January 18, 2005

# 2D separations and analysis of proteins in biological samples

**Helen Kim**

**934-3880**

**helenkim@uab.edu**

**McCallum Building, room 460**

# Learning objectives

**2-D electrophoresis: What it involves:**

- **2D separation**
- **image, statistical analysis**
- **MS**

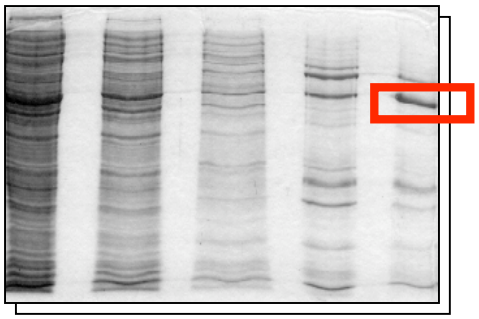
**Other types of 2D protein separations;**

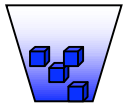
- **MUDPIT**
- **2D LC-LC**

**Other types of protein analysis technologies:**

- **protein arrays**
- **antibody arrays**

**Initially: The objective of protein separation in proteomics was to get a protein "spot" or "band", for mass spectrometry analysis, to identify the protein/its modifications**

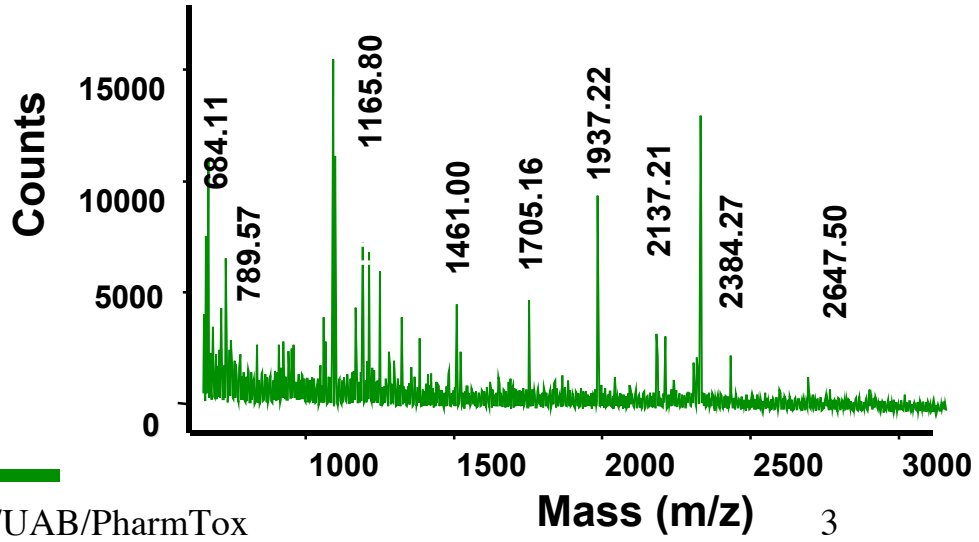




trypsin

MALDI-TOF mass spectrometry

ID of parent polypeptide



# Types of high-throughput separation & analysis technologies

I. 2-dimensional electrophoresis (2-DE)
   A. "regular" IEF/SDS-PAGE
   B. 2D-blue-native electrophoresis
   C. Visualization methods

II. 2-dimensional liquid-based LC/LC

III. Free-flow electrophoresis

IV. "Chip" technology: arrays of ligands for proteins

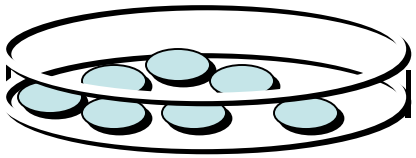
# Parameters that govern the choice of protein separation method

- Purity of protein
- Speed of purification
- Quantity of protein
- What is the question:most important
  - Discovering a new protein/proteome
  - Identifying protein-protein interactions
  - Identifying potential modifications of known proteins

# The Elements of any 2-D separation in proteomics

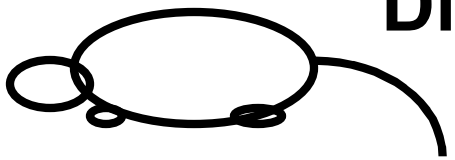
- **The experiment!! that generates the "signal"**
- **Subfractionation to enrich for suspected proteins**
- **(Trypsin-digestion to generate peptides of the parent proteins)**
- **Sample work-up**
- **2-D separation**
- **Image or other analysis to identify gel "spot" differences between untreated & treated**
- **MS of (trypsin-digests) spots/proteins/peptides of interest, to identify and characterize the protein**

# Sample preparation for 2DE:



**Harvest, rinse, and pellet the cells;**

**or**



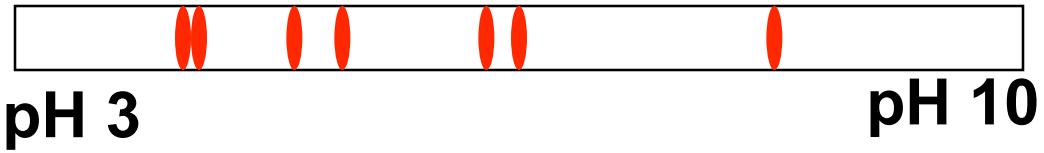
**Dissect out tissue, organ, or fluids;**

- **Homogenize/lyse in buffer that *unfolds the proteins w/o adding or disturbing the charges*:**
  - **High urea usually 5-8 M---unfolds the protein**
  - **Sometimes 2 M thiourea--unfolds the protein**
  - **1-4% detergent--solubilizes hydrophobic components**
  - **Beta-mercaptoethanol or other reductant, such as TBP, DTT**
  - **Inhibitors: of proteases, kinases, & phosphatases**
- **Clarify by centrifugation to get rid of insoluble matter;**
- **Protein assay to know how much and how concentrated**

