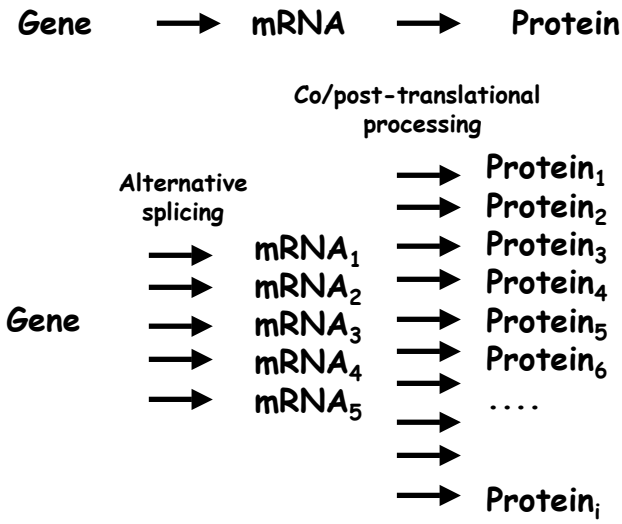
Applications of Mass Spectrometry to Proteomics

Akhilesh Pandey, M.D., Ph.D.

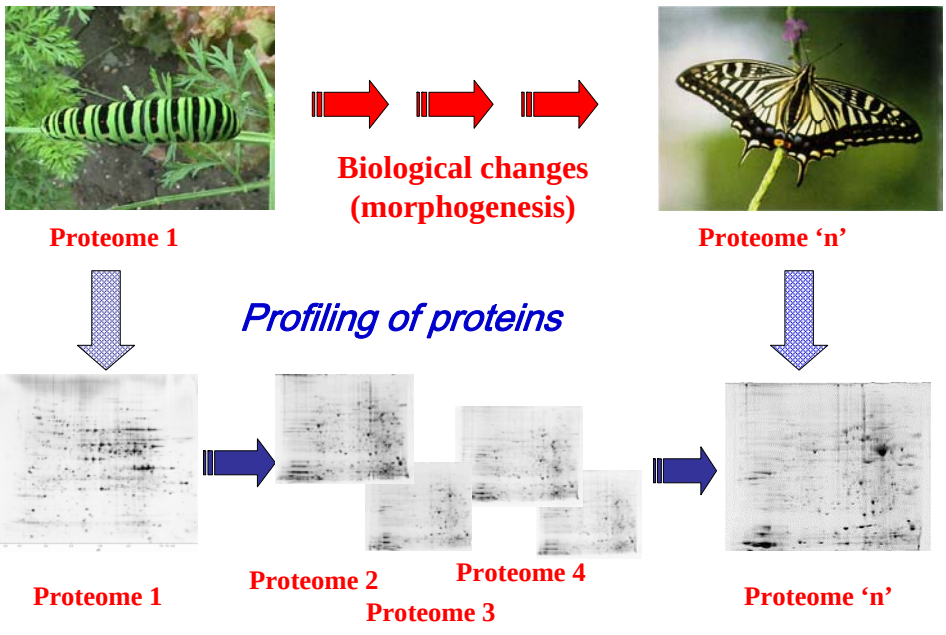
McKusick-Nathans Institute of Genetic Medicine and the Department of Biological Chemistry

Why Proteomics?

