November 13, 2015

Protein purification and top-down proteomics to integrate with other omics strategies

Helen Kim

934-3880

helenkim@uab.edu

Dept of Pharmacology and Toxicology

McCallum Building, room 454

http://www.uab.edu/proteomics

Objectives of HK part of 11-13-15 lecture

- **Purification of proteins based on their properties**
- **1D- and 2D- separations of intact proteins (topdown) for indication of differences in abundance or PTMs;**
- **Different 2D topdown methods for quantitative intact proteomics (Top-down)**
- **Discovery to detect differences, vs discovery to catalog maximum number of proteoforms**
- **Everything old is new again: modern immunoprecipitation + LC-MSMS to define nuclear transcription factor components**

Biological question determines extent and nature of protein purification for proteomic analysis

Different rationales for studying proteins in biological samples:

A. Nothing is known, and all proteins are of interest;

B. Nothing is known, but only proteins that are different between treatment groups are of interest;

C. PTM of a single protein or a PARTICULAR PTM category; *a genomics analysis may have indicated that a KINASE is upregulated, but NOT which proteins are increased in phosphorylations.*

D. Protein interactions involving a specific protein are of interest

Properties of polypeptides that enable separation from each other

- **Intrinsic properties**
  - **Size—number of amino acids**
  - **Net Charge**

- **Biological/functional properties**
  - **Intracellular location**
  - **Enzyme activity**
  - **Undergoes oligomerization**
  - **Undergoes modification**

Separating proteins by size

- **Gel filtration chromatography:**
- **Separate under native or denaturing conditions: when do you want to do which?**
- **SDS-PAGE of fractions allows assessment of complexity in each, extent of separation**

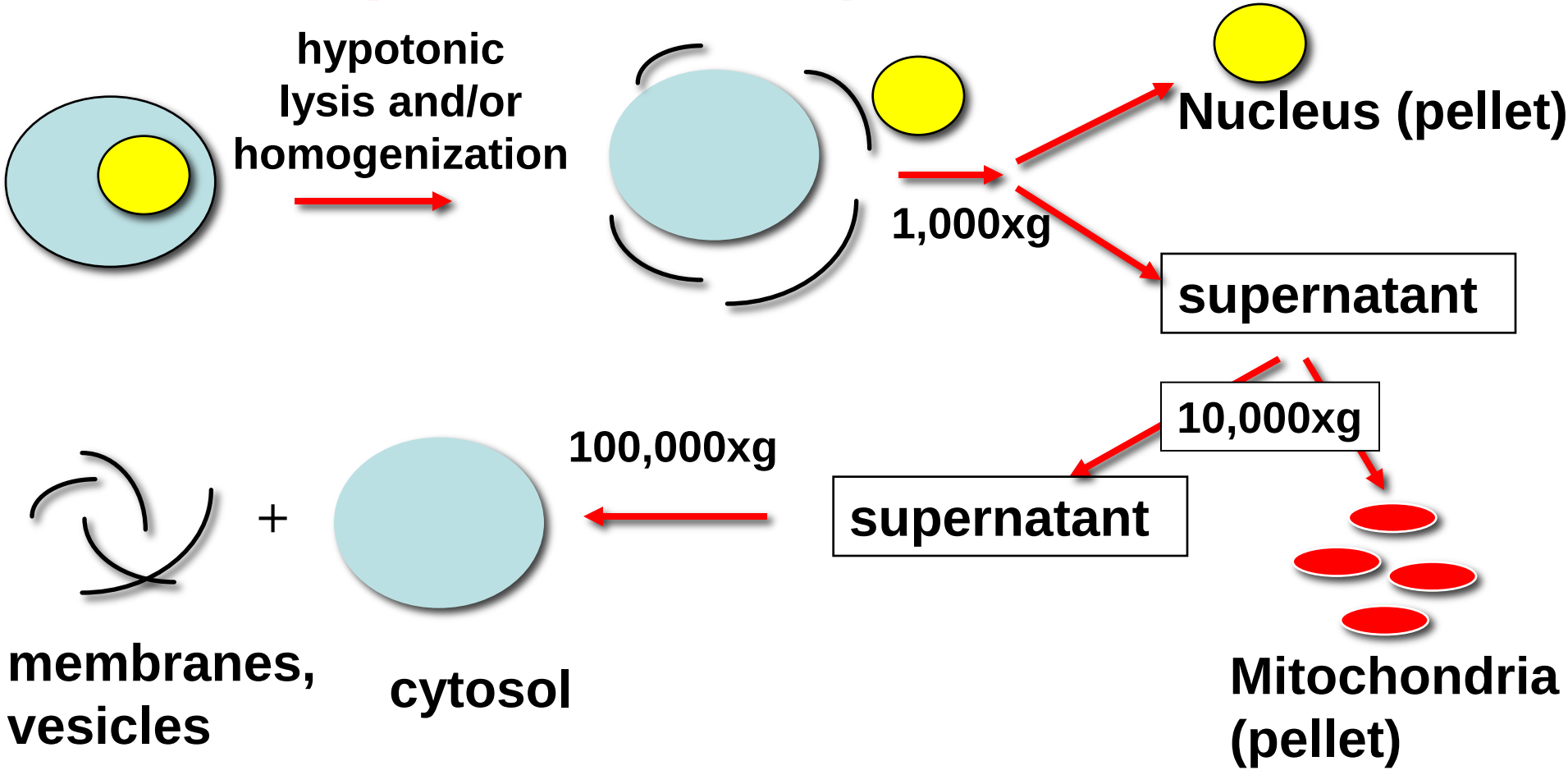
Separating proteins by charge

- **Ion exchange chromatography**
- **Takes advantage of the charged character of proteins; can greatly concentrate one fraction from the other, by the former binding to the ion exchange resin;**

