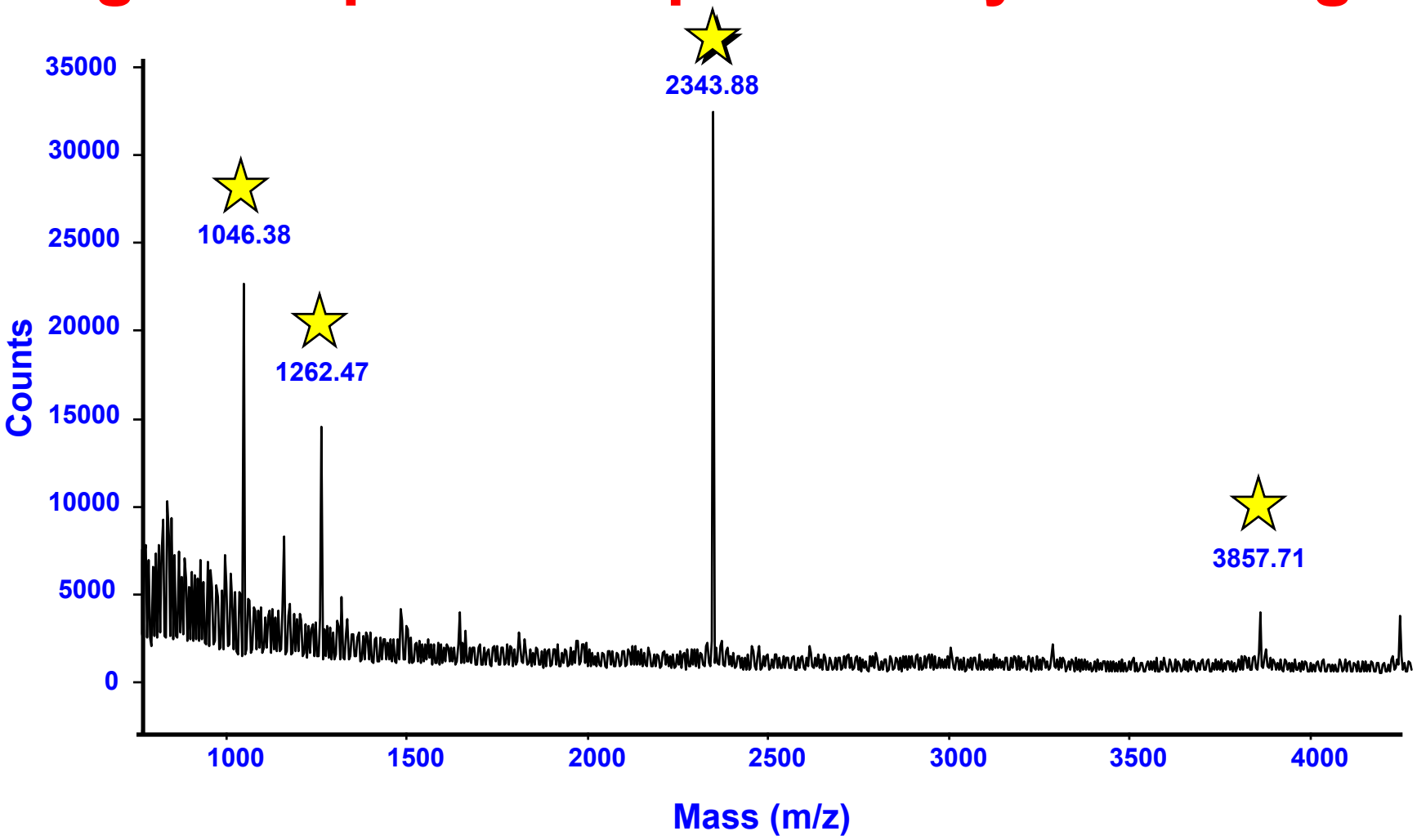
From Proteins to Sequence Tags

- If each protein (average 500 residues) had a cleavage site every 10 residues, then about 1.5-3.0 million peptides describe the expressed products of the human genome
- Each peptide has a molecular weight value that is its individual sequence tag
- Any modification will increase the peptide's molecular weight, e.g., a nitro group adds 45

Matrix-Assisted Laser Desorption Ionization (MALDI)



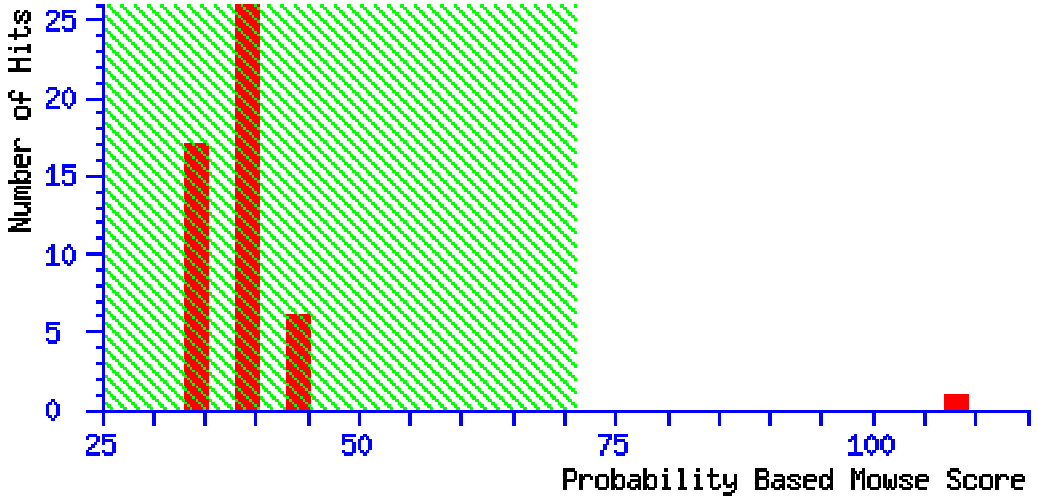
MALDI-TOF mass spectrum of tryptic digest of p22 band purified by 6xHis-tag



Probability Based Mowse Score

Score is -10*Log(P), where P is the probability that the observed match is a random event.