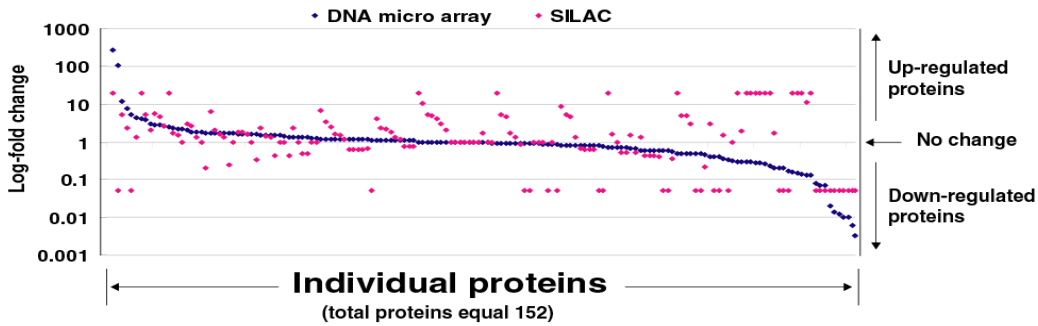
One Gene, Many Proteins



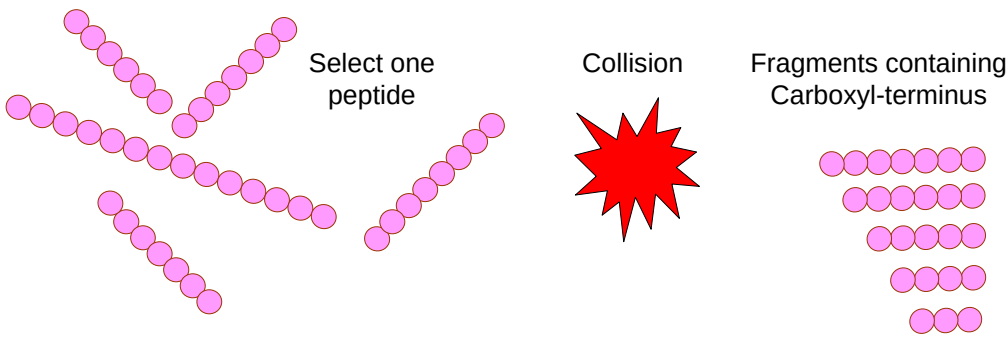
Genomics and Proteomics



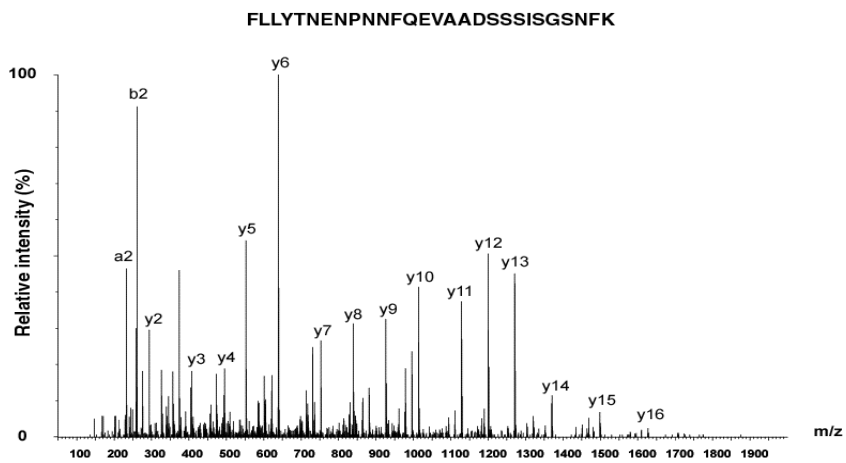
mRNA and Protein Correlation

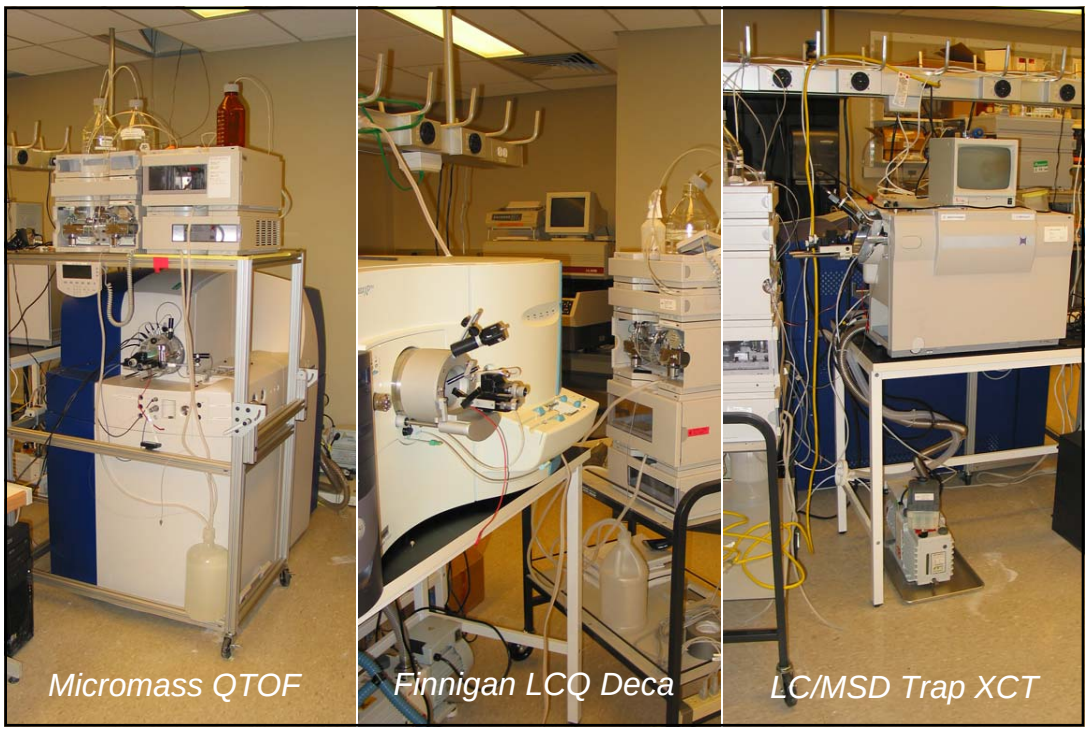


Peptide Sequencing by MS/MS

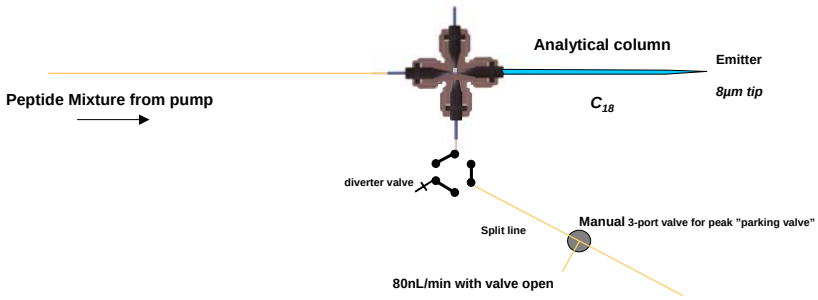


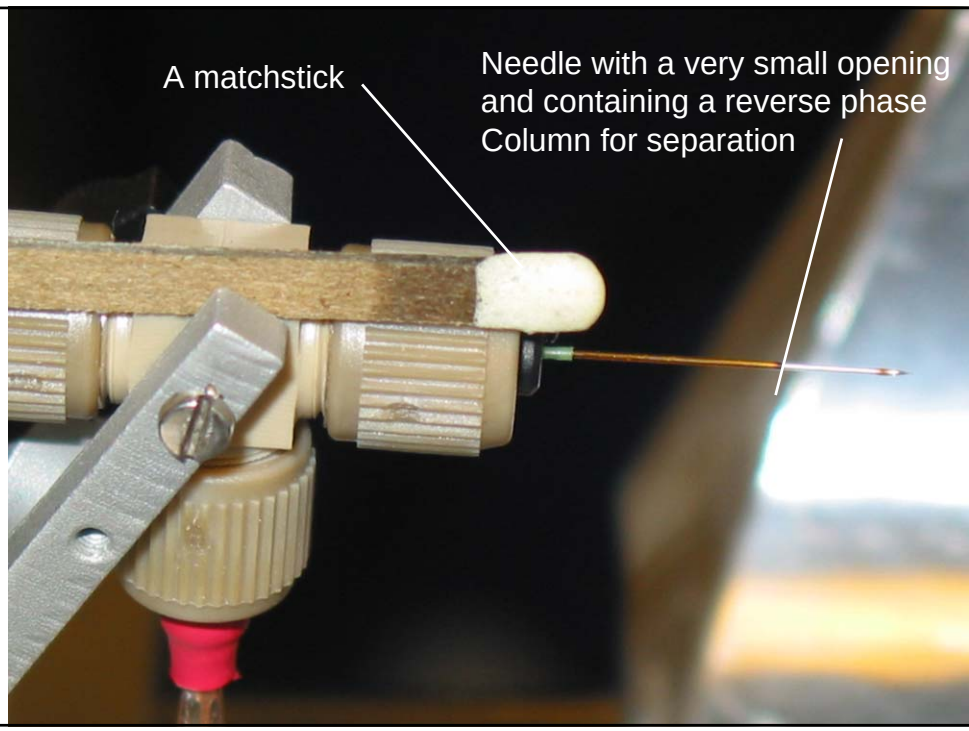
MS/MS spectrum (sequencing)