Separating proteins by charge,part II

- **Isoelectric focussing:**
- **As with SDS-PAGE, was an analytical method;**
- **With the sensitivity of MS, can be a purification method for resolving multiple proteins**
- **Particularly effective at resolving proteins that differ by PTMs**

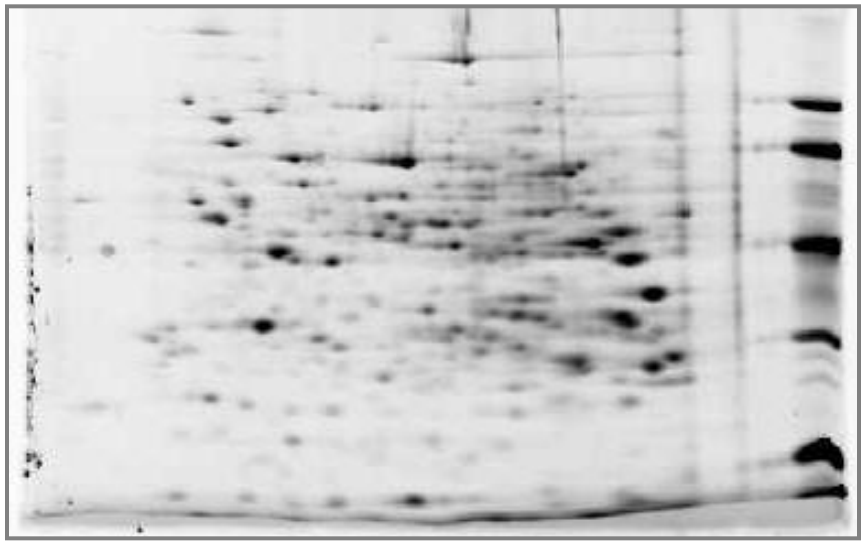
Don’t discount “conventional” or “old” methods like differential centrifugation to enhance proteomics experiments.



Analysis of mitochondrial proteins enhanced by purifying that subproteome

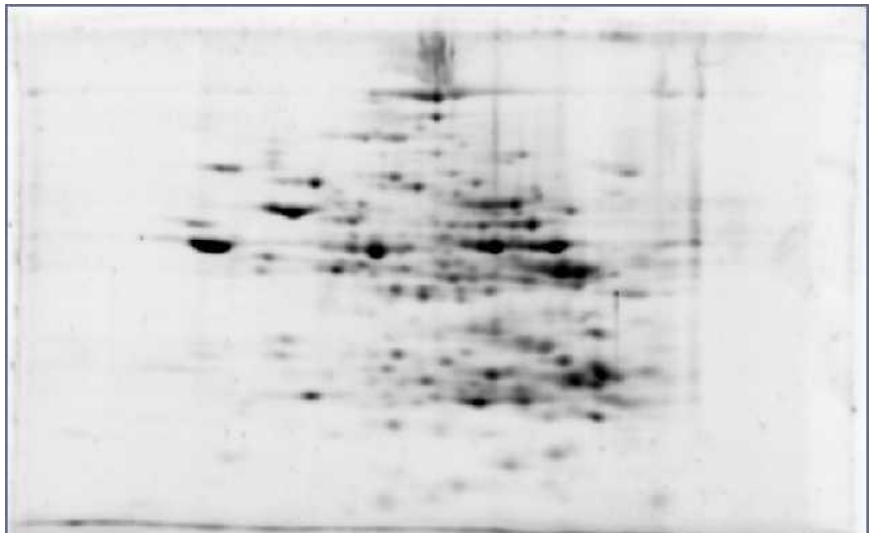
Whole Rat Liver

3 10



Rat Liver Mitochondria

3 10

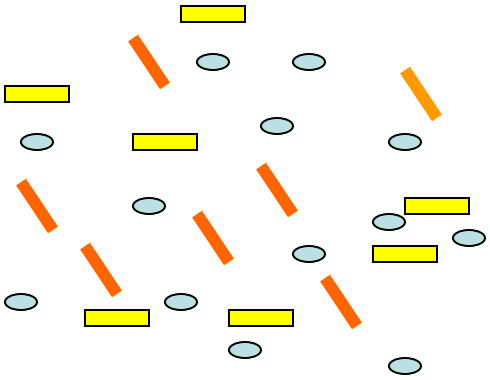


Superimposition of these two images reveals little overlap.