Protein scores greater than 71 are significant (p<0.05).



Accession	Mass	Score	Description
1. gi|548939	20840	108	FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE SLYD (PPIASE) (ROTAMA
2. gi|13384624	46931	45	myocyte enhancer factor 2C [Mus musculus]
3. gi|5257384	43424	44	(AF137308) phytochrome B [Lolium perenne]
4. gi|4505147	50305	44	MADS box transcription enhancer factor 2, polypeptide C (myocyte enhan
5. gi|1515365	44552	43	(U52596) nucleocapsid protein [Avian infectious bronchitis virus]
6. gi|6093850	49443	42	PRESENILIN 2 (PS-2)
7. gi|15225198	47999	42	hypothetical protein [Arabidopsis thaliana]
8. gi|113854	58376	41	NITROGENASE IRON-IRON PROTEIN ALPHA CHAIN (NITROGENASE COMPONENT I)
9. gi|13928425	13831	40	(AB040419) envelope protein [Bovine immunodeficiency virus]
10. gi|4389228	56064	40	Chain Z, Crystal Structure Of The Complex Between Escherichia Coli Glycerol

E. coli: FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE

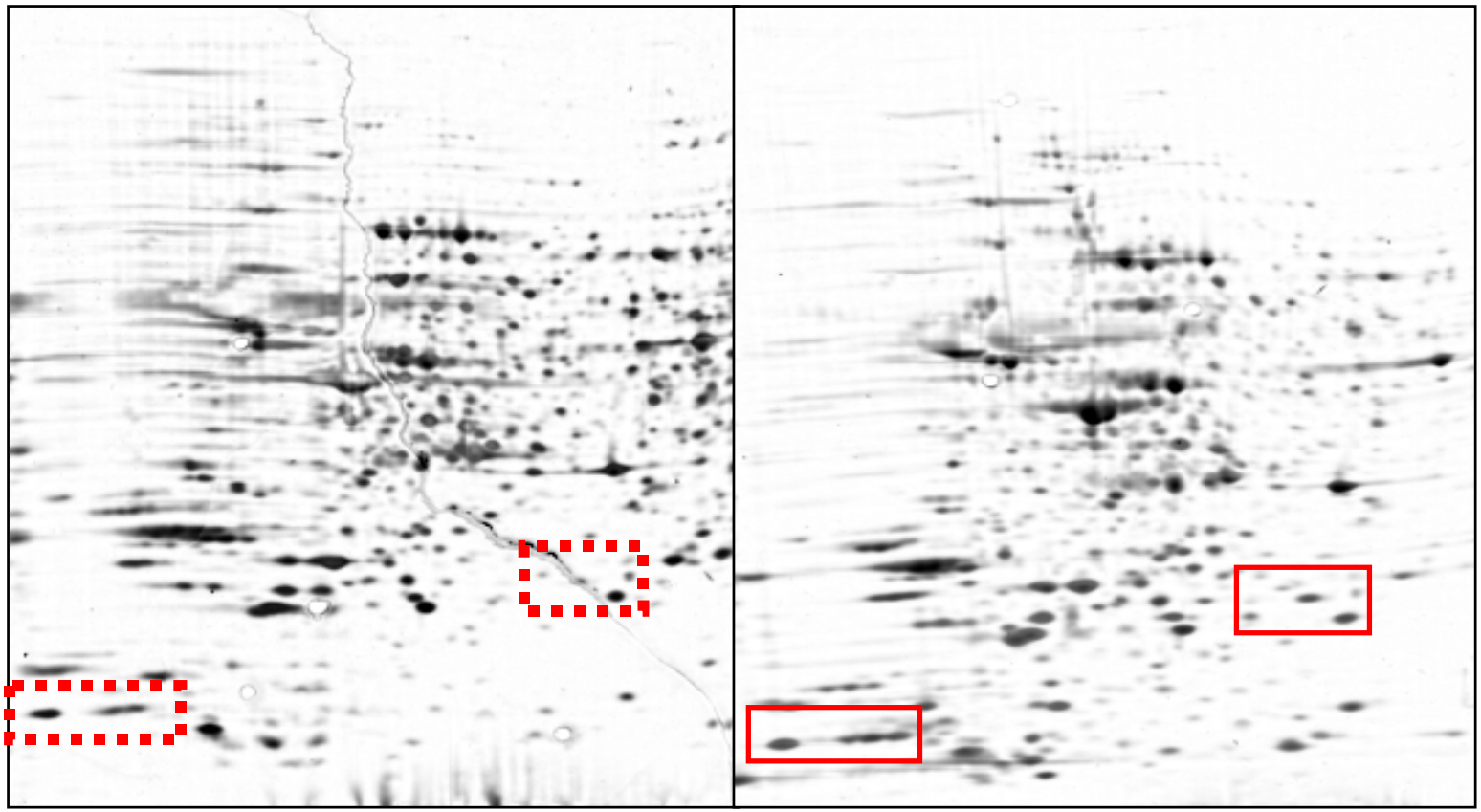
Nominal mass of protein (Mr): 20840

```
  1 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS
 41 LISGLETALE GHEVGDKFDV AVGANDAYGQ YDENLVQRVP
 81 KDVFMGVDEL QVGMRFLAET DQGPVPVEIT AVEDDHVVVD
121 GNHMLAGQNL KFNVEVVAIR EATEEELAHG HVHGAHDHHH
161 DHDHDGCCGG HGHDHGHEHG GEGCCGGKGN GGCGCH
```

Tryptic fragments detected by MALDI-TOF-MS

```
132-140 FNVEVVAIR
  6- 16 DLVVSLAYQVR
 58- 78 FDVAVGANDAYGQYDENLVQR
 96-131 FLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLK
```