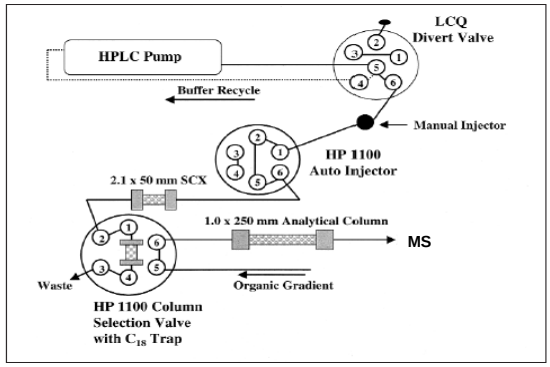


1D Liquid Chromatography Setup



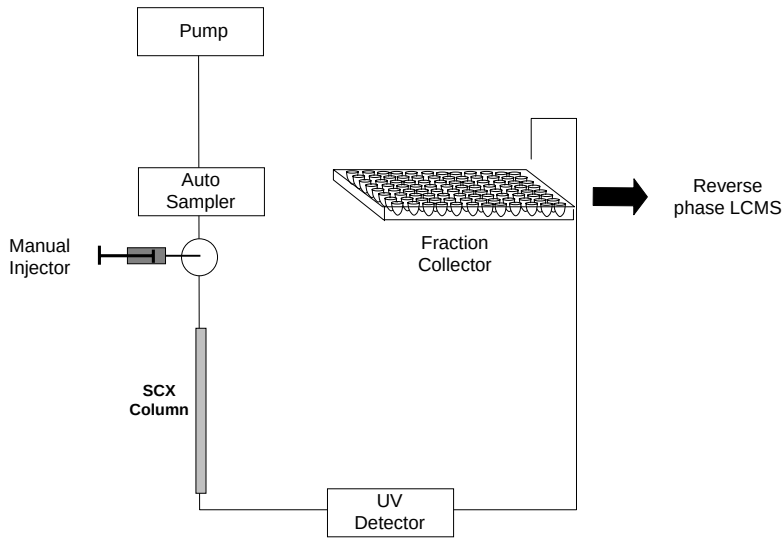


Multidimensional Liquid Chromatography (MudPIT) Setup

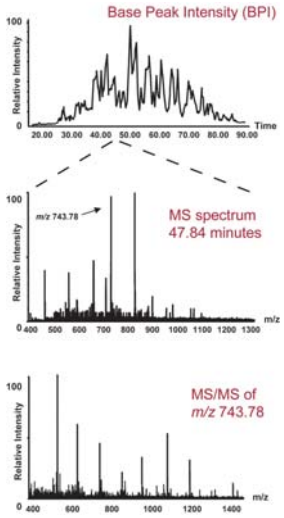


Liu, H. et al. (2002)

Offline Fractionation in the First Dimension



Automated nanoLC-MS/MS



- Purification of sample
- Very complex samples can be analyzed
- Identical peptides will elute in one peak (enhanced signal)
- Automated
- Many proteins can be identified in one run (100-2,000 proteins)

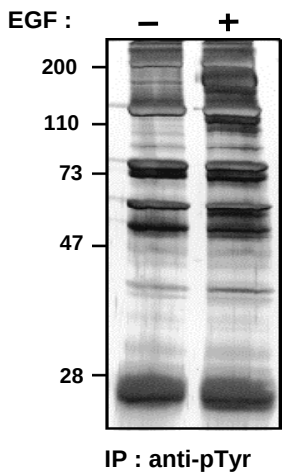
Quantitative Proteomics

- Based on difference in intensity
- By relative quantitation using MS

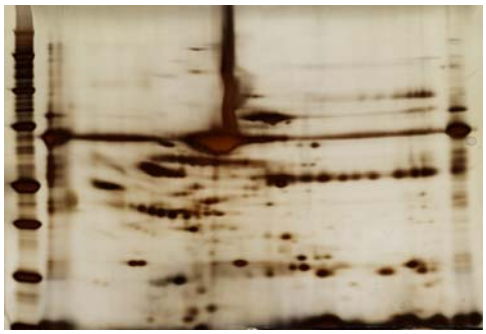
Quantitative Proteomics

- Based on difference in intensity
  - 1D gels
  - 2D gels
  - Use of fluorescence-based methods

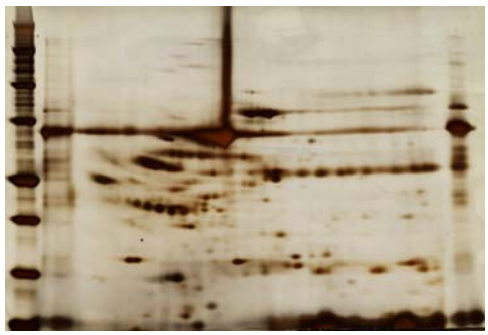
1D Gel-based Comparison



2D Gel-based Comparison



Normal



Cancer

2D Gels - Limitations

- Sample preparation - lot of optimization required
- Loading capacity limited
- Do not resolve very small (<10 kD) or large (>100 kD) proteins
- Do not resolve hydrophobic (e.g. membrane) proteins
- Issues with reproducibility