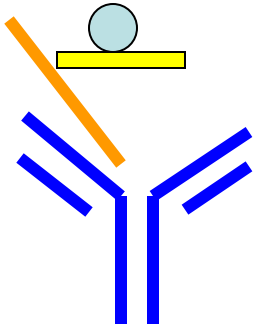
2D gel analysis using the Invitrogen ZOOM system

(Courtesy of Shannon Bailey Lab – Whitney Theis and Kelly Andringa)

Antibodies can reduce the complexity of the proteome, as well as enhance biological specificity, by 10,000-fold

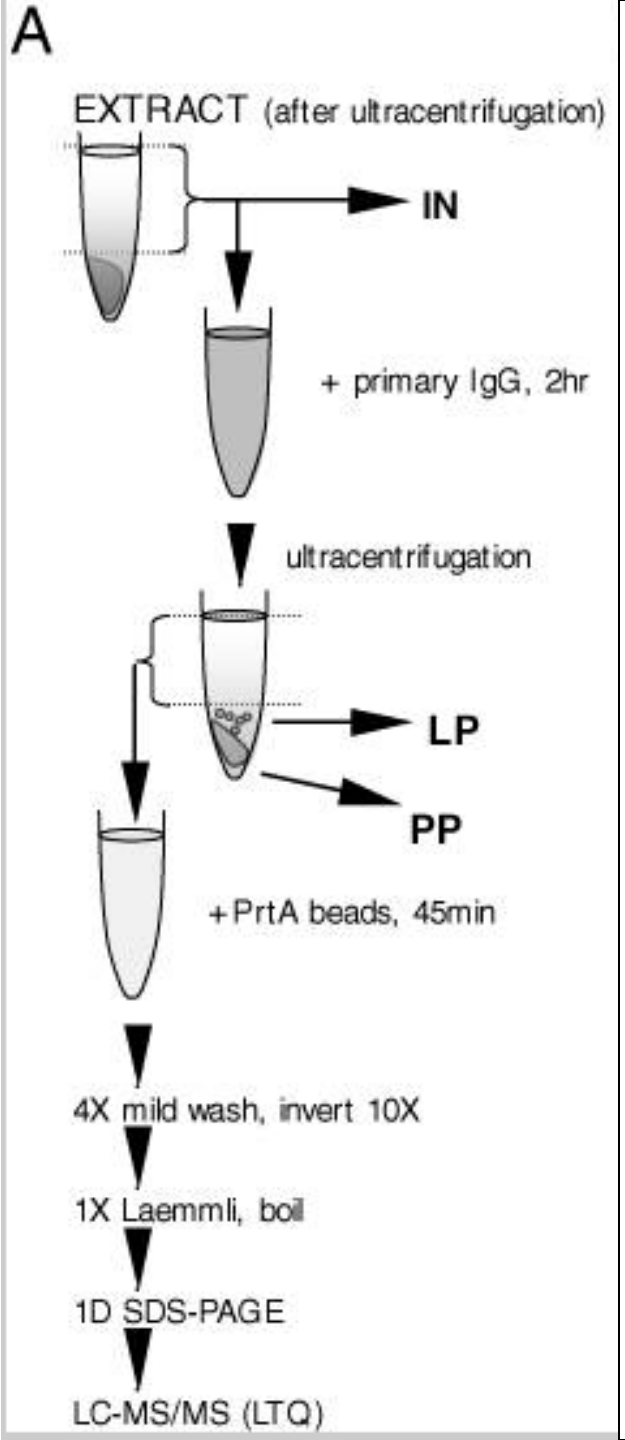


A cell lysate: 6,739 polypeptides



An immune complex of 1-10 polypeptides (why might there be more than one polypeptide?)

Rationale for affinity purification: Average K_d for an antibody-antigen complex is 10^{-12}M.



Issues in IMMUNOPRECIPITATION STUDIES until recently

NONSPECIFIC PROTEINS (see Left):

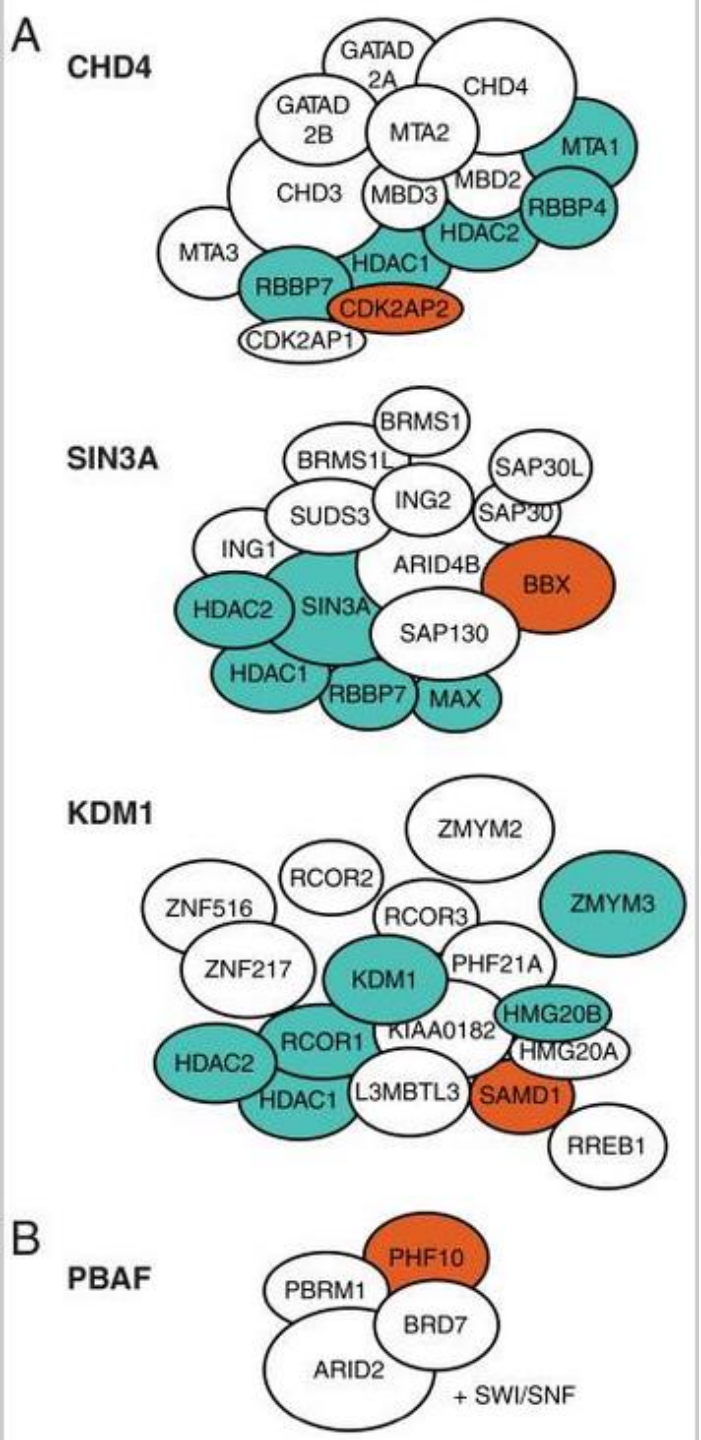
1.IN: input… proteins intrinsic to the sample that aggregate over time during the course of incubation with the antibody, and spin down with the immune complex: LP: loose pellet; PP: packed pellet