Silver-stain detection of soy-induced protein changes in brain proteins

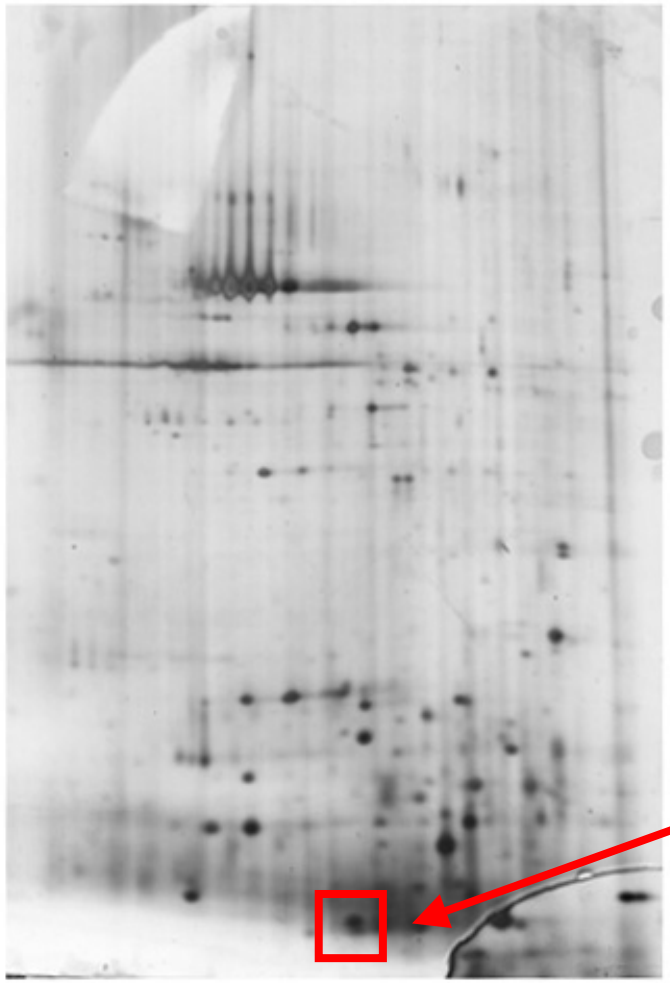


Soy + isoflavones

Soy - isoflavones

Helen Kim

2D-proteomics of HL-60 cells differentiated with DMSO



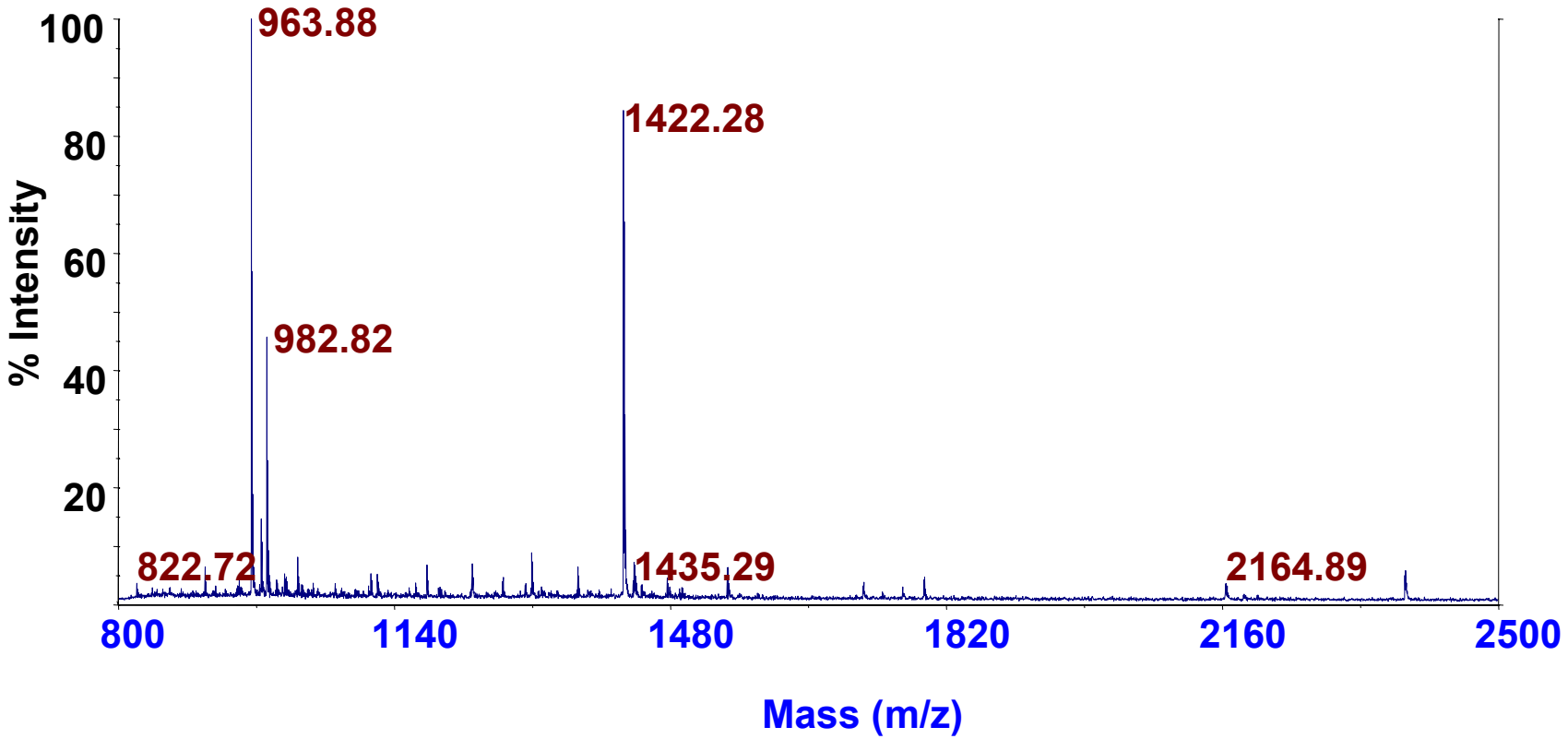
HL-60 cells were treated with DMSO for 7 days, causing their differentiation into a neutrophil phenotype