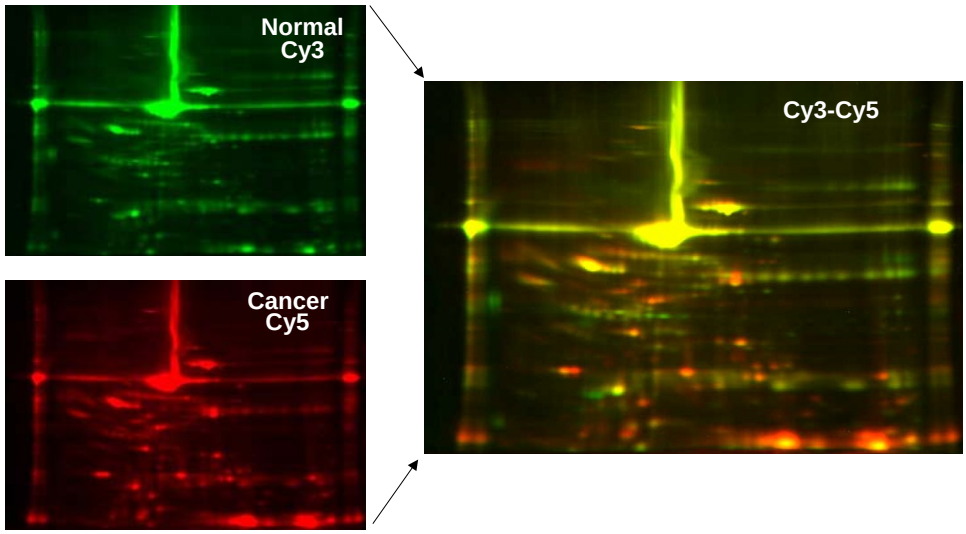
Quantitative Proteomics

- Fluorescence-based quantitation
  - DIGE (Difference in-gel electrophoresis)

DIGE

- Samples to be labeled are labeled with Cy3 (green) and Cy5 (red)
- Samples are ‘mixed’ and resolved by 2D gels
- Fluorescence measured and quantitated

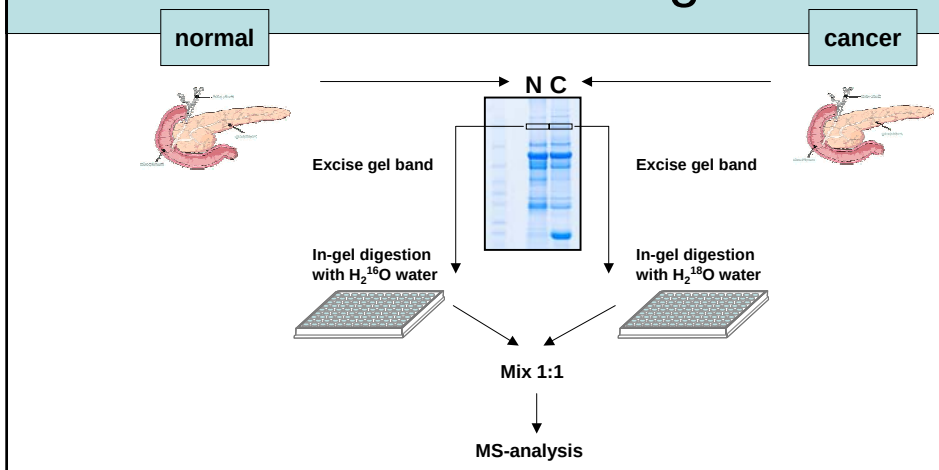
2D-DIGE of Pancreatic Juice



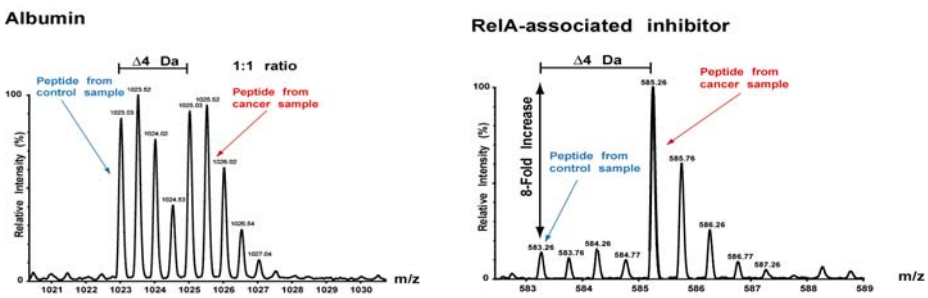
Quantitative Proteomics

- By relative quantitation using MS
  - *in vitro* labeling
    - ^{18}O-labeling
    - Peptide mass tagging (ICAT)
  - *in vivo* labeling
    - Labeling with stable isotope containing amino acids (SILAC)

