What Follows 'Simple' Protein Identification?

Protein Characterization

- Defining N- and C-termini
- Classification of splice variants
- Characterization of protein modifications

Protein Modifications

- Identification of modified protein
- Localization of modification
- Structure elucidation of the modification

The second level of proteome analysis - analysis of modified proteins

Some of the most common modifications

Acylation
- N-terminus
- Lysine

Phosphorylation
- Serine
- Threonine
- Tyrosine
- (Aspartic acid, histidine and lysine)

Proteolytic processing
- Specific
- Non-specific
- C- and/or N-terminal

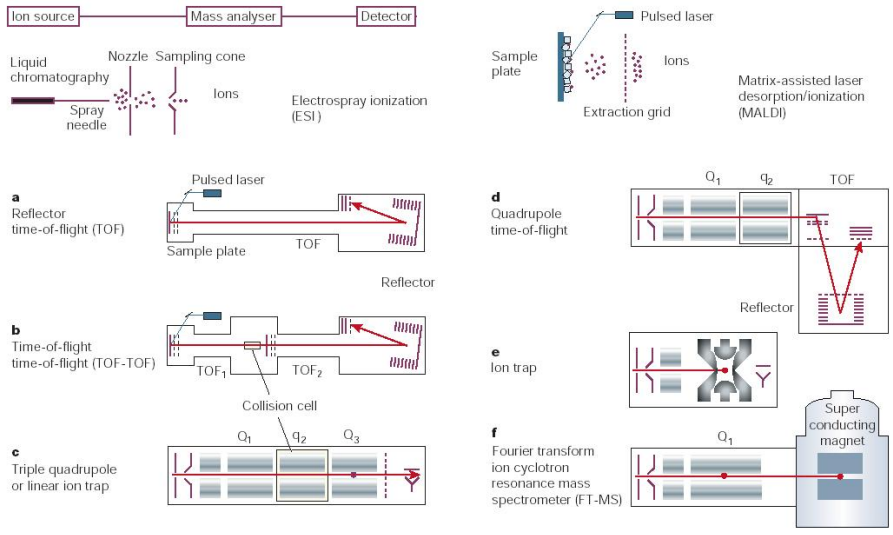
Glycosylation
- O-linked (Serine, Threonine, ...)
- N-linked (Asparagine)

"Glyco-lipids"
- Glycosyl-phosphatidylinositol anc
- Farnesyl anchor

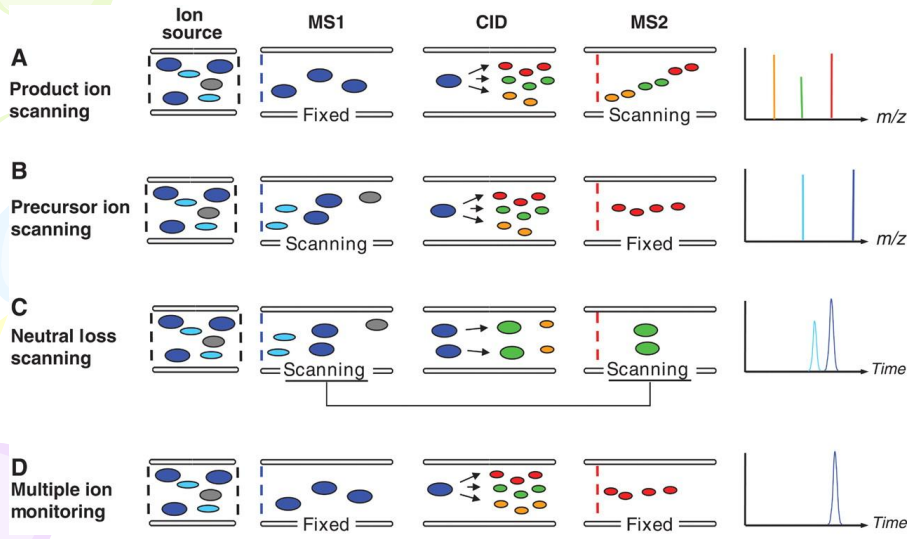
Analysis of modified proteins

- Specific detection in gels
  - Radiolabeling
  - Fluorescent labeling
  - Western blotting
  - Modification specific stains
- Affinity fishing
  - Immune precipitation of proteins
  - Affinity purification of proteins/peptides
    - -phosphopeptide isolation: IMAC, TiO_2
    - -phosphoprotein isolation: phosphospecific antibodies
  - Selective tagging followed by affinity purification
- Selective mass spectrometry
  - Precursor ion scanning
  - Neutral loss scanning
  - Stable isotope labeling