2D-proteomics analysis revealed spots that both increased and decreased in amount

This spot was absent in control HL-60 cells

Tracy D'Alessandro

MALDI-TOF analysis of trypsinized spot from 2D-analysis of DMSO-differentiated HL-60 cells



DMSO-treated HL-60 cell spot analysis

- DMSO-induced spot was shown to be S-100 (or calgranulin A) - a calcium binding protein

- BLAST showed that the sequence of S-100 is shared by migratory inhibitory factor related protein 8

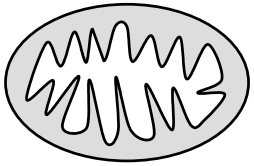
- Two of these entries have a pdb entry (4-letter alphanumeric descriptor) - this means there is a molecular structure available

Problems in all proteomics work

- **Non-stoichiometric recovery from the tissue/cell matrix**
- **Failure of proteins to absorb into or elute from IEF or SDS-PAGE gels, or column packing materials**
- **Non-linear response of the detecting method - coomassie blue, Sypro ruby, silver stain, etc.**

Complexes in lipid membranes

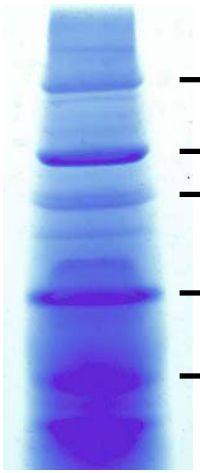
(Paul Brookes, Pathology)