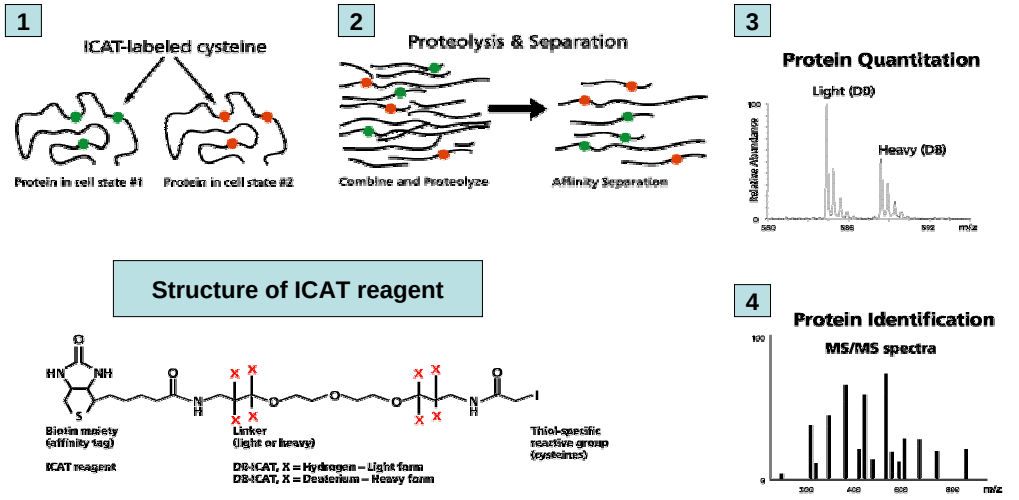
^{18}O-labeling



^{18}O-labeling



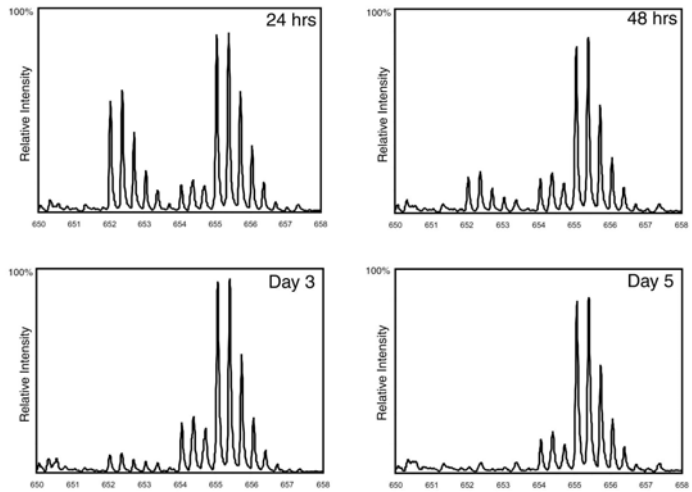
Isotope-Coded Affinity Tag (ICAT)



Stable Isotope Labeling in Cell Culture (SILAC) for Protein Quantitation

- Mammalian cell culture models are used for studying a number of biological processes
- In the SILAC approach, cells are grown continuously in media containing one or more stable isotopes (e.g. ^{13}C). All the proteins in the cells are heavier and can be used to 'mark' a given state in mass spectrometric analysis

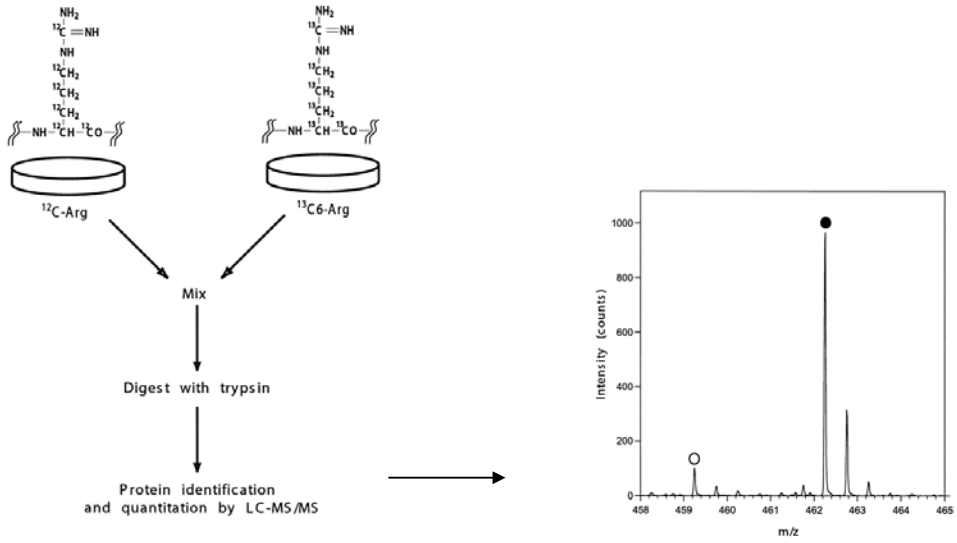
Time Course of Heavy Amino Acid Incorporation



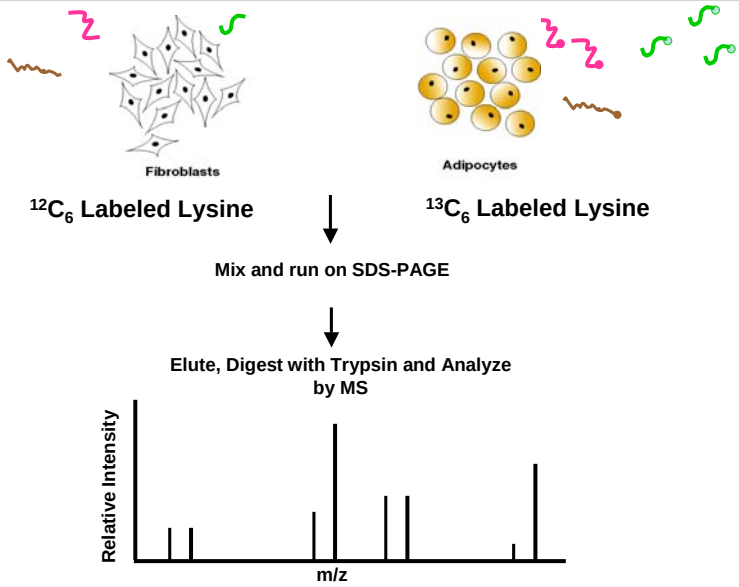
Advantages of the SILAC Method

- Simple
- *In vivo*
- Does not require any extra processing steps
- All proteins are uniformly labeled
- Complete and predictable incorporation
- Choice of labeled amino acids
- De novo sequencing can be performed

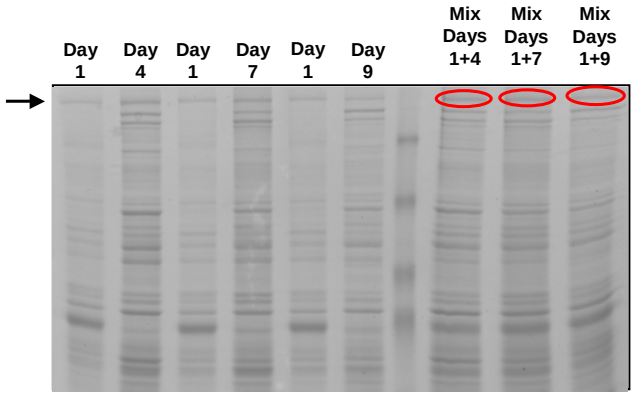
General strategy for stable isotope labeling by amino acids