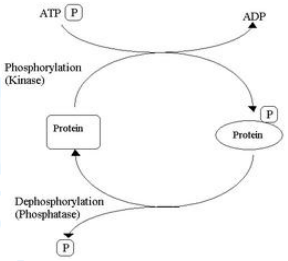
Mass analyser types



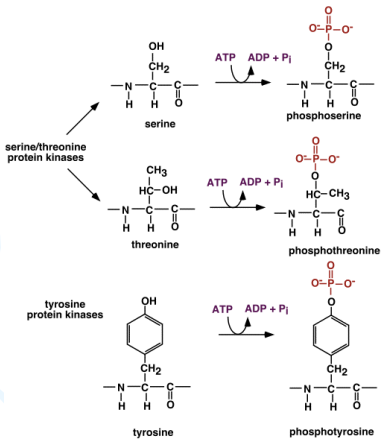
Tandem mass spectrometry scan types



Protein phosphorylation



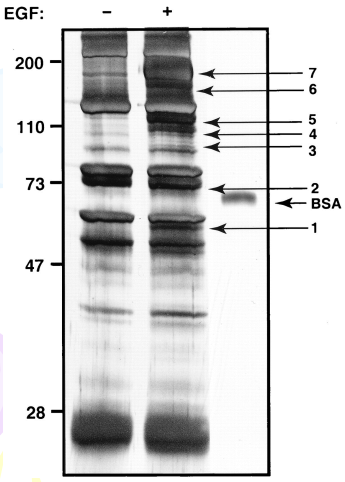
- Phosphorylation is a reversible modification, afffecting both the folding and function of proteins
- Phosphorylation is involved in:
  - Enzyme activation / inhibition
  - Protein-protein interaction
  - Protein degradation
  - Signal transduction
- Phosphorylation regulates:
  - Cell growth
  - Cell differentiation
  - Migration
  - Metabolism
  - apoptosis



- Ratio of pSer/pThr/pTyr is 1800:200:1
- Phosphorylation is usually not 100%

Identification of EGF signaling molecules by immunoprecipitation and mass spectrometry

Immunoprecipitation with anti-phosphotyrosine antibody



- **EGF Receptor**
- **Hrs***
- **Cbl***
- **Eps15**
- **p62**
- **Shc**
- **p85 subunit of PI 3-Kinase**
- **SHIP-2**
- **STAM***
- **Vav-2**
- **STAM2***
- **Odin**
- **Ku 70 Autoantigen**
- **Hsp 70**

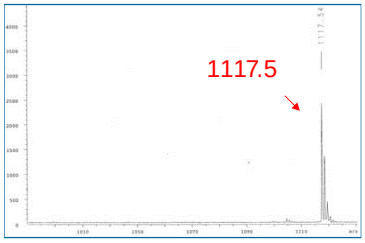
Pandey *et al* (2000) *PNAS* 97 179-184
***Steen *et al* (2002) *J. Biol. Chem.* 277 1031-9.**

Phosphorylation site analysis by MALDI-TOF

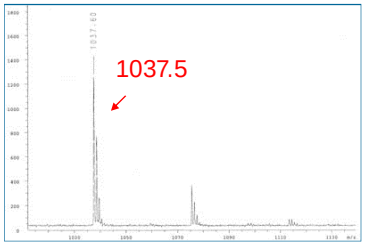
- Digestion of phosphorylated protein by trypsin
- MALDI-TOF mass analysis of peptides and comparison
  - to database calculated masses
  - to nonphosphorylated peptide masses
- To find a peptide with a mass addition of 80 (=HPO_3)
- Phosphatase treatment of the sample on the MALDI plate should result in the disappearance phosphate group

Phosphatase treatment of a phosphorylated peptide on the MALDI plate

- Peptide
  ALIHLSDLR
  $[M+H]^+$ 1037.5
- Phosphorylated peptide
  ALIHLpSDLR
  $[M+H]^+$ 1117.5