2. IgG-binding: may be similar across multiple samples

3. Bead-binding: may be similar across multiple samples

The above were re-visited, and MS used to identify nonspecific proteins specific to one experiment, vs across many IPs.

(taken from Malovannaya et al., 2010)



Immunoprecipitation coupled with mass spectrometry revealed multiple components in nuclear transcription complexes.

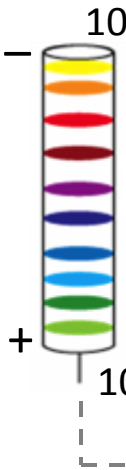
Antibodies to the indicated proteins were used in followup immunoprecipitation experiments to corroborate protein composition in related complexes.

(taken from Malovannaya et al., 2010)

GELFrEE technology can generate intact proteins between 50-5 kDa, for direct analysis by MS

Cell lysate

(200 – 400 µg)



GELFrEE intact protein fractionation

(60-90 min separation)

In this study, 1000 proteins, and >3000 proteoforms were characterized.

Online nanocapillary LC-MS/MS

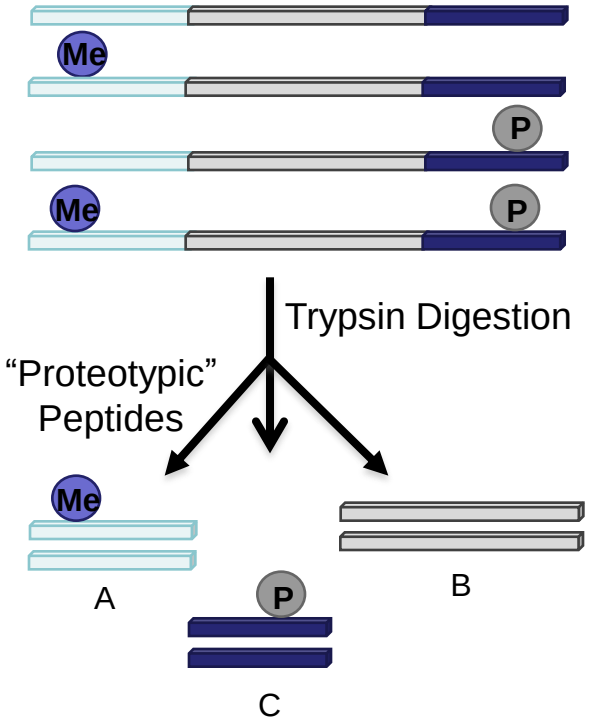
High-throughput search using ProSightPC

Tran, J.C. et al. *Nature* **2011**, 480, 254-258.

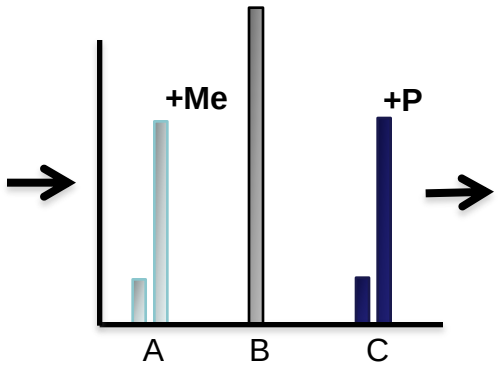
(Courtesy of Neil Kelleher)

Topdown vs bottom up, part I

One protein; four proteoforms



Bottom-Up MS analysis



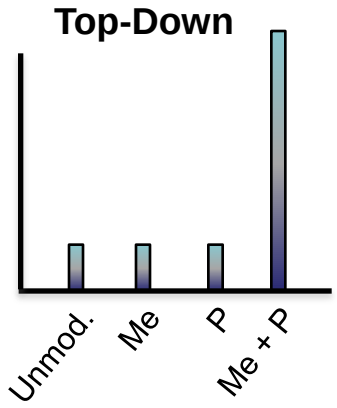
Ambiguous Result: Methylated and phosphorylated peptides of the protein were increased, but the specific proteoforms were unknown.