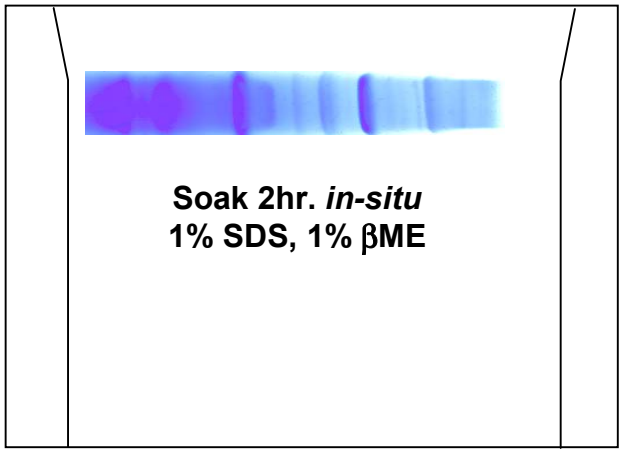


1.25% Laurylmaltoside
0.75M Aminocaproic acid
50mM BisTris
0.25% Coomassie Blue

4 – 12% Gradient Gel

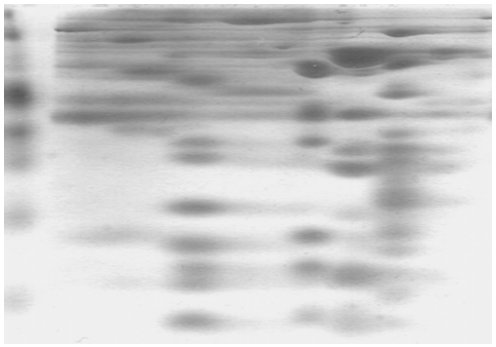
20V
15h
4°C





16% Tris-Tricine gel
10% native stacker gel

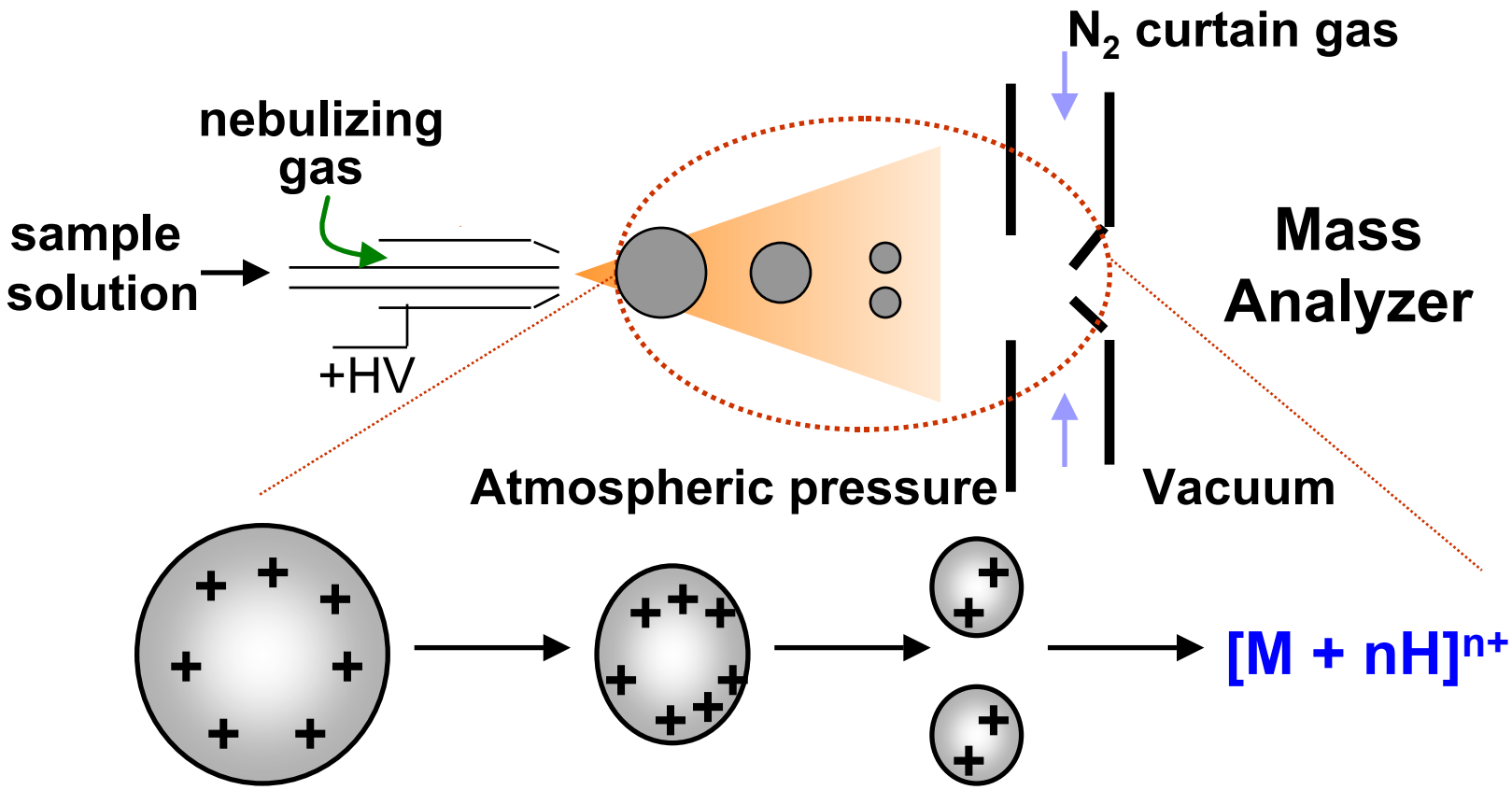
15h, 25V, 25°C



Options in proteomics analysis

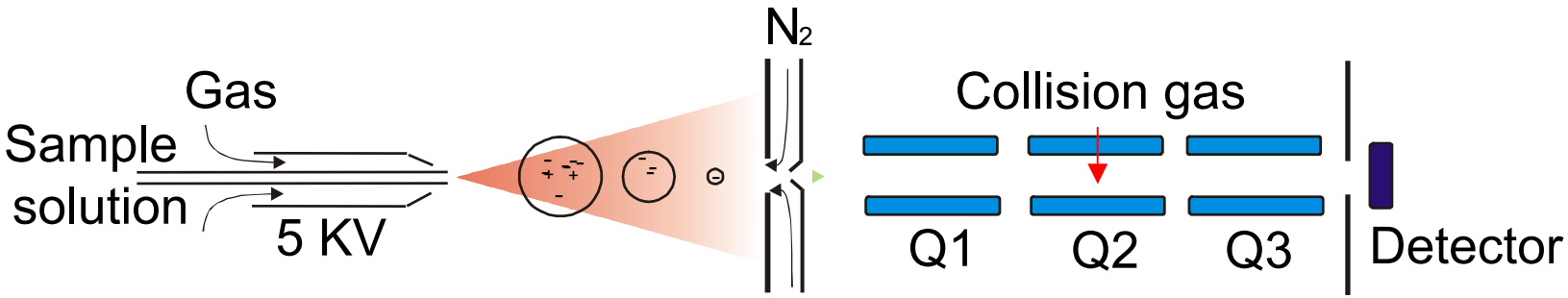
- **Methods for protein separation and identification**
  - ***2D-electrophoresis of proteins***
  - **Reverse phase nanoLC-MSMS of peptides**
  - **Ion exchange/reverse phase LC-LC-ESI-MSMS**
  - **Isotope-coded affinity tagging LC-ESI-MSMS**
  - **CE- or reverse phase nanoLC/MALDI-TOF-MS**
- **Radical methods on the horizon**

Electrospray Ionization (ESI)

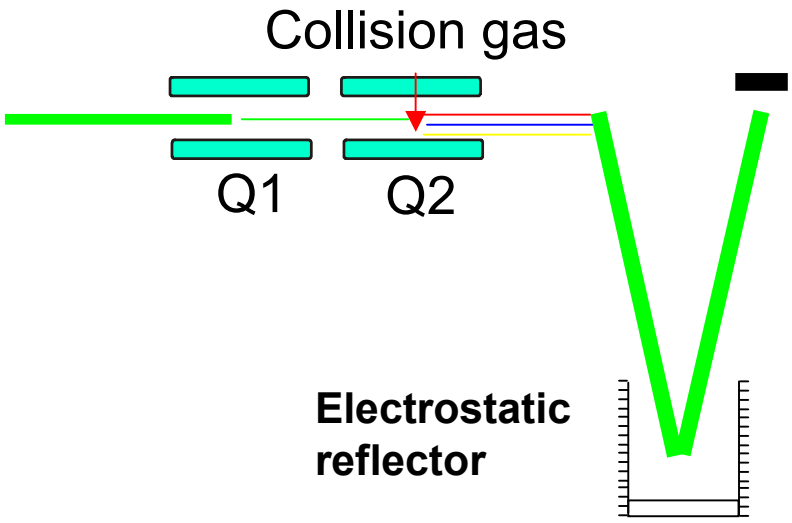


1. Solvent evaporation
2. Coulombic repulsion

Triple quad versus Q-tof and sensitivity



The quadrupole analyzer (Q3) is slow and insensitive - it's a filter - thus throws away large amounts of data



TOF detector collects all ions generated and yields fmol rather than pmol sensitivity