Mass spectra of the phosphorylated peptide



Mass spectra of the same sample after phosphatase treatment on the MALDI plate

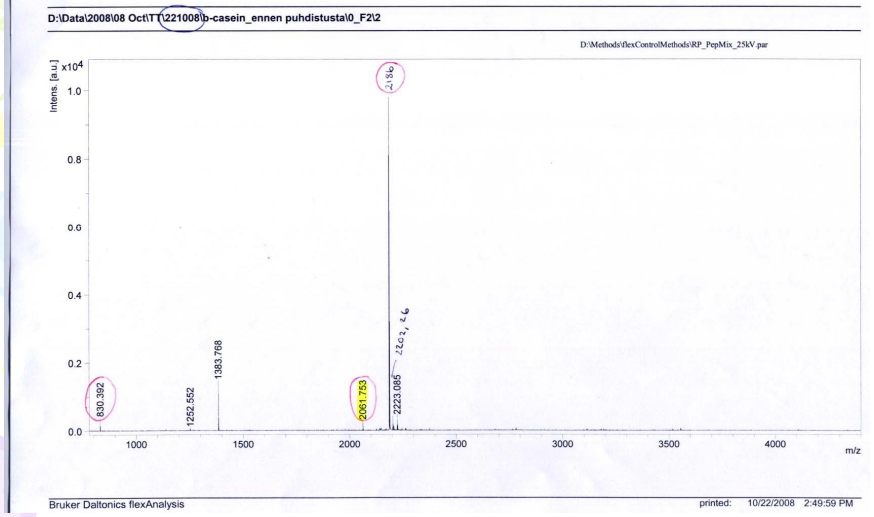
Problems in phosphorylation site analysis (by MALDI-TOF)

- Phosphorylated peptides ionize poorly in a positive mode
- Phosphorylated peptides are supressed in a total protein digest analyses
  - enrichment of phosphopeptides
  - prefractioning the digest into peptides (HPLC)

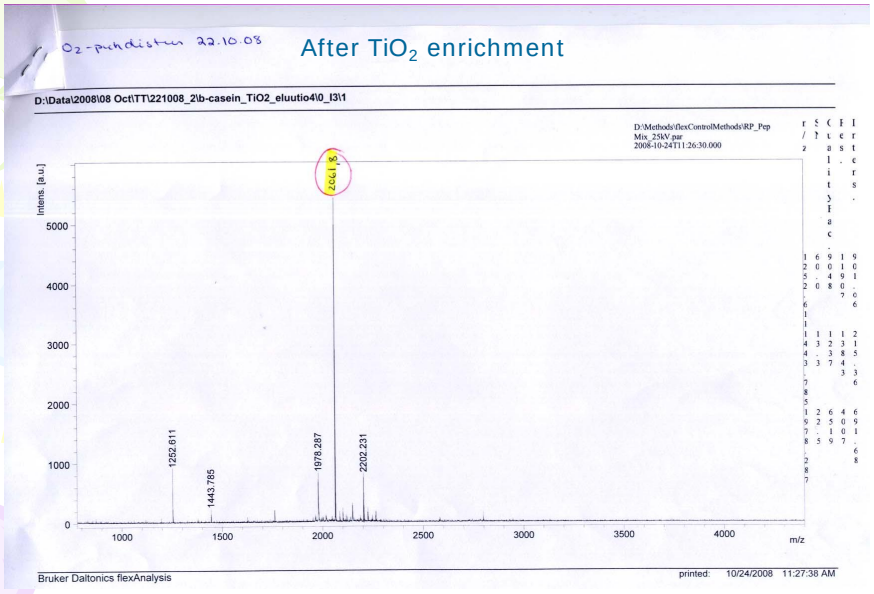
Selective enrichment of phosphopeptides

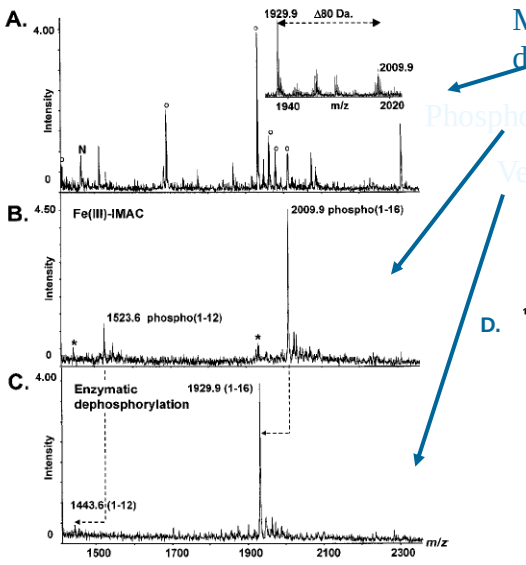
- IMAC (Immobilized metal-ion affinity chromatography)
- SAX + IMAC
- SCX + IMAC
- TiO_2
- SCX + TiO_2
- HILIC (Hydrophilic interaction chromatography) + IMAC
- IMAC + TiO_2 (SIMAC)
- α-pTyr ab + IMAC
- ZrO_2, other metal oxides
- Calcium-precipitation
- Phosphoramidate chemistry