Savaryn *et al.* ***Genome Med.*** **2013 Jun 27;5(6):53.**

(Modified from N. Kelleher)

Topdown vs bottom up, part II

One protein; four proteoforms

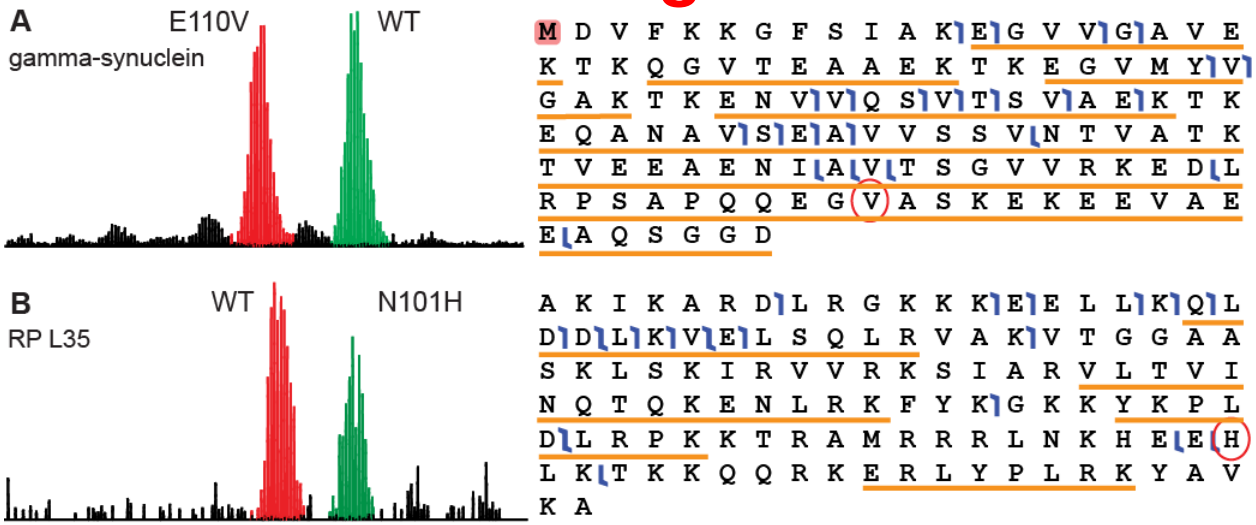


Meaningful Result: Proteoform with co-occuring methylation and phosphorylation is upregulated, but not the other three.

Savaryn *et al.* ***Genome Med.*** **2013 Jun 27;5(6):53.**

(Modified from N. Kelleher)

Sometimes both Topdown and Bottomup approaches are needed to confirm genomic information.



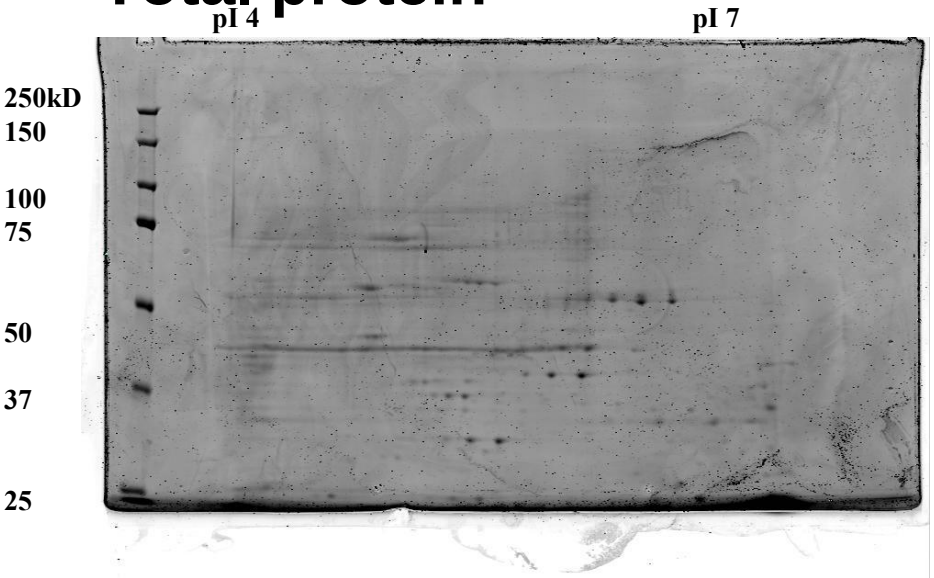
A) TD MS data of γ-synuclein showing the distinctive pattern of a heterozygote genotype at this locus, and the sequence of γ-synuclein detected by TD and peptide sequences (underlined) detected by BU. The highlighted N-terminal amino acid indicates an N-terminal acetylation. The cSNP E110V is circled. *Both technologies provided evidence of the cSNP*.