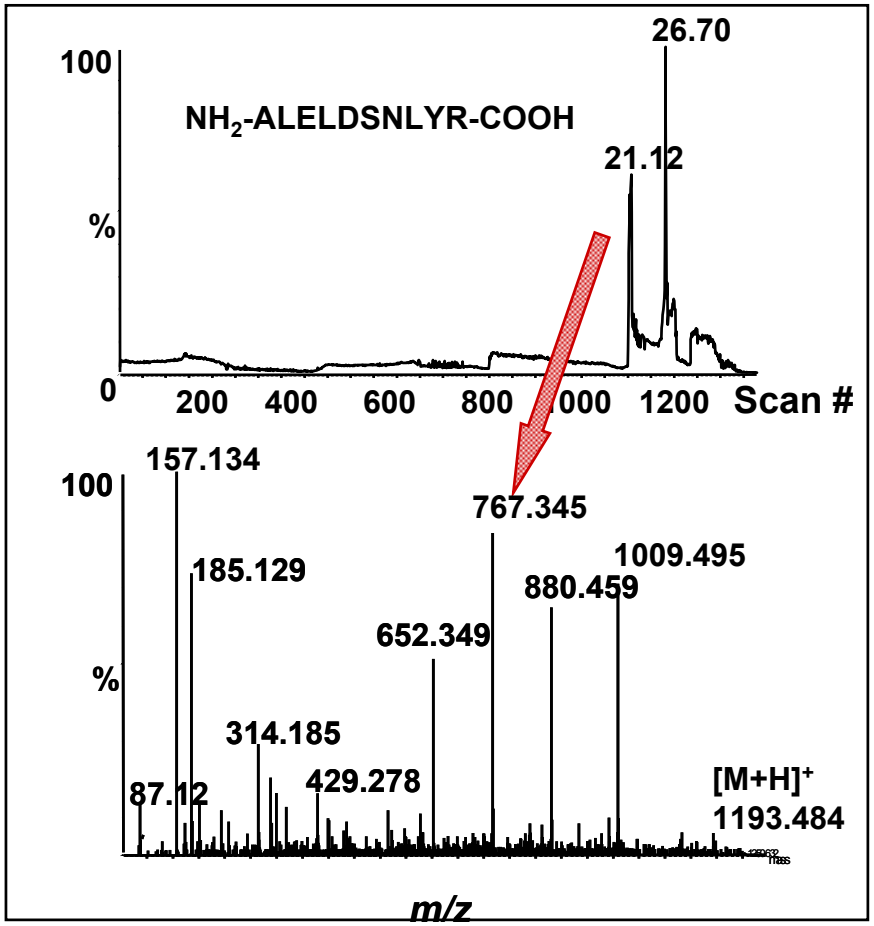
Also gives far greater mass accuracy - from 1000 ppm on the triple quad to 5-10 ppm on the Q-tof

Crucially important for automated interpretation of MS-MS spectra to yield amino acid sequence

Reverse phase nanoLC-MSMS

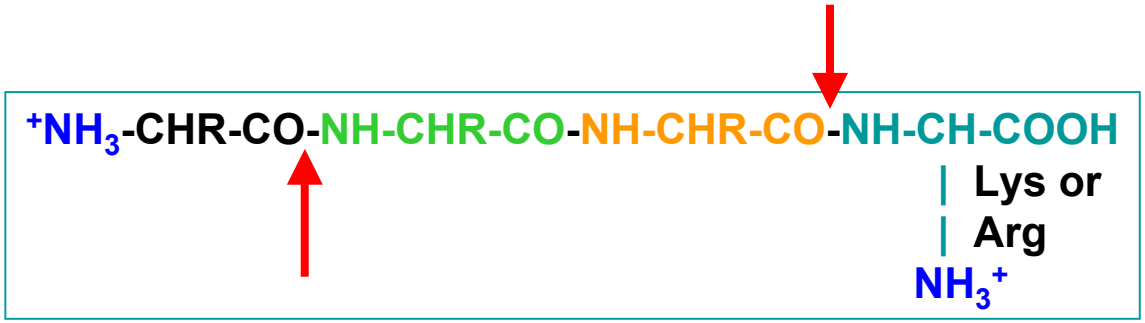
- Peptides are separated on an acetonitrile gradient using columns with i.d.s of 0.05-0.30 mm. These operate at 200-2000 nL/min

- Peptides are introduced by electrospray and analyzed on a Qqtof. Ions are selected by a quadrupole filter, collision-dissociated and analyzed by time-of-flight (accuracy 5-10 ppm)