Before TiO_2 enrichment



After TiO_2 enrichment



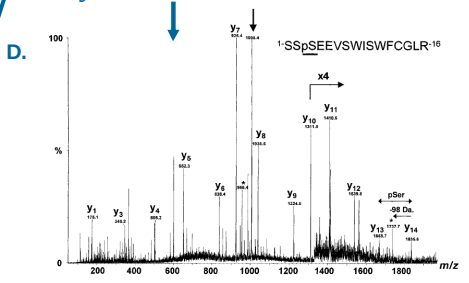


MALDI mass map of in-gel digested protein (CK2)

Phosphopeptides enriched using IMAC

Verification of phosphorylation

followed by sequencing by ESI-MS/MS



A. Stensballe & O.N. Jensen

Proteomics levels

Expression proteomics

Which gene products are expressed, when and how much

PTM-omics, "Modificomics"

Which variants are present of each protein, when and how much

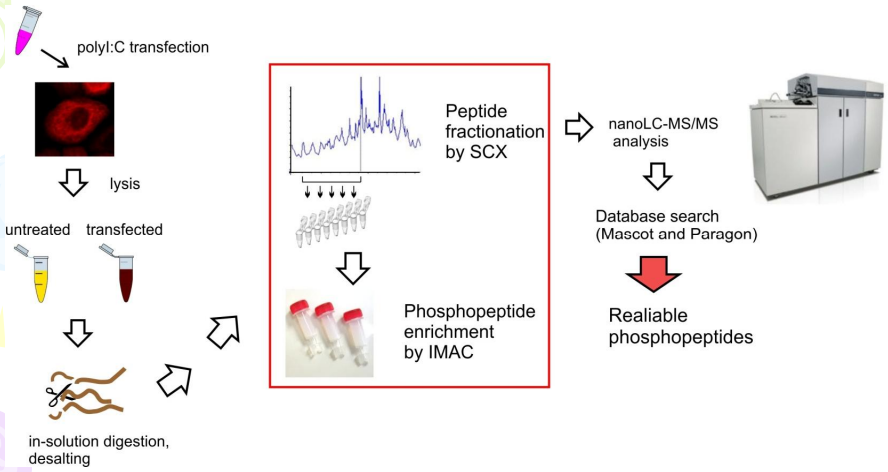
Cell map proteomics,"Interactomics"

Who interacts, when and where

PHOSPHOPROTEOMICS

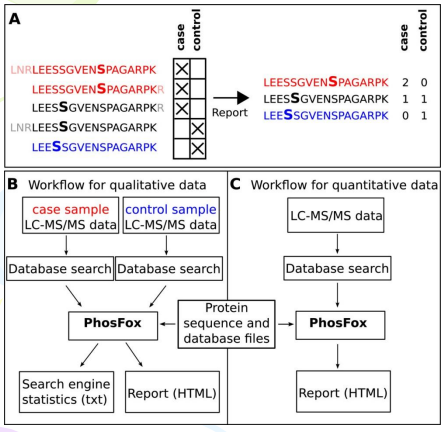
- The human proteome is estimated to include up to 500,000 phosphorylation sites
- The advances in phosphopeptide enrichment procedures and high-throughput mass spectrometry instrumentation have led to rapid development of MS-based phosphoproteomics during the last few years, and currently thousands of phosphorylation sites can be detected from a single sample.
- Current bottleneck is the downstream data analysis which is often laborious and requires a number of manual steps

Phosphoproteomics workflow



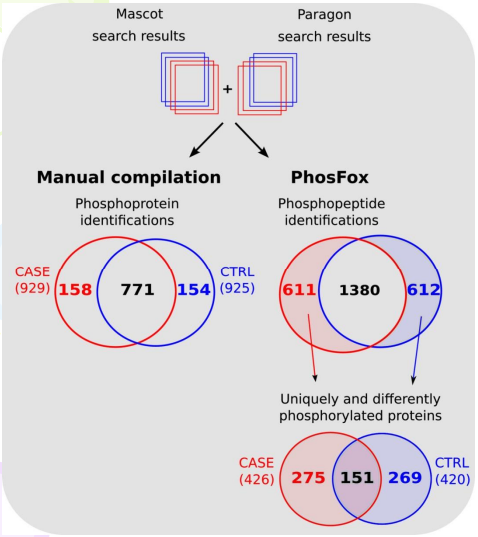
CHALLENGES IN PHOSPHOPROTEOME DATA ANALYSIS

- Compared to protein identification in phosphoproteomics it is essential to determine also the phosphorylation site of the modified peptide/protein.
- Phosphorylation site localisation is fairly easy for individual protein identifications but poses additional challenges for large-scale proteome studies.
- Protein/peptide can be differentially phosphorylated e.g. upon viral infection, and these changes are not easily found from the identification lists.
- Need to estimate both false identification rate AND false localization rate



A 'uniquely phosphorylated peptide' is a phosphopeptide with a unique phosphorylation or phosphorylations either in the case or control sample

The term 'uniquely phosphorylated protein' is used for describing a protein with at least one uniquely phosphorylated peptide, which has uniquely been matched to that particular protein in a particular sample.