SILAC for Quantitation of Secreted Proteins

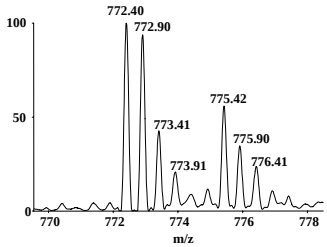


Profile of Proteins Secreted by Adipocytes

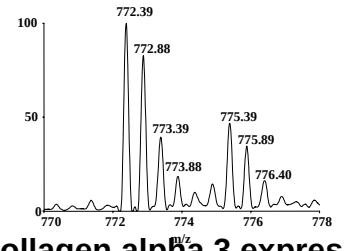


Fibronectin expression is downregulated

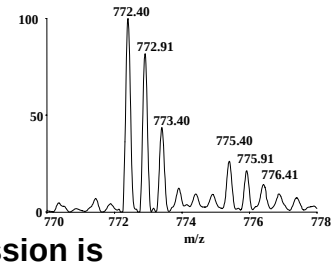
Ratio 1:0.5



Ratio 1:0.4

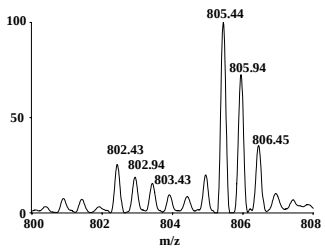


Ratio 1:0.24

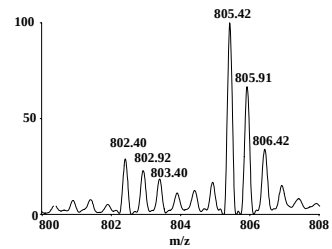


Collagen alpha 3 expression is upregulated

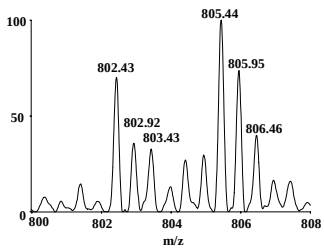
Ratio 1:2.4



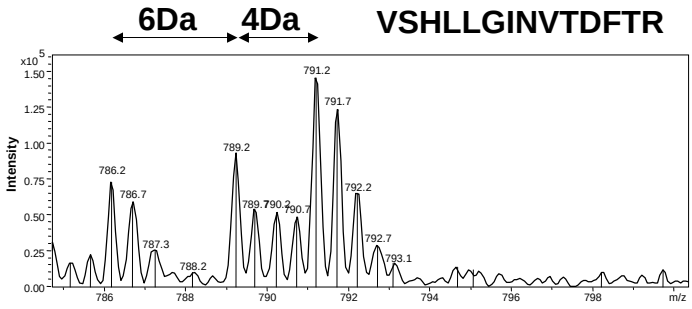
Ratio 1:2.7



Ratio 1:1.6



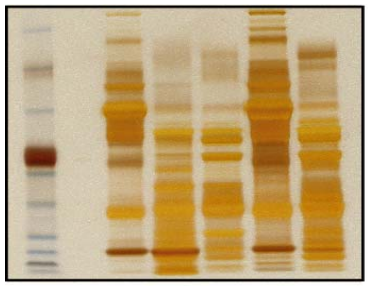
Studying Dynamics Using SILAC



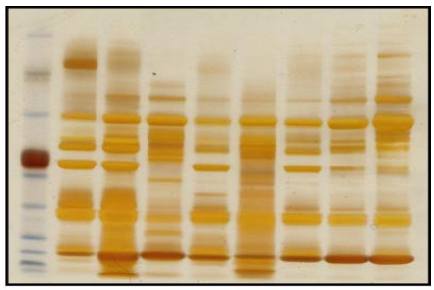
Biomarker Discovery Using Proteomics

- Ideal targets for biomarker
  - Protein (differential expressed proteins)
  - DNA (mutations, methylation)
  - RNA (differential expressed genes)
- Biological specimen
  - Tissue (whole tissue or isolated tumor cells)
  - Pancreatic juice
  - Serum
  - Plasma

Pancreatic juice

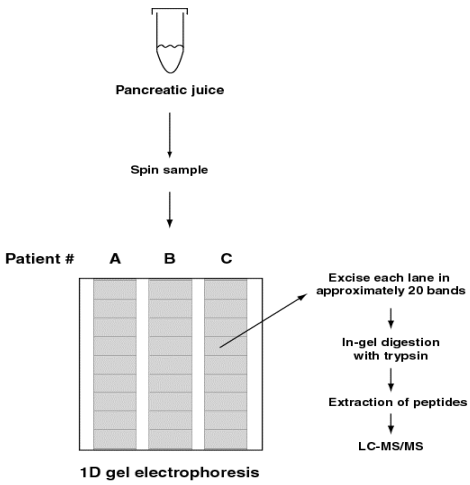


Chronic Pancreatitis



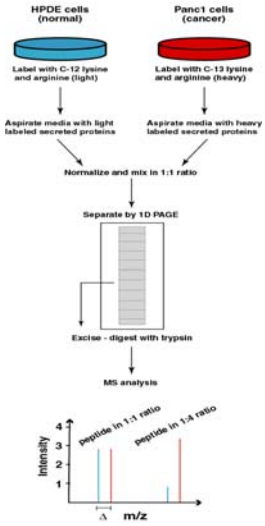
Pancreatic Cancer

Physiological proteome of human pancreatic juice



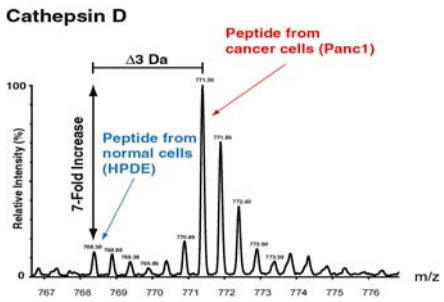
- Collection of pancreatic juice (cancer) during surgery
- Run on 1D gel
- LC-MS/MS analysis
- Bioinformatic analysis of identified proteins
- Compare data with known microarray data