# What 2-D electrophoresis involves:

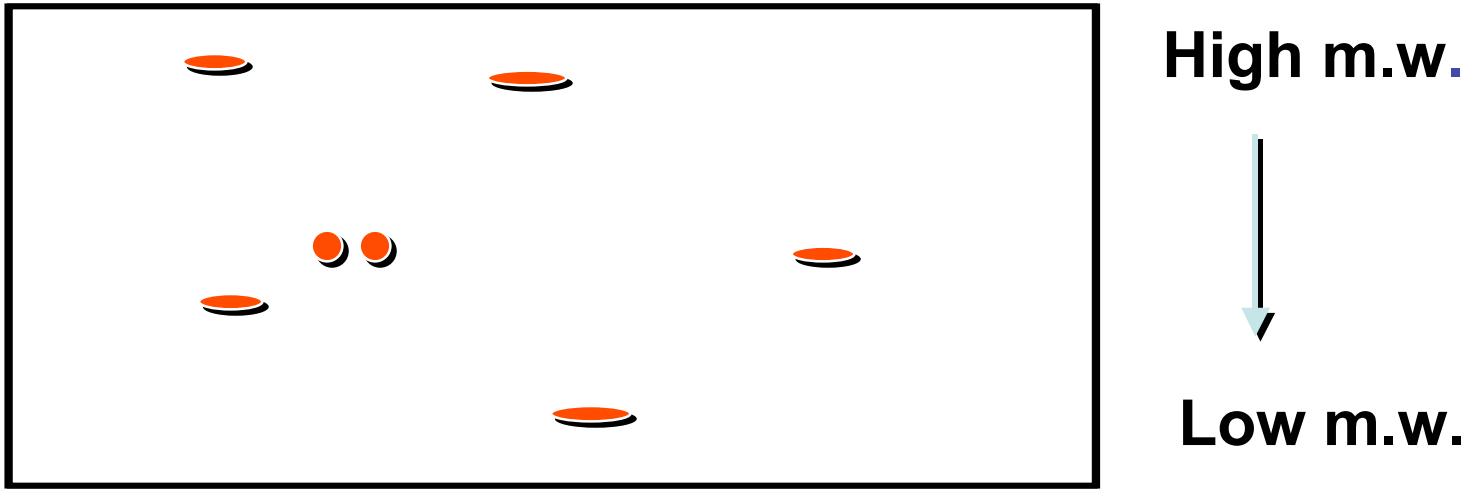
- 1st dimension: **Isoelectric focussing**

**(separation according to charge )**

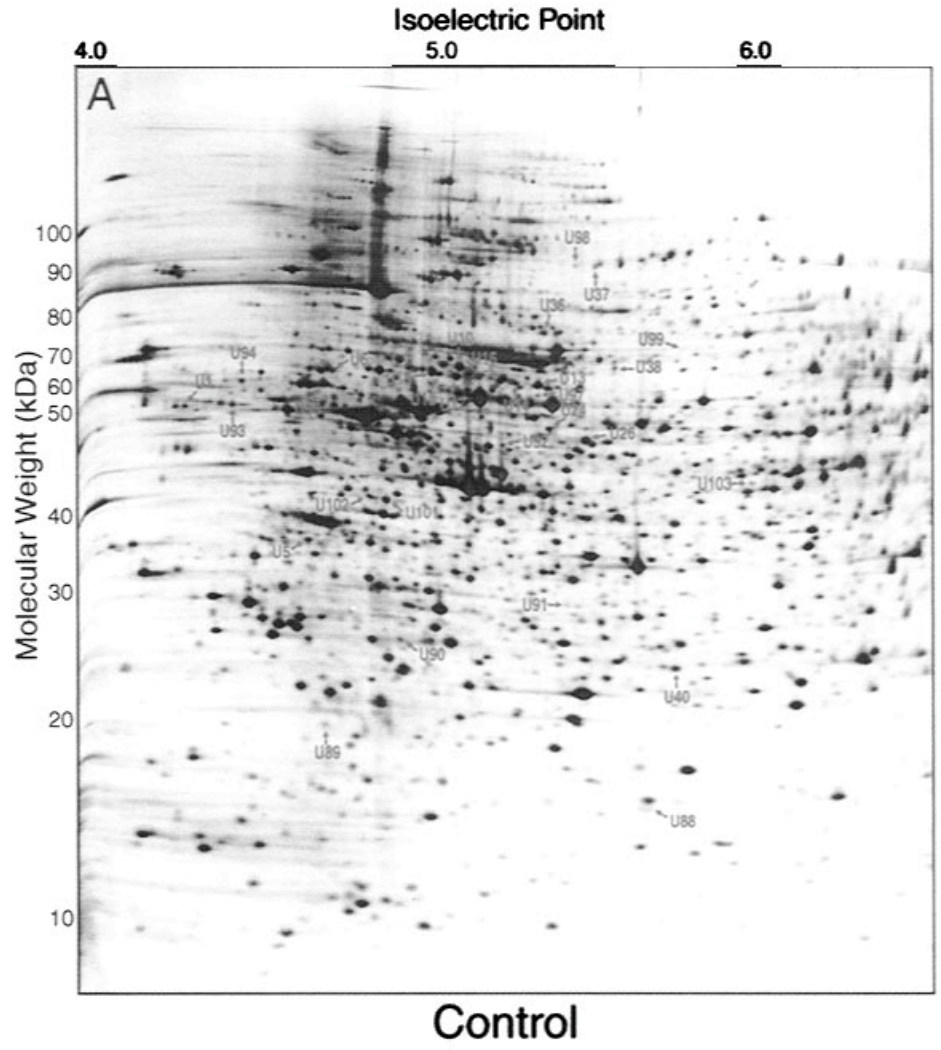



- 2nd dimension: **(SDS)-PAGE**

**(separation according to size )**



# A real 2-D gel



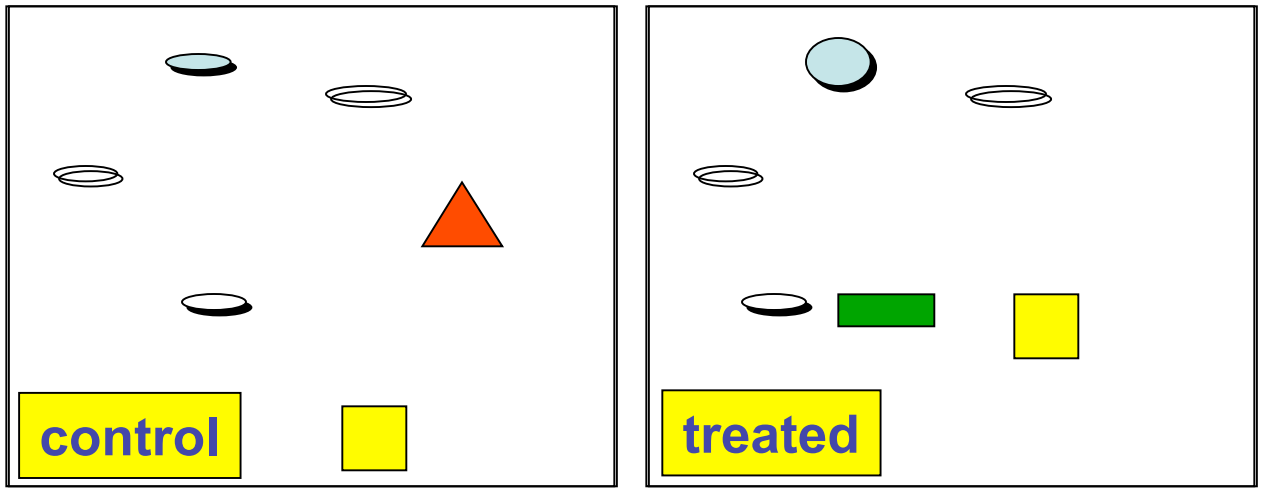

**Find this and other 2-D gels at http:www.expasy.org**

**Lewis et al., [2000] Molec. Cell, 6)**

**(from Natalie Ahn's lab)**

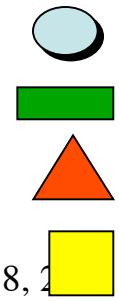
**The pattern itself is information; a change in intensity of a spot is not meaningful unless you know it is the same spot.**