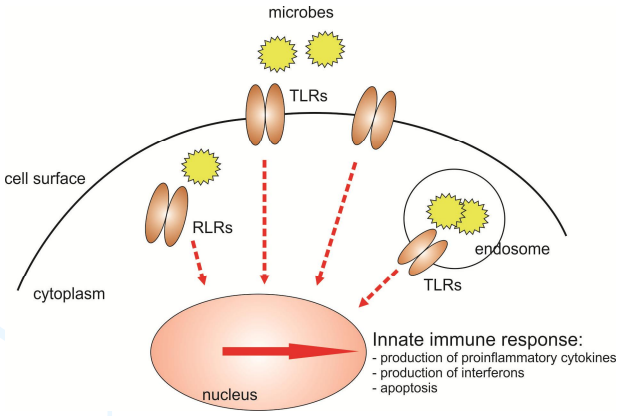
PhosFox: a bioinformatics tool for peptide-level processing of LC-MS/MS-based phosphoproteomic data

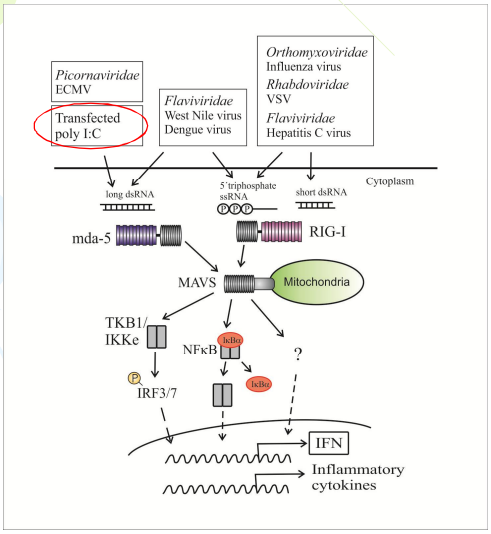
Söderholm *et al*, 2014

Functional proteomics to identify new regulators of antiviral immune response

Tiina Öhman
Institute of Biotechnology
Tuula Nymans group

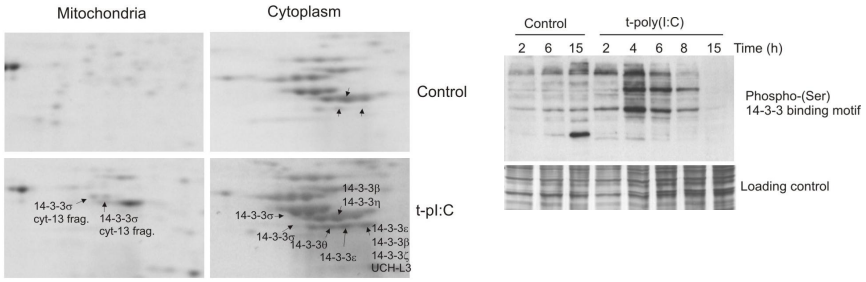
Innate immune response





Viral double-stranded RNA (dsRNA) is the most important viral structure involved in activation of innate immune response

Cytosolic RNA recognition pathway activates 14-3-3 protein mediated signaling and caspase-dependent disruption of cytoskeleton network in human keratinocytes



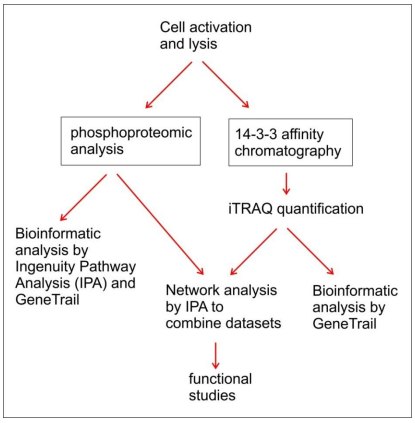
Öhman et al, 2010

Aims of this study:

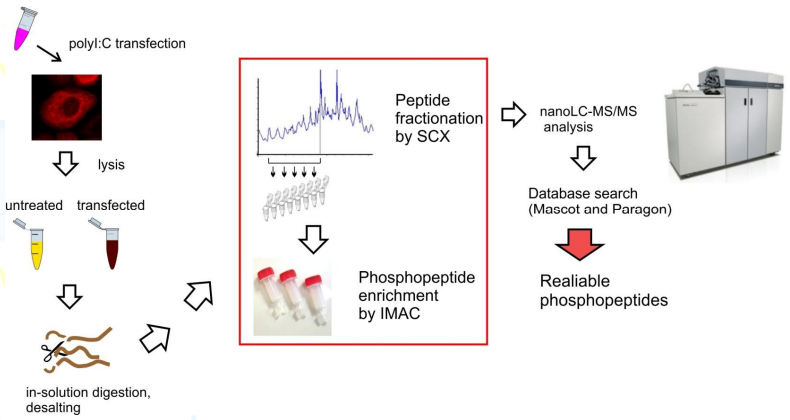
-to characterize the signaling pathways activated in dsRNA-stimulated keratinocytes