B) B) TD MS data of ribosomal protein L35 showing the distinctive pattern of a heterozygote genotype at this locus, and the sequence of L35 detected by TD and peptide sequences (underlined) detected by BU. The cSNP N101H is circled. *In this case, only TD provided evidence of the cSNP.*

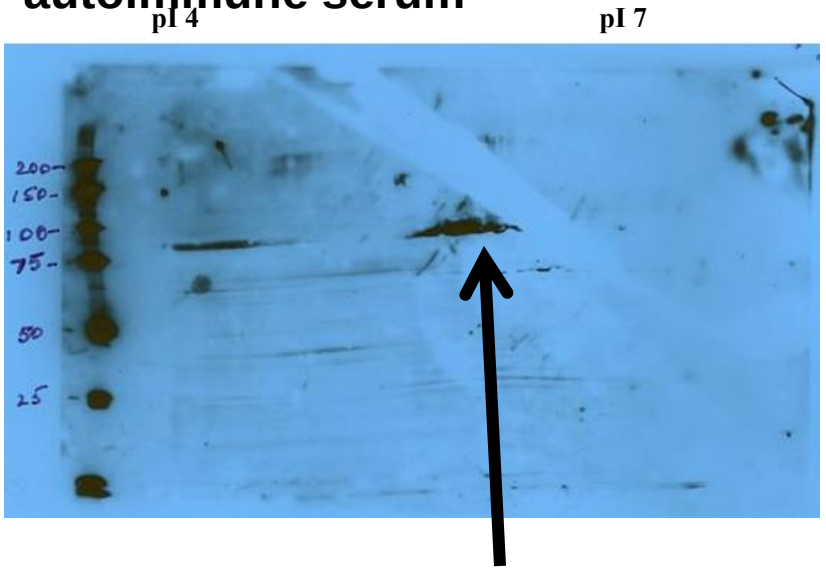
(Modified from N. Kelleher)

Topdown in immunology studies: 2D Western blot detection of the major antigen in an autoimmune disease

Total protein



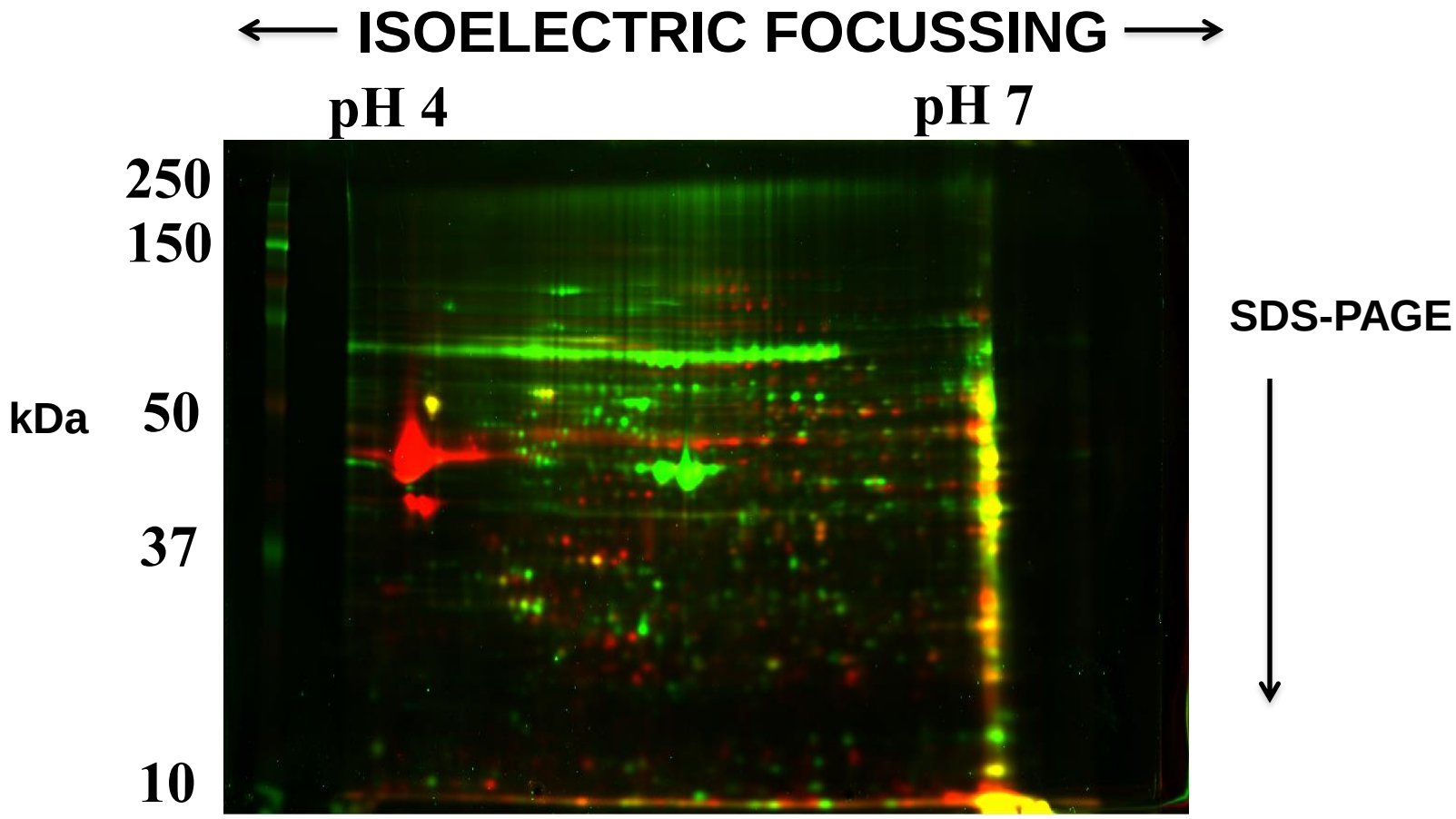
Western blot with mouse autoimmune serum



A protein spot(s) barely visualized in the total protein stain.