# Critical part of 2-D gel proteomics: Image analysis




**Either manually or with software: "compare" the images.**

## Types of information:



------ Upregulation of gene

------ Posttranslational modificatiion

------ Downregulation of gene

------ Aberrant processing

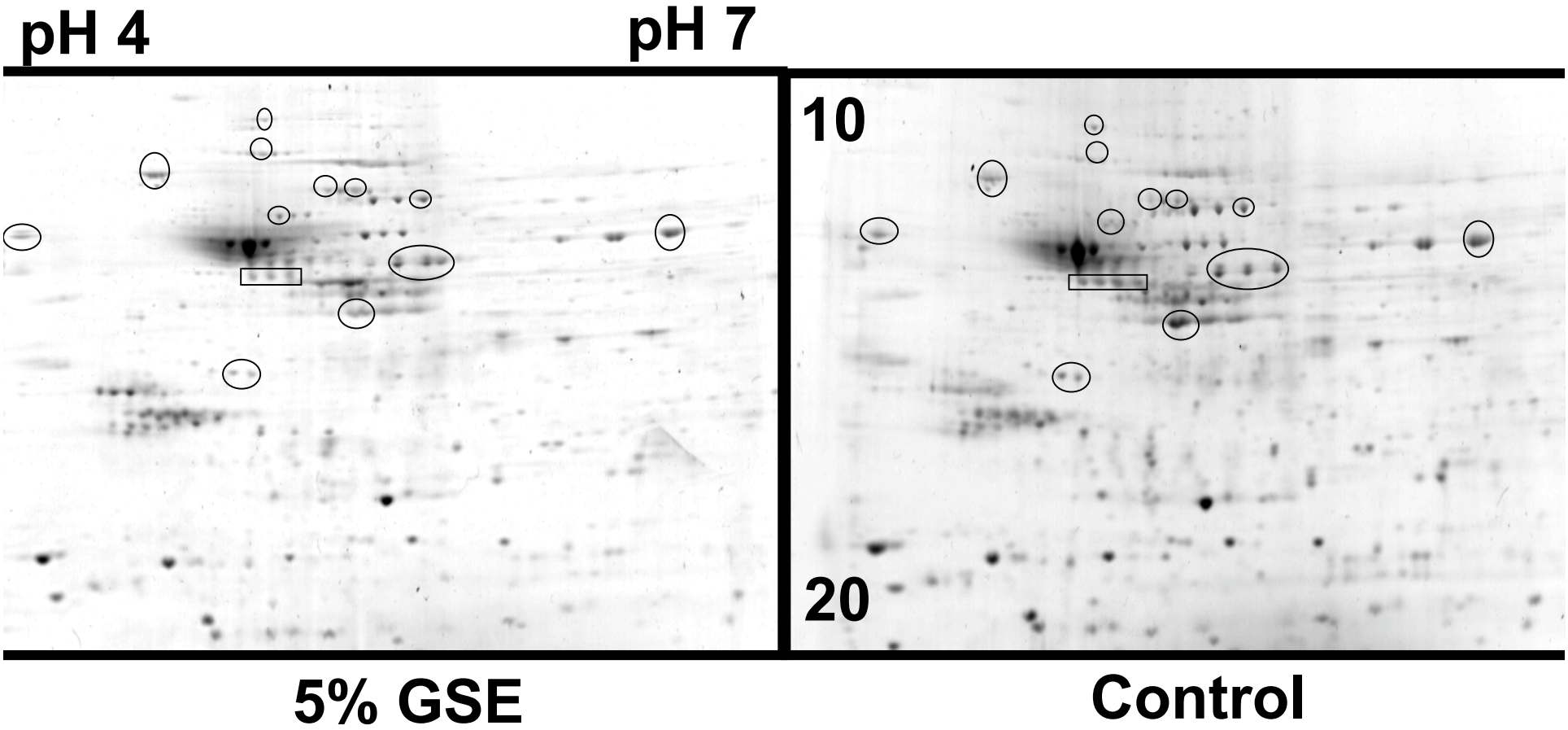
## Elements of image analysis of "regular" 2D gels:

1. **Compare the 2D displays of spots**
2. **Determine total spot number for each display**
3. **Quantify spot intensities, identify differences**
4. **Identify spots that may have " moved" horizontally; these are candidates for those may have alterations in charge.**

## Ultimate and simple goal of image analysis

**to answer the question,**
**"What is changing, and by how much?"**

# A pair of 2D gels representing rat brain protein changes induced by ingestion of grape seed



# Categories of data generated by 2D gel image analysis

| SSP | intensity | X coordinate | Y coordinate | Peak quality |
|---|---|---|---|---|
| 1369 | 12,269 | 2.375 | 1.279 | 75 |
| | | | | |

Differential expression?

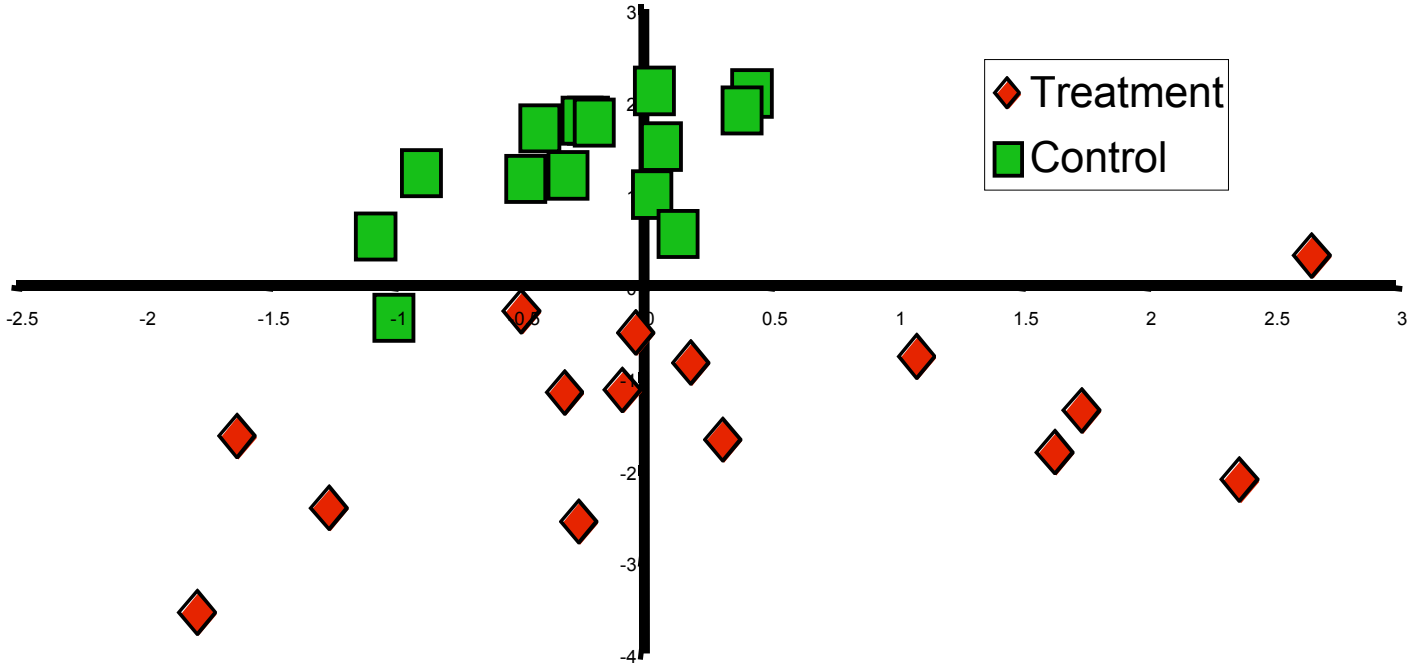
Modifications?