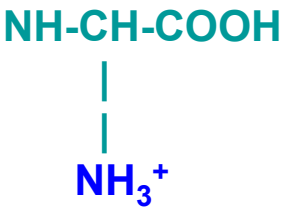
Fragmentation of peptide ions

Tryptic peptides are charged at each end



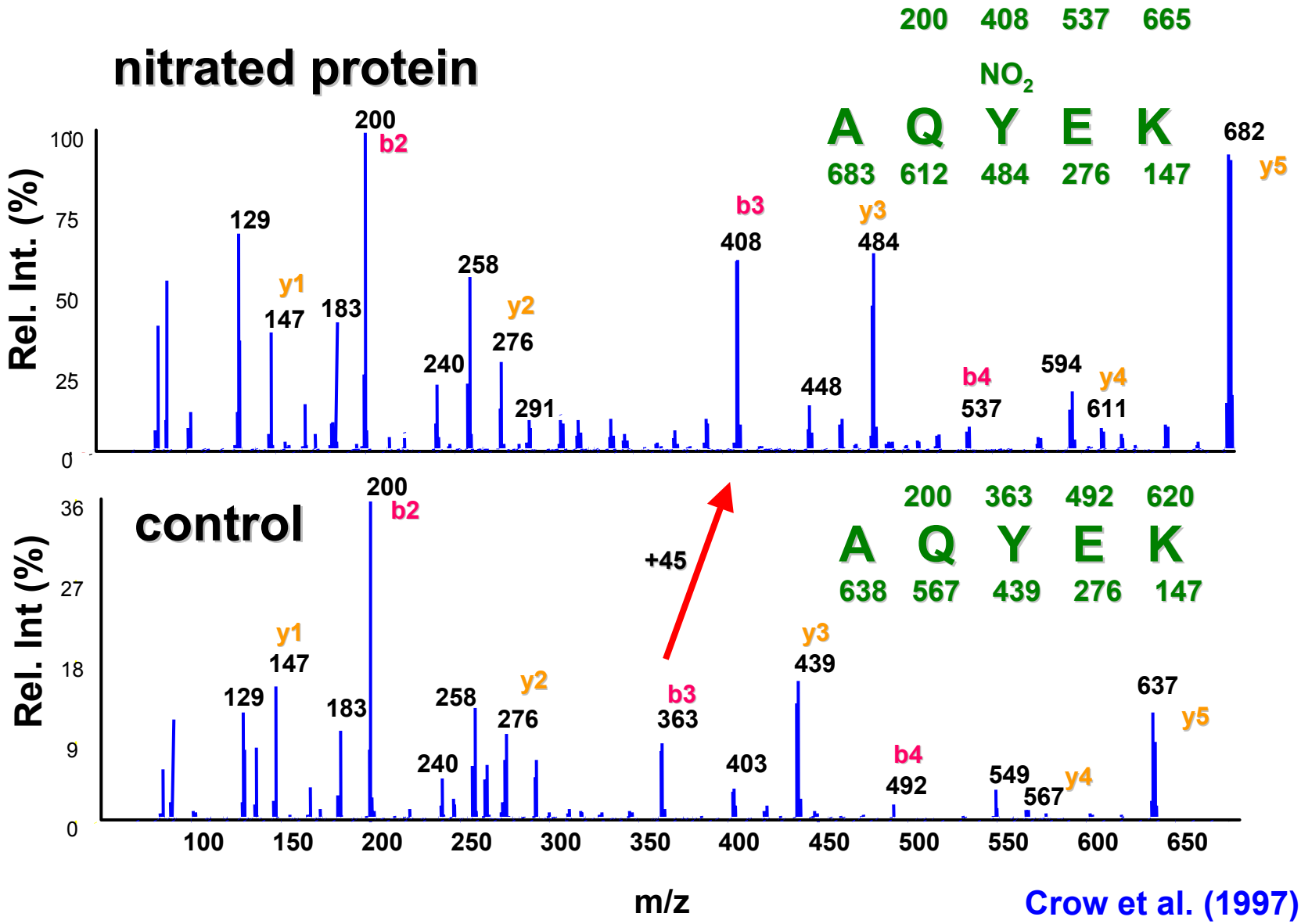
+NH3-CHR-CO

immonium ion



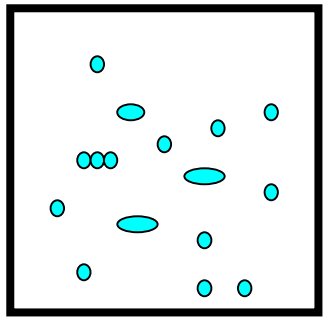
y1 ion

Site-specific nitration of a tyrosine-containing peptide using CID MS-MS spectra



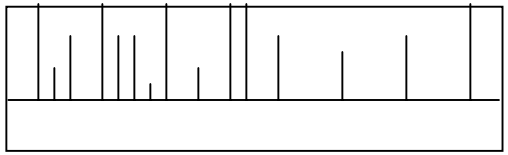
Crow et al. (1997)

When to use which technology?

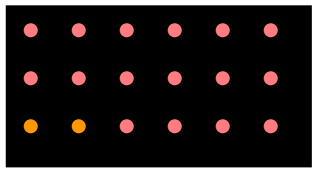


2-D gels (& MS):

- Available today; ***NCRR grant to Helen Kim for robotics***
- Is the only method that readily indicates posttranslational modifications;
- ***Least high throughput for the whole proteome, but the most informative***

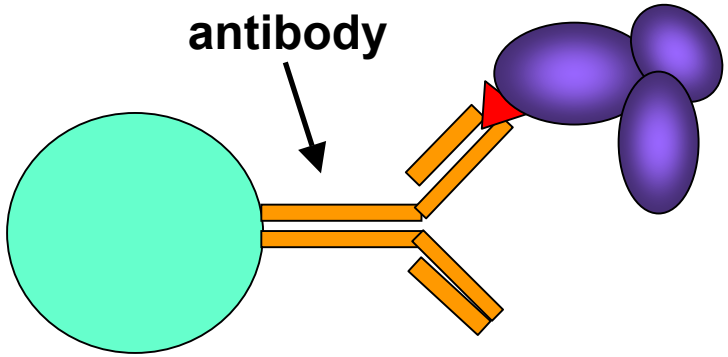


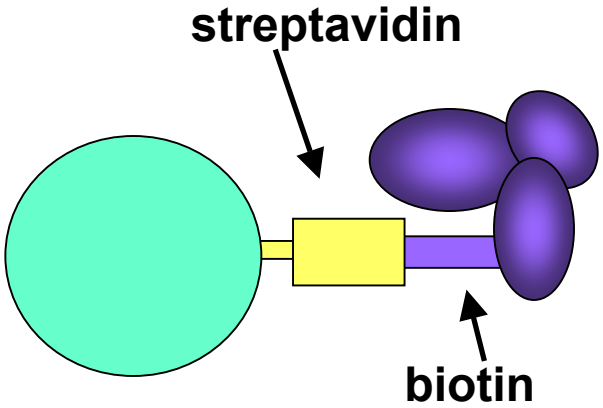
LC/LC-MS/MS: To catalogue a new proteome; HIGH THROUGHPUT; possible quantitation - at a price

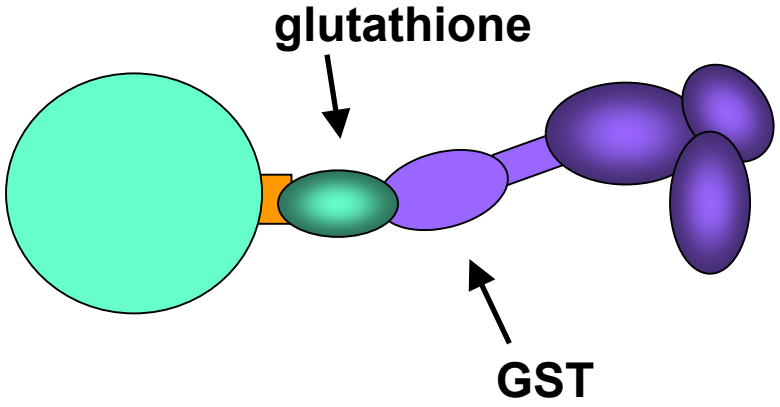


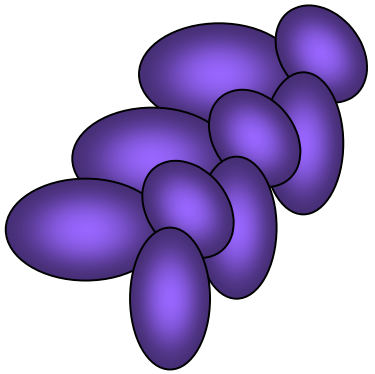
Protein Chip: To identify all possible targets of a drug, or ligand, or protein; HIGH THROUGHPUT