-to identify new players in antiviral innate immune responses

For this, we used
14-3-3 affinity chromatography and phosphoproteomics combined with bioinformatics and functional studies

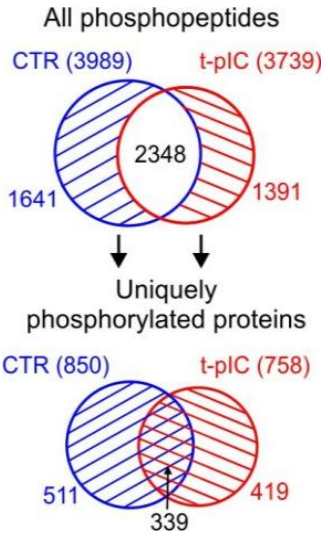


Phosphoproteomics

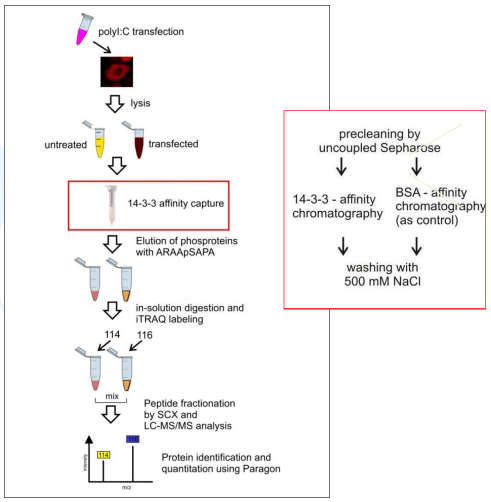


Phosphoproteomics results

All phosphoproteins	1614
All phosphopeptides	5380
Distinct phosphosites	5214
Unreported phosphosites	668



Quantitative 14-3-3 affinity capture



14-3-3 affinity capture results

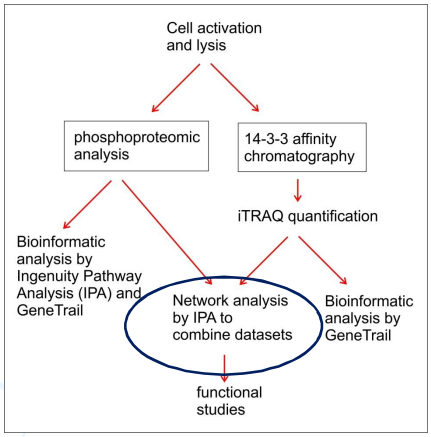
Altogether 646 proteins were identified and quantified with high-confidence

373 (58%) were published previously

209 proteins whose binding to 14-3-3 proteins was changed after dsRNA transfection were identified

- 147 proteins have higher affinity (FC > 1.3)
- 62 proteins have lower affinity (FC < 0.77)

Network analysis to combine datasets

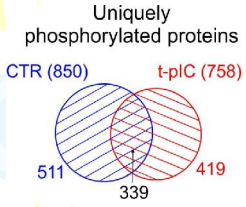


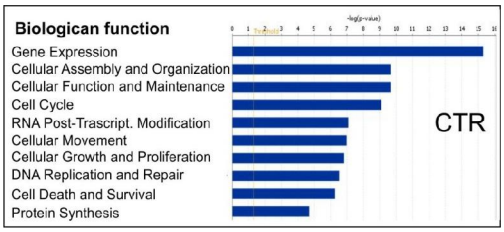
IPA® It's just good science. INGENUITY® SYSTEMS

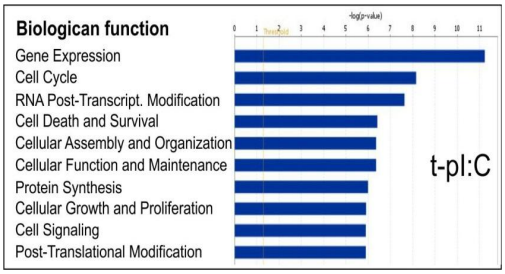
IPA Ingenuity Pathway Analysis

- Bioinformatic software that is based on information from published literature
- Core analysis:
  - Cellular distribution
  - Biological functions
  - Canonical pathways
  - Molecular networks
- Comparison analysis (control vr. stimulated sample)