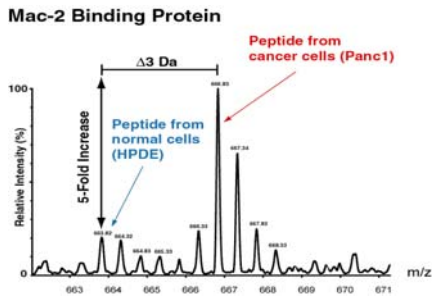
Differential Proteomic Analysis of Pancreatic Cancer Secretome



- *In-vivo* labeling with both arginine and lysine
- Collect conditioned media and concentrate with centricon 3,000 Da MWCO
- Normalize and mix in 1:1 ratio
- Resolve proteins by 1D gel electrophoresis
- Excise bands and digest proteins by trypsin
- Identify proteins by nanoLC-MS/MS (2x30 bands)
- Verify identified proteins (manually)
- Relative quantitation of ID proteins (manually)

Quantitation of Secreted Proteins





Post-translational Modifications

- Peptides can have a number of modifications
- During database searching, a variable modification has to be specified – otherwise, no 'hit'
- Common PTMs are phosphorylation, acetylation, ubiquitination, glycosylation etc.

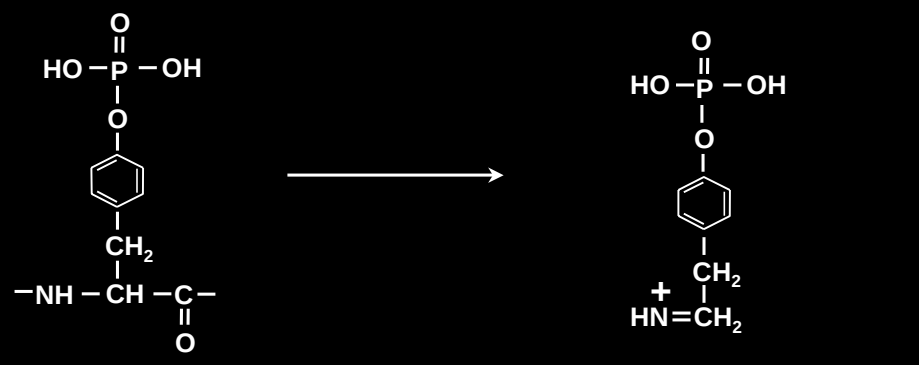
Protein Phosphorylation

- One-third of all cellular proteins are phosphorylated at one time or another
- Phosphoamino acid content of a vertebrate cell:
  Serine - 90%; Threonine - 10%; Tyrosine - 0.05%
- Ser:Thr:Tyr - 1800:200:1
- Tyrosine phosphorylation is tightly regulated

Why is Phosphorylation Analysis Difficult?

- Stoichiometry of phosphorylation is low
- Complete coverage of proteins is difficult to obtain
- Phosphorylated serine and threonine residues are labile whereas phosphotyrosines are more stable
- Phosphoserines and phosphothreonine residues can be subjected to a beta-elimination reaction but not phosphotyrosine residues
- Antibodies to enrich for serine and threonine phosphorylated proteins are not available
- Phosphopeptides are ‘suppressed’ in a mass spectrum

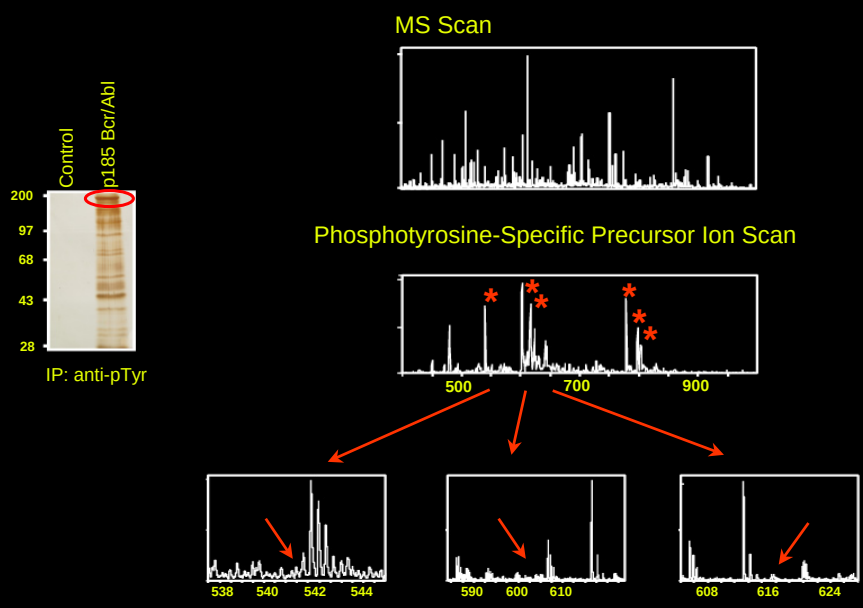
Immonium Ion of Phosphotyrosine as a Reporter Ion



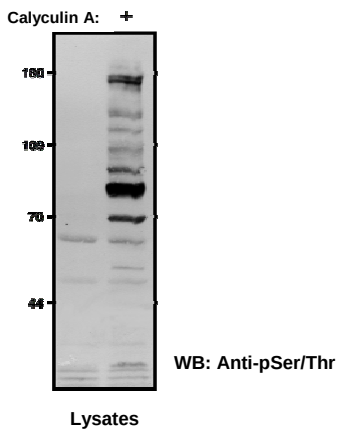
Phosphotyrosine

Immonium Ion (216.043 Da)

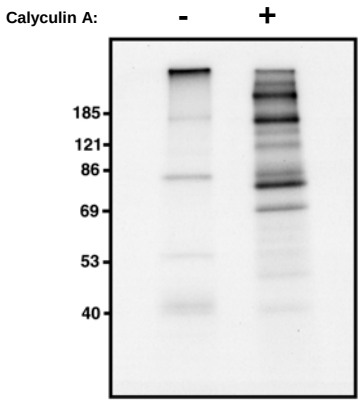
Phosphotyrosine-Specific Precursor Ion Scanning – Bcr/Abl