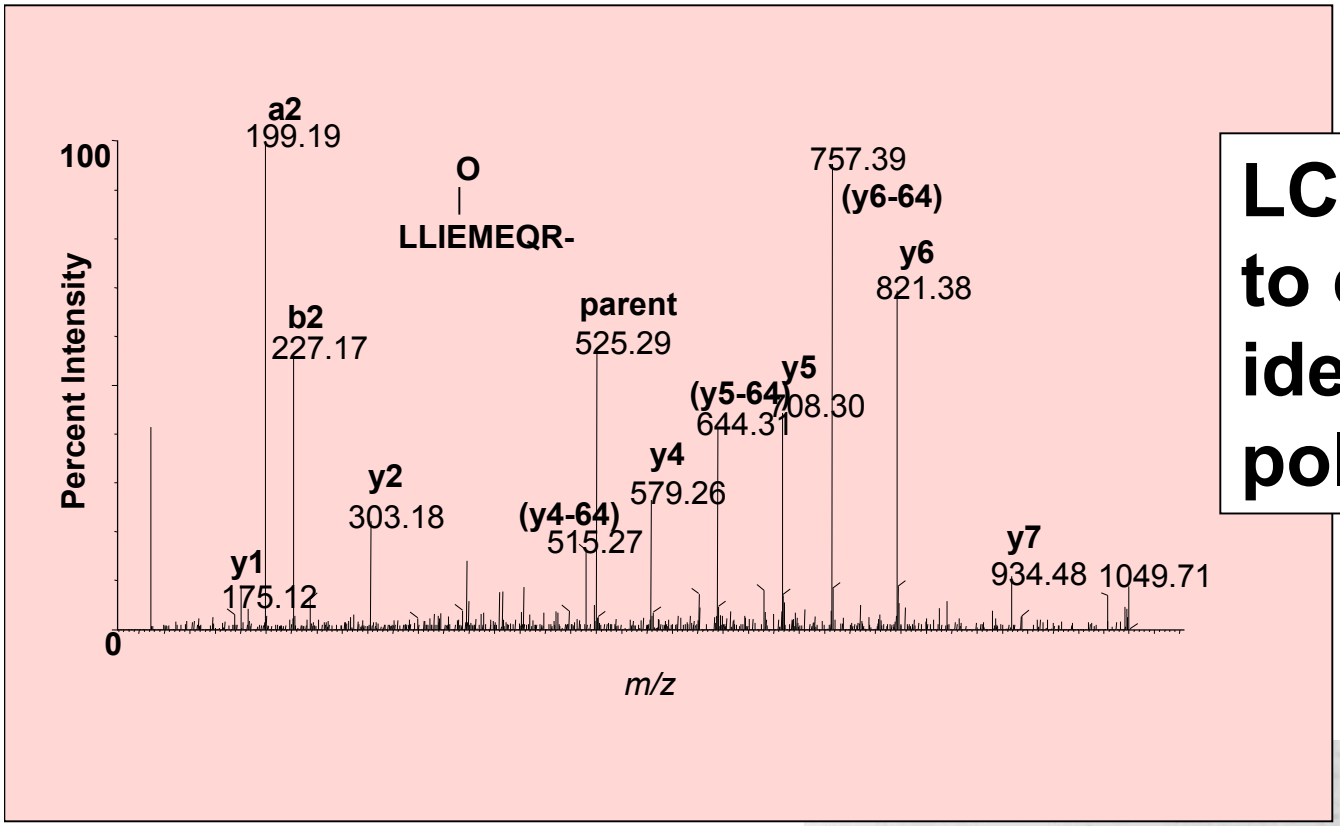
Change in Complexity?

**typical dataset: >100,000 rows of data**

**Statistical analysis**

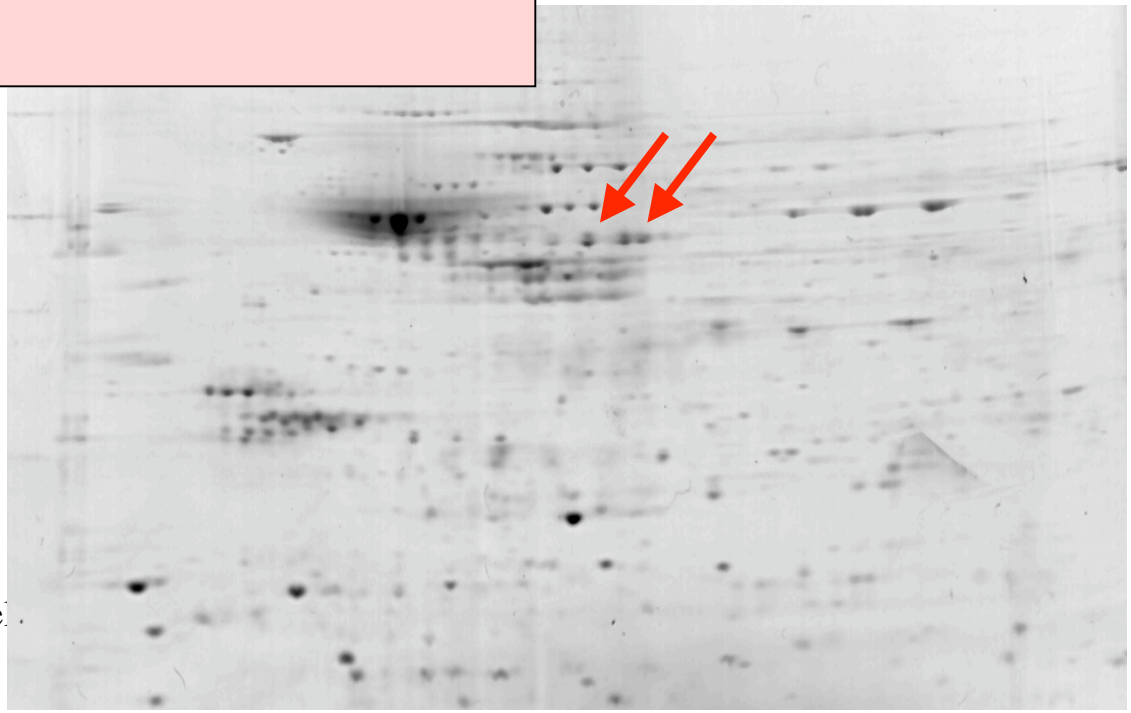
# Principal Components analysis assists in determining whether the gel patterns distinguish between the experimental groups.