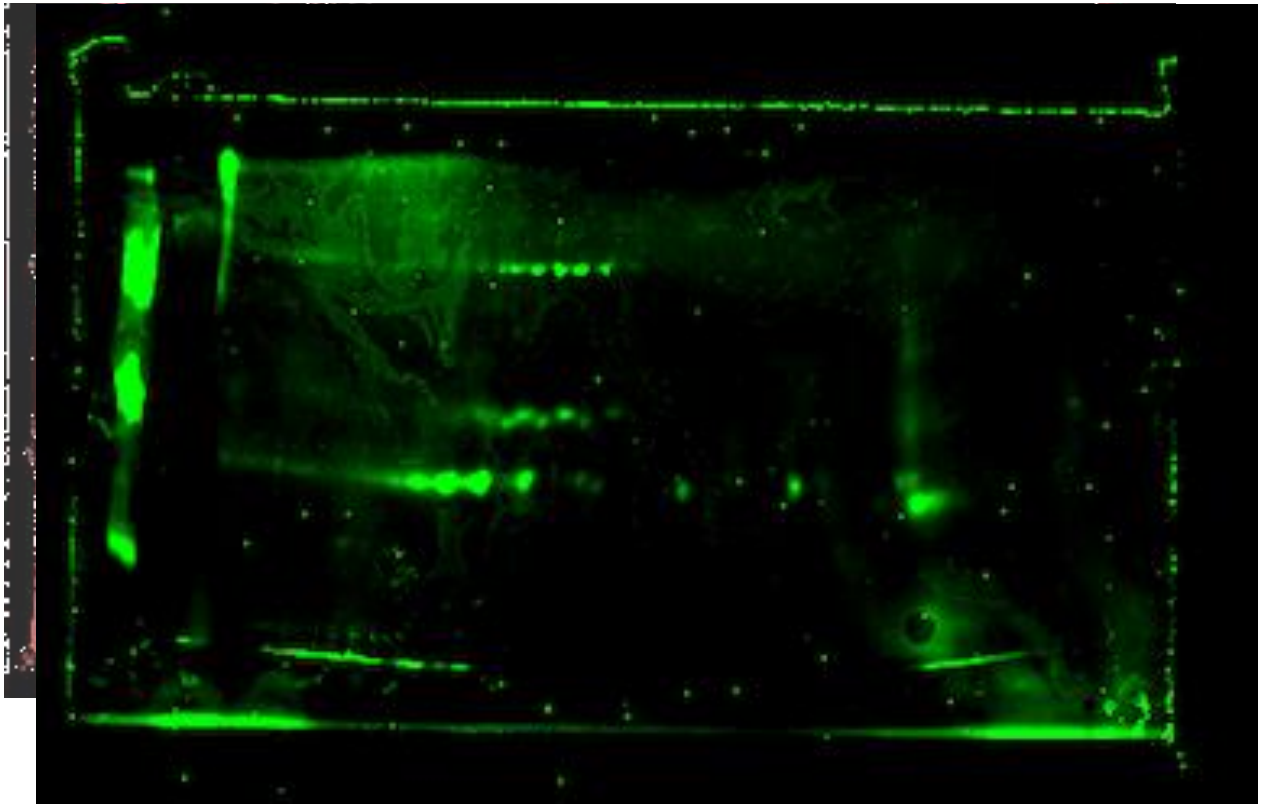
(Courtesy of M. Stoll)