IPA results for phosphoproteomics data

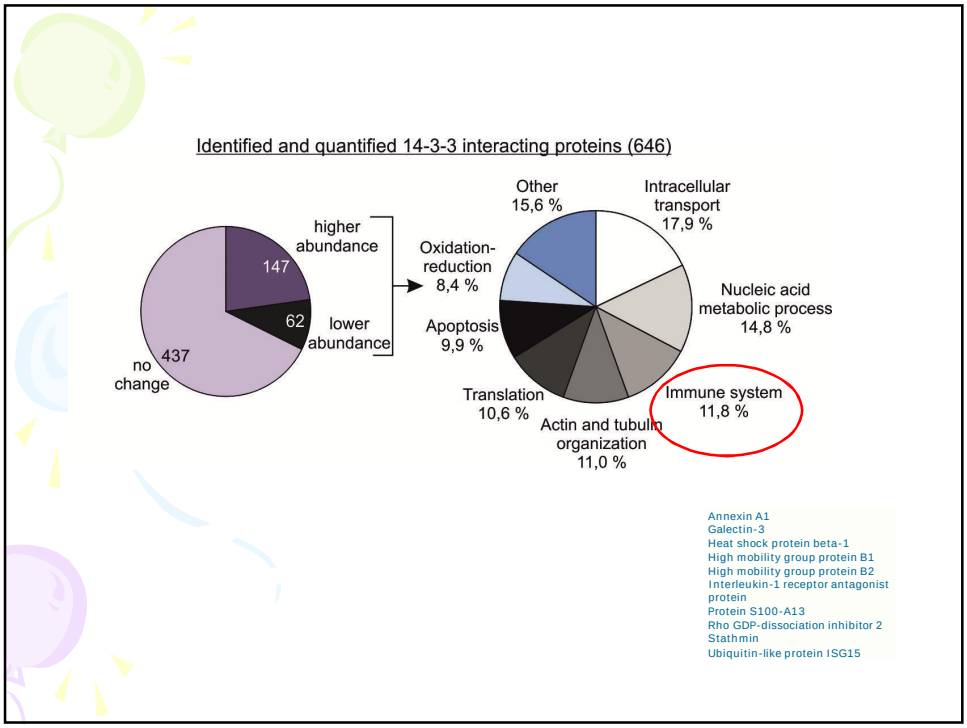






Signaling pathways activated after dsRNA induction based on phosphoproteome data

Ingenuity Canonical Pathways	p-value
ERK/MAPK Signaling	8,16E-10
Insulin Receptor Signaling	2,55E-08
Tight Junction Signaling	1,97E-07
ERK5 Signaling	1,99E-07
Cdc42 Signaling	2,35E-07
Signaling by Rho Family GTPases	1,82E-05
IL-1 Signaling	3,82E-05
ILK Signaling	4,54E-05
Calcium Signaling	1,32E-04
Actin Cytoskeleton Signaling	1,72E-04
p38 MAPK Signaling	2,12E-04
p53 Signaling	3,10E-04
FAK Signaling	1,24E-03
NF-κB Signaling	1,92E-03
14-3-3-mediated Signaling	5,58E-03
Apoptosis Signaling	7,58E-03
Virus Entry via Endocytic Pathways	7,94E-03
SAPK/JNK Signaling	9,51E-03



14-3-3 interacting proteins

Accession #	Name	FC	Published*
P02538	Keratin, type II cytoskeletal 6A	9,24	+
Q04695	Keratin, type I cytoskeletal 17	8,62	+
Q8WUF5	RelA-associated inhibitor *	4,35	
Q9NP97	Dynein light chain roadblock-type 1	2,70	+
P17096	High mobility group protein HMG-I/HMG-Y **	2,39	
Q9UBS4	DnaJ homolog subfamily B member 11	2,27	+
P68371	Tubulin beta-4B chain	2,15	+
Q9BQE3	Tubulin alpha-1C chain	2,15	+
P16949	Stathmin OS=Homo sapiens	2,09	
P07919	Cytochrome b-c1 complex subunit 6, mitochondrial	2,03	

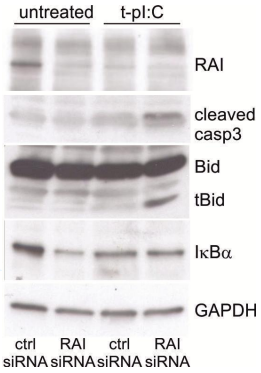
* regulates NFkB signaling and apoptosis