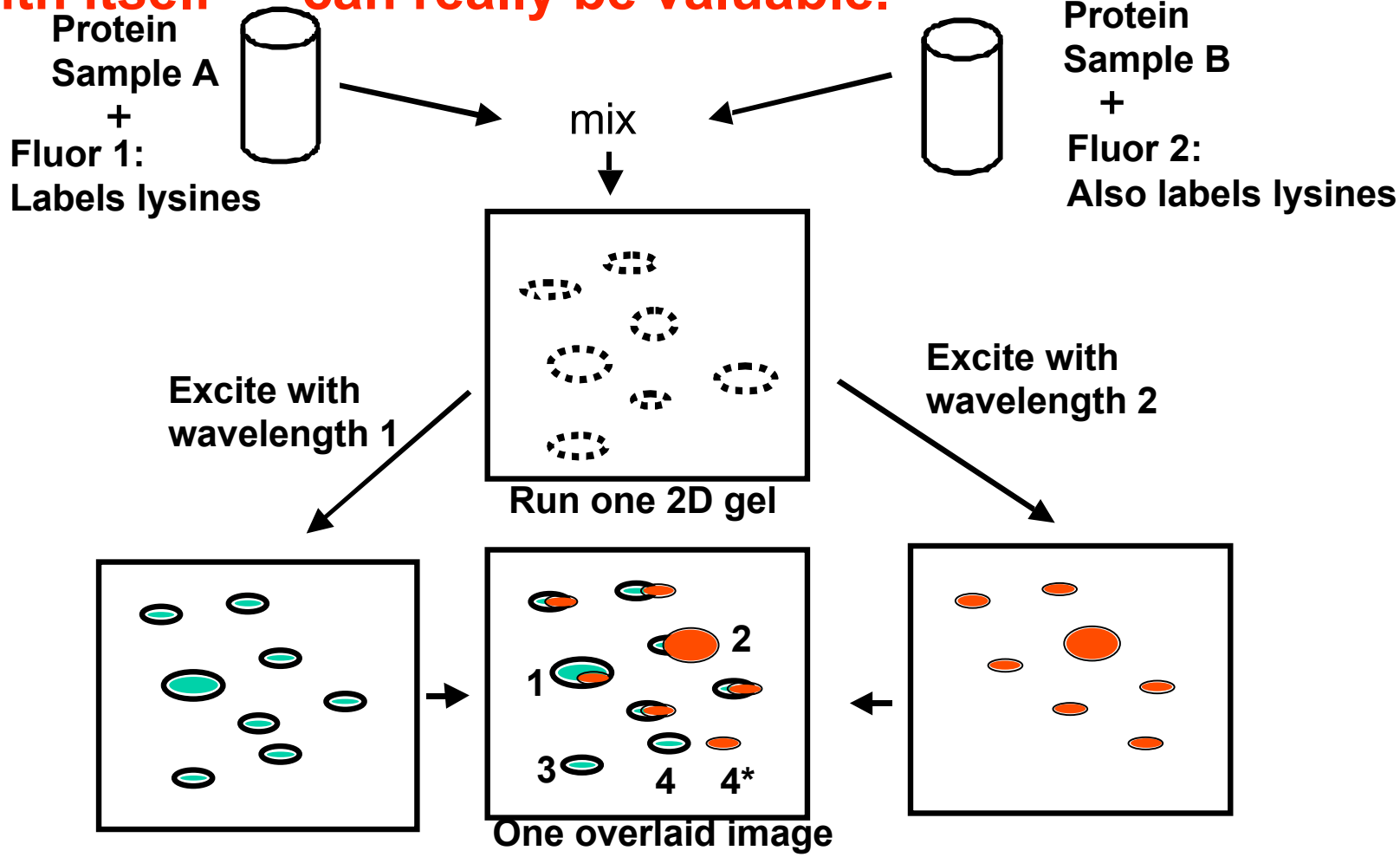



**LC-tandem MS to conclusively identify the polypeptide**

**Creatine kinase brain beta (BB) chain**



# Difference gel electrophoresis (DIGE); a protein migrates with itself---- can really be valuable.



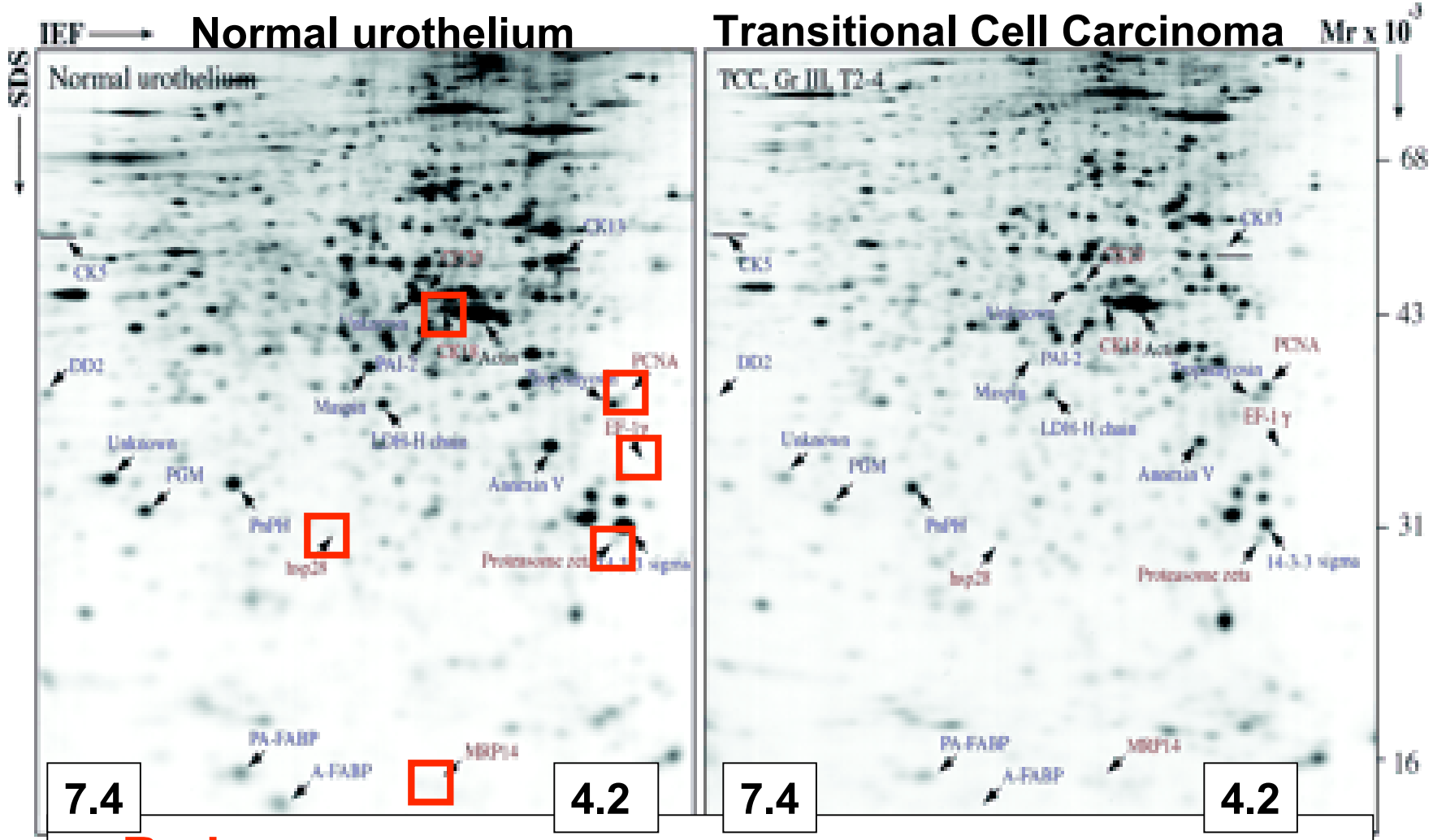

**Single biggest bottleneck, even with DIGE: even commercially made gels (1st dimension, 2nd dimension) are not perfect. 2D displays of proteins across large datas**

# Database obtained from proteomics analysis of breast cancer cell lines

| Sample | Spot# | Protein Name | MOWSE | Accession# | Obs: kDa | Pred: kDa | Obs: pI | Pred: pI |
|---|---|---|---|---|---|---|---|---|
| MCF-7 | BR-8 | Cytoskeletal keratin-7 | 132 | gi12803727 | 20.7 | 51.3 | 8.6 | 5.3 |
| MCF-10AT | BR-9 | Alpha-1 anti-trypsin | 90 | gi1942953 | 19.7 | 44.3 | 9.6 | |