Large scale IP with an anti-pSer/Thr antibody

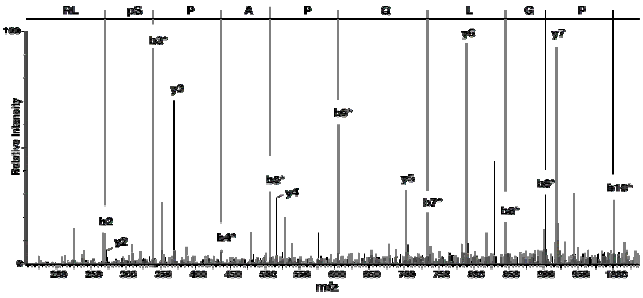


In Vivo Labeling with ^{32}P



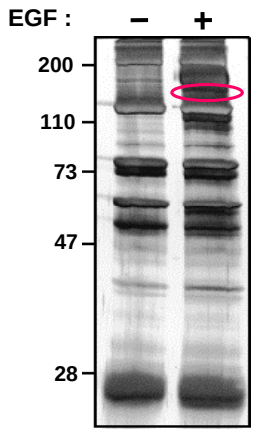
IP: Anti-pSer/Thr

Identification of Phosphorylated Ser/Thr residues



^{10}RLpSPAPQLGP^{19}

MS-Based Identification of a 130 kDa Protein in the EGF Receptor Signaling Pathway



Silver stained gel

IP : anti-pTyr

Assignment of the initiator methionine in a cDNA 'fragment' based on an N-terminal peptide

>KIAA0229 (1180 residues) FRAGMENT

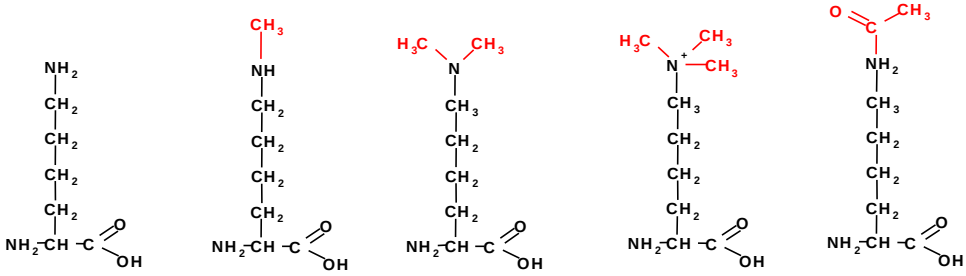
SWGKGREGVVSPAGLGGALPGDGKFGSPSRLGCSLGEGVQRVAALGMGKEQ
ELLRAARTGHLPAVEKLLSGKRLSSGFGGGGGGGSGGGGGGSGGGGGGLGS
SSHPLSSLLSMWRGPNVNCVDSTGYTPLHHAALNGHHRRSSSSRSQDSAEGQ
DGQVPEQFSGLLHGSSPVCEVGQDPFQLLCTAGQSHPDGSPQQGACHKASM
QLEETGVHAPGASQPSALDOSKRVGYLTGLPTTNSRSHPETLTHTASPHPGGA
EEGDRSGAR

Assignment of the initiator methionine in a cDNA 'fragment' based on an N-terminal peptide

>KIAA0229 (1180 residues) FRAGMENT

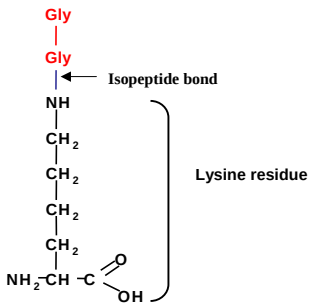
SWGKGREGVVSPAGLGGALPGDGKFGSPSRLGCSLGEGVQRVAALG**MGKEQ**
LLRAARTGHLPAVEKLLSGKRLSSGFGGGGGGGSGGGGGGSGGGGGGLGSS
SHPLSSLLSMWRGPNVNCVDSTGYTPLHHAALNGHHRRSSSSRSQDSAEGQD
GQVPEQFSGLLHGSSPVCEVGQDPFQLLCTAGQSHPDGSPQQGACHKASMQL
EETGVHAPGASQPSALDQSKRVGYLTGLPTTNSRSHPETLTHTASPHPGGAEE
GDRSGAR

Lysine Modifications



	Lysine	Mono-methylated lysine	Di-methylated lysine	Tri-methylated lysine	Acetylated lysine
Mass (Da)	128.095	142.111	156.127	170.143	170.105
Mass gain	-	14.016 Da	28.032 Da	42.048 Da	42.010 Da

Lysine Modifications

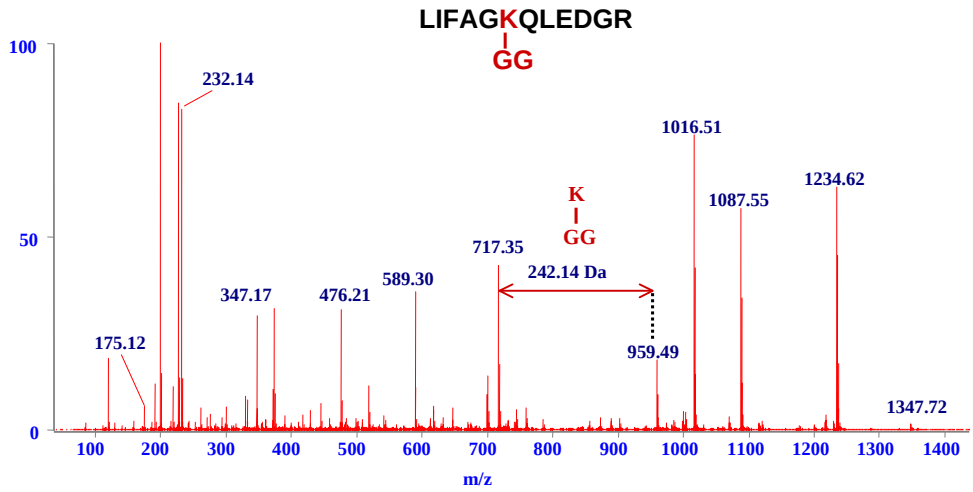


Di-glycyl lysine

Mass gain

114.05 Da

Tryptic Digest of Ubiquitylated Peptide



Signature of a Ubiquitylated Peptide