2D DIGE: enables RAPID visual discovery of differences in abundance or PTM



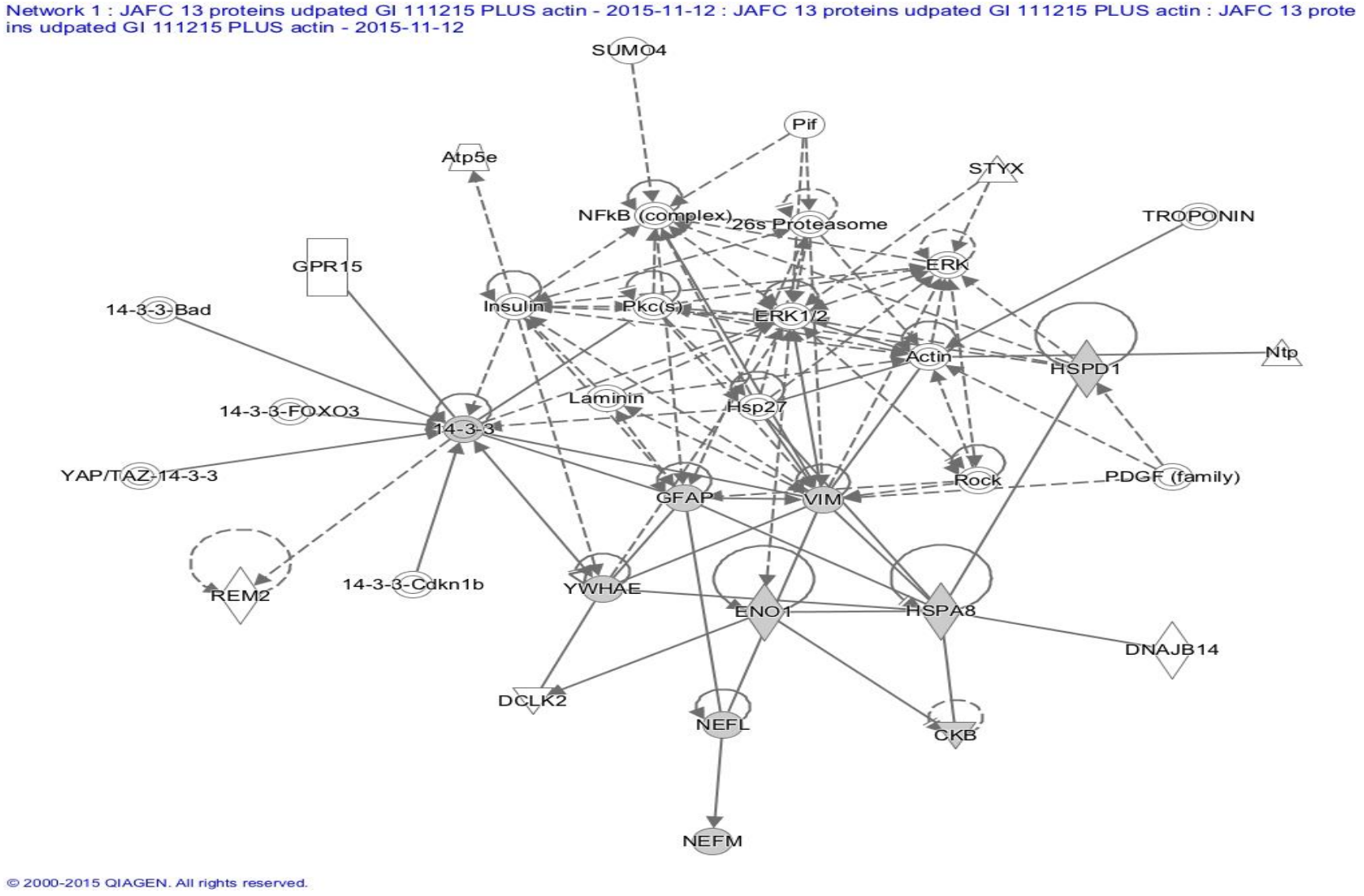
(Diane Bimczok, unpublished data)

Topdown Multiplex approach: with scarce samples



In MULTIPLEX analysis, one gel is stained for phosphoproteome, then total protein. Overlay of images indicates which proteins/proteoforms are phosphorylated, which are not.

A list of proteins that are changed can be enhanced through software-data mining using Ingenuity, i.e.



(Courtesy of D. Crossman)

How to address quality control in topdown proteomics, especially where multiple samples are involved

- **Minimize variance;**
  - **Start with collecting and storing samples:**
    - **If -80, always -80**
    - **Small aliquots to reduce number of thaws a sample undergoes**
  - **"Help" a labmate by letting him do all of one step, and you do all of another step… everyone pipets differently.**
- **Sometimes we can ELIMINATE VARIANCE:**
  - **If you freeze IEF strips overnight at -80, always freeze overnight at -80;**
  - **If you blot at 200V for 1.5 h, always blot at 200V for 1.5 h.**
  - **If you blot in the cold room, always blot in the cold room.**
- **Randomize sample analysis wherever possible; don't sacrifice all the control mice one day, and all the treatment groups the next day.**

Take home messages regarding protein purification

- **Goals in protein purification for proteomics analysis:**
  - **Reduce the complexity of the sample by taking advantage of extrinsic and intrinsic properties of the proteins**
  - **Be mindful of what questions you are asking…. Do you even need to run a 2D gel, or will a 1D (cheaper and faster) and SWATH MS suffice.**
- **Choice/extent of purification governed by**
  - **Abundance of sample**
  - **Abundance (if known) of protein in question**
  - **What technologies you can access readily**

Take home messages regarding protein purification, part II

- What proteomics technology gives back is like any other analytical approach: the quality is as good as what you put in;
- "Conventional" approaches like immunoprecipitation can be powerful when combined with MS;
- The "end result" in proteomics may just a beginning:

I. Some changes are causal to the disease/phenotype;

II. Some are "real" but not causal;

III. Some could be response of the cells/tissues TO DEAL WITH the disease, not causing the disease.

Suggested readings

Malovannaya, A, Li, Y, Bulynko, Y, Jung, S Y, Wang, Y, Lanz, RB, … O'Malley, B and Qin, J. (2010). ***Streamlined analysis schema for high-throughput identification of endogenous protein complexes.*** **Proc Natl Acad Sci, USA, 107(6), 2431–2436.**

Ntai, I, LeDuc, RD, Fellers, RT, Erdmann-Gilmore, P, Davies, SR, Rumsey, J, Early, BP, Thomas, PM, Li, S, Compton, PD, Ellis, MJC, Ruggles, KV, Fenyö, D, Boja, ES, Rodriguez, H, Townsend, RR, and Kelleher, NL. ***Integrated bottom-up and top-down proteomics of patient-derived breast tumor xenografts.*** **Molec Cell Proteomics. published online Oct 26, 2015.**