| | BR-1 | Human Cytoskeletal keratin-8 | 94 | gi87303 | 56.2 | 53.5 | 5.9 | 5.6 |
| | BR-6 | Human | 74 | gi18573275 | 23.4 | 24.1 | 6.3 | |
| | BR-5 | hypo:XP109048 IgG heavy chain | 60 | gi5106591 | 18.1 | 6.4 | 5.7 | |
| | BR-7 | variable region | 105 | gi1942953 | 37.0 | 44.3 | 6.9 | |
| | BR-4 | Intact recomb: alpha-1 antitrypsin mutant F-L IgG heavy chain var:reg | 65 | gi5106591 | 10.0 | 6.4 | 4.6 | |
| MCF-10A | BR-3 | Human cytokeratin-8 | 65 | gi87303 | 56.0 | 53.5 | 5.8 | 5.6 |
| | BR-2 | DNA replication silencing factor MCM-4 | 58 | gi1705520 | 22.8 | 96.5 | 6 | 6.6 |

**(Deshane, Johanning, and Kim, unpublished data)**

# Metabolic labelling can enhance 2D gel analysis: i.e. $^{35}$S-methionine-labelling




**Red** = upregulated in the tumor; **blue** = down-regulated.



(Celis et al., Mol.Cell Proteomics, 2002, 1:4)