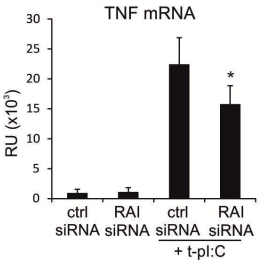
** DNA binding proteins which have role in innate immune response

RelA-associated inhibitor, RAI

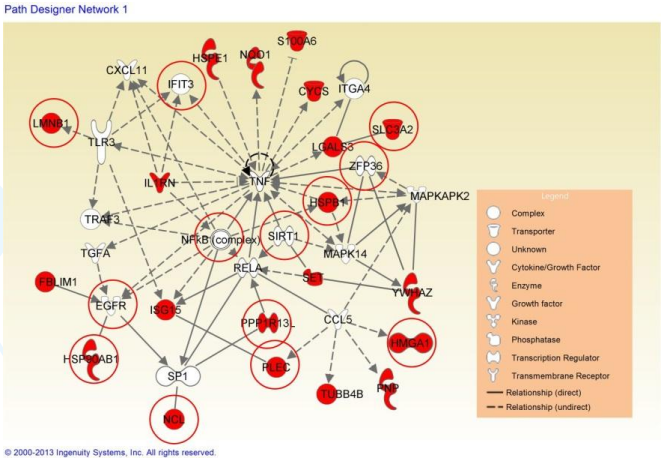
- Inhibitor of apoptosis-stimulating protein of p53 (iASPP)
- Plays a central role in regulation of apoptosis and transcription via its interaction with NF-κB subunit p65 and p53/TP53 proteins
- Our phosphoproteomic data show that RAI has five phosphorylated sites that were uniquely identified after dsRNA stimulation

Functional studies show that RAI negatively regulates dsRNA-induced apoptosis and TNF production





Integrated network analysis



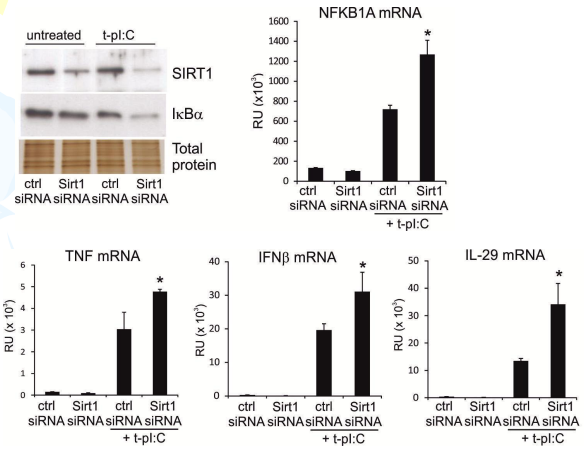
The most significant network of 14-3-3 interacting proteins was related to cell death and survival, cell-to-cell signaling and interaction, and cell morphology.

SIRT1 (= Sirtuin 1)

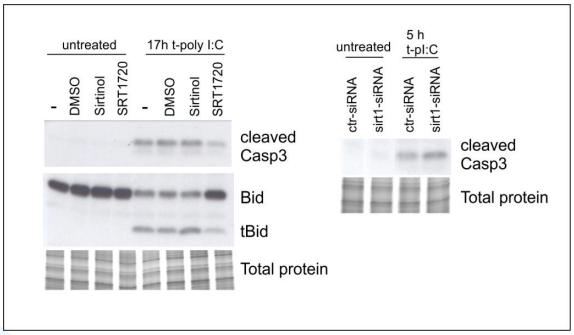
- NAD-dependent protein deacetylase
- participates in the coordination of several separated cellular functions such as cell cycle, response to DNA damage, metobolism, apoptosis, autophagy and aging a variety of age-associated diseases such as type 2 diabetes, obesity, osteoporosis, and Alzheimer's disease are associated with sirtuins

- Our datasets did not identify SIRT1, bioinformatics was needed to find out the relationship of SIRT1 to dsRNA activated signaling related to 14-3-3 proteins

Sirtuin 1 regulates NFκB-signaling and cytokine production upon dsRNA stimulation



SIRT1 negatively regulates dsRNA-induced apoptosis



Similarly, Sirtuin 1 negatively regulates innate immune responses in EMCV (Encephalomyocarditis virus) infected cells!

CONLUSIONS:

RAI and Sirtuin 1 were identified as novel regulators of antiviral innate immune responses

Functional studies showed that RAI inhibits dsRNA-induced apoptosis and contributes to dsRNA-induced TNF cytokine response

Sirtuin 1 is a central molecule regulated by 14-3-3 proteins and functional studies show that it negatively regulates virus-induced cytokine production and protects cells from apoptosis in viral-infected keratinocytes