Affinity methods for recovering complexes







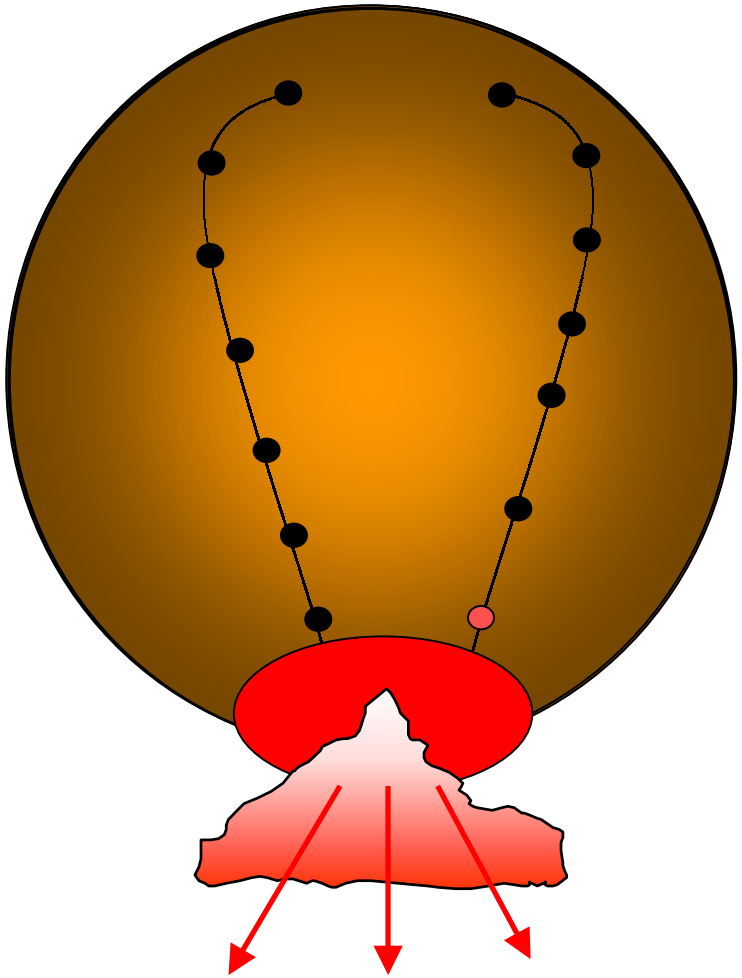


Multiprotein complex

Summary

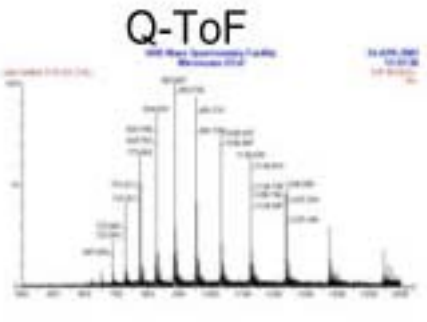
- Powerful proteomics and bioinformatics tools exist to separate and identify proteins
- 2D-proteomics and tandem mass spectrometry enables detection of protein modifications
- Protein-protein interactions are readily determined, even for membrane-associated complexes
- Proteomics has the power to assess where drugs act and where they shouldn’t

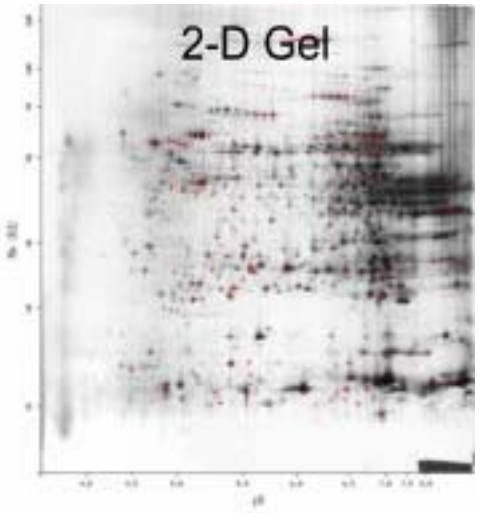
Visualization at the whole cell level

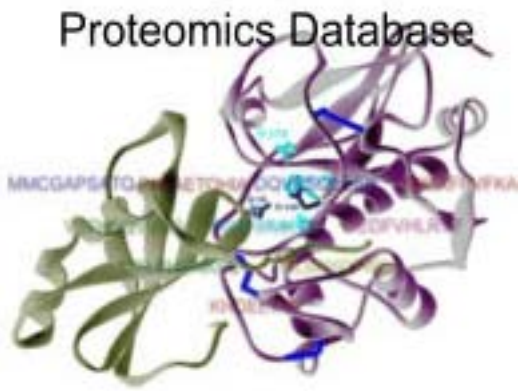


Mass Spectrometry website

http://www.uab.edu/proteomics







resources

news

charges

Acknowledgements

Helen Kim Lab
Lisa Chaves
Patti Hall
Robert Mills
Kiran Varma

Barnes' Lab
Mindan Sfakianos
Tracy D'Alessandro
Amanda Foxwell
Steven Burke

Colleagues:
Joe Beckman (Oregon St.)
Paul Brookes (UAB)
Peter Prevelige (UAB)

Proteomics & MS Facility:
Marion Kirk
Lori Coward
Tivanka DeSilva
Scott Isbell
Ramu Vempati

Support:
NCI
NIDDK
NIH-NCCAM-ODS
United Soybean Board
UAB-HSF-GEF