# Make use of databases and the internet:

**I. Check existing databases and web-links:**

**www.expasy.org**

***many are annotated***

**helpful links:**

**proteomics tools**

**II. Keep up with the literature/ competition:**

***Electrophoresis***

***Proteomics***

***Molecular & Cellular Proteomics***

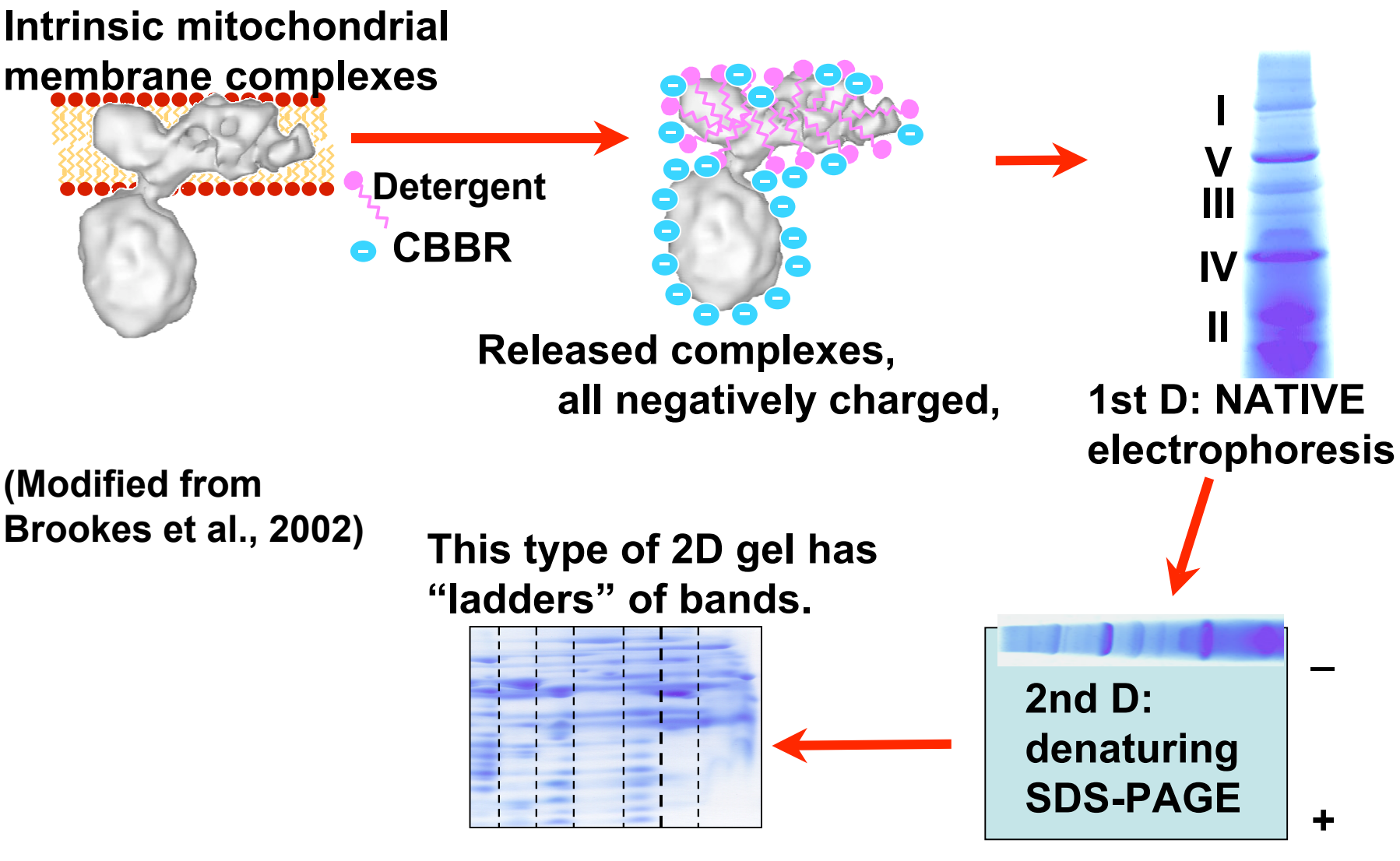
***J. Proteome Research***

**III. Use genomics information when available:**

**The polypeptide sequence (from the cDNA) can predict electrophoretic parameters-- m.w. & pI;**

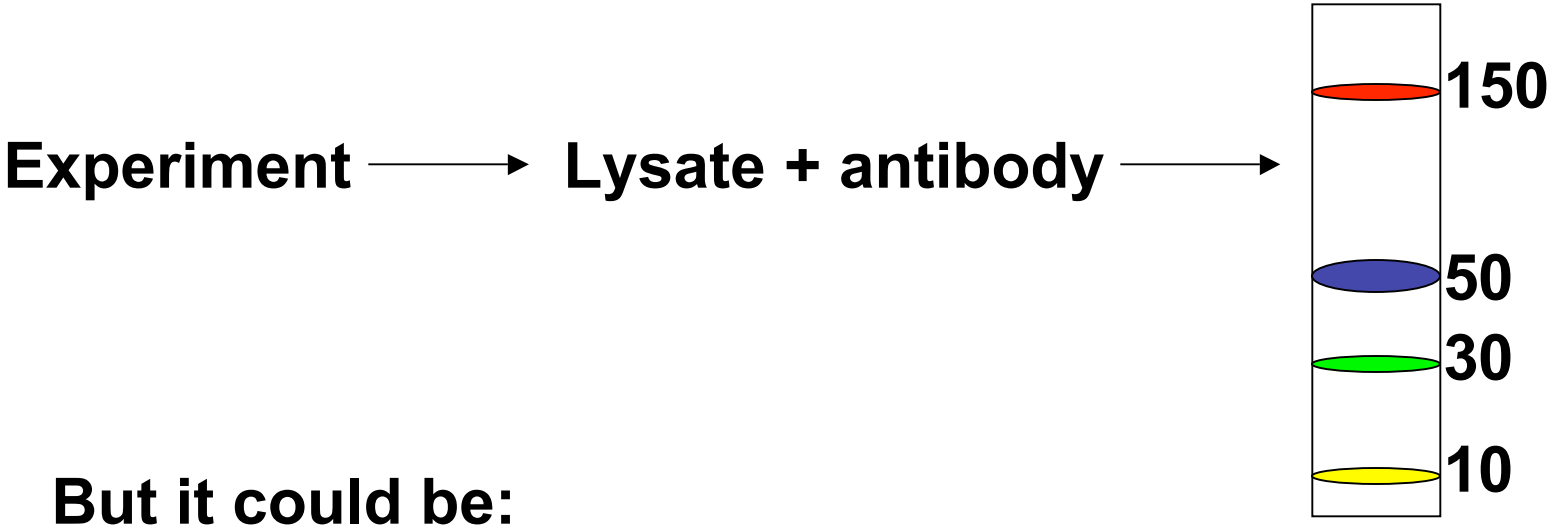
***helpful in setting up 2D gel conditions***

# 2D-Blue-Native gels: for hydrophobic proteins

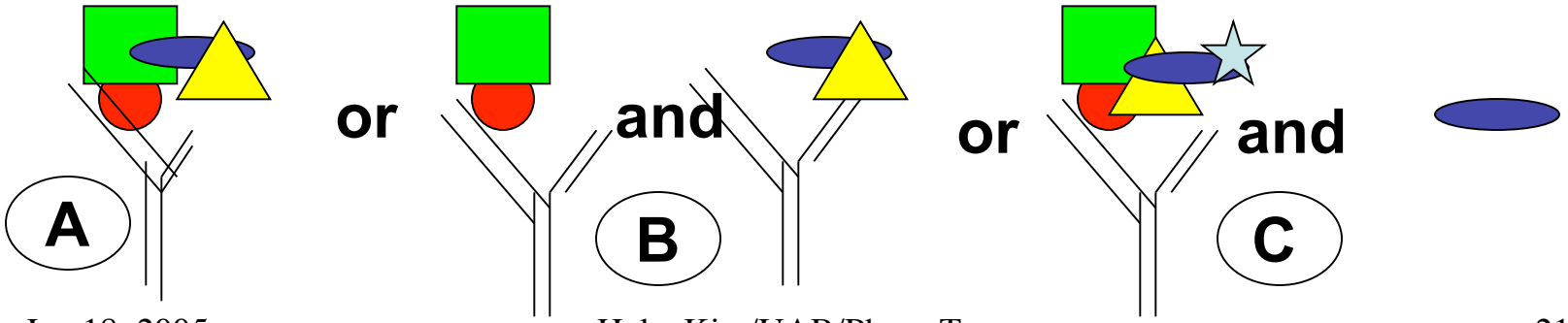


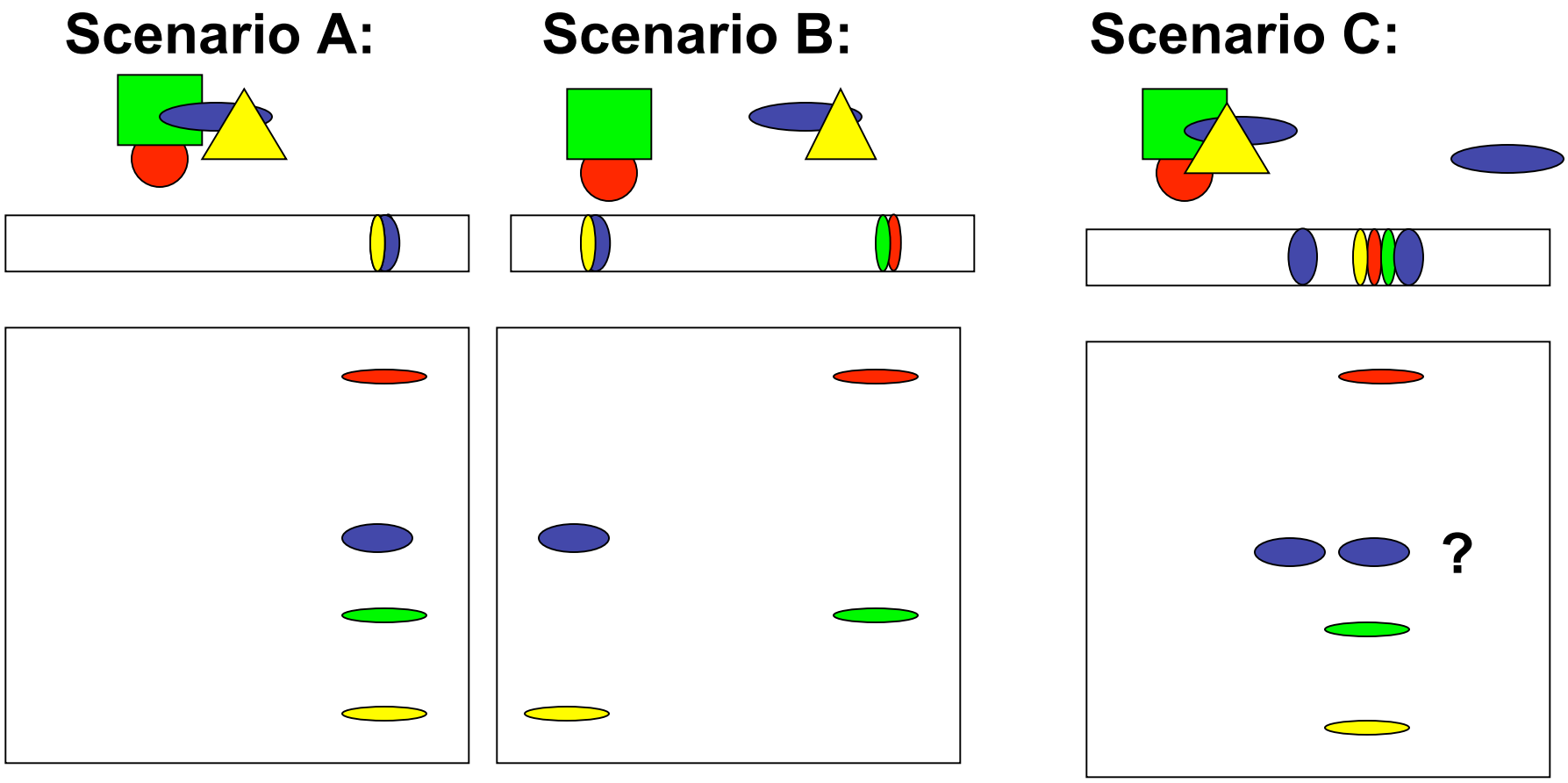
# What questions does BN electrophoresis address:

## (a) Which proteins are actually interacting with which?




**But it could be:**





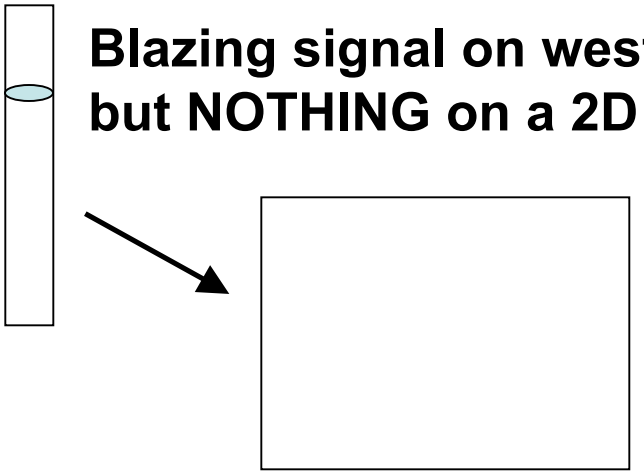

**When would you see the same protein in two lanes on the 2n**

# Take home message

- **Use of proteomics technologies enables *global* analysis of protein changes ;**
- **2D electrophoresis can indicate both differential expression or posttranslational modifications;**
- **Choice of separation governed by**
  - **Abundance of sample**
  - **Question being asked**
  - **Technology available to you**
  - **Cost is a factor;**

# Issues in 2D gel analysis:

**I. Now you see it, now you don't:**



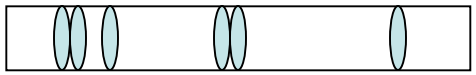
**Blazing signal on western blot**
**but NOTHING on a 2D gel of the same lp:**

Probable answers/what are the solutions:

1. Protein not solubilized by IEF buffer
2. Not enough protein!!!
3. Modification lost during preparation?

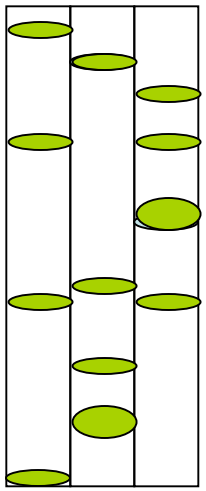
**II. I see the band on my 1D gel that runs with my western blot band; Why do I need to run a 2D gel anyway?**

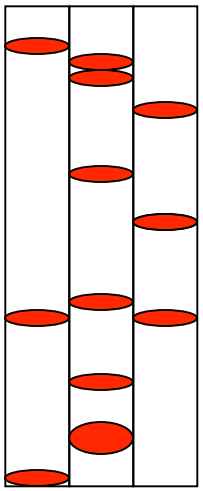
# 2D-LC-LC

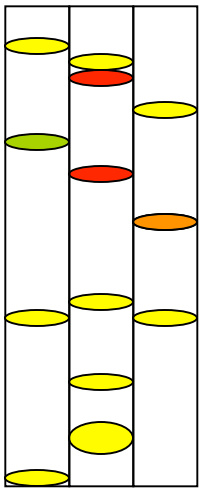


1st dimension: chromatofocussing (like IEF, but *in solution*, so can take higher protein loads)

**2nd D: RPLC**







**normal** **disease**

**Digitized super-imposed images; Image analysis fundamentally different from 2D gels; Quantitation MUCH Faster; CAVEATS?**

**cost**

**relatively new**