V6: Protein phosphorylation during cell cycle

Protein **phosphorylation** and **dephosphorylation** are highly controlled biochemical processes that respond to various intracellular and extracellular stimuli.

Phosphorylation status modulates protein activity,
- influencing the tertiary and quaternary **structure** of a protein,
- controlling **subcellular distribution**, and
- regulating **interactions** with other proteins.

Regulatory protein phosphorylation is a transient modification
that is often of low occupancy or “stoichiometry”

This means that only a fraction of a particular protein may be phosphorylated on a given site at any particular time, and that occurs on regulatory proteins of low abundance, such as protein kinases and transcription factors.

Olsen Science
Signaling 3 (2010)

CDK inhibitors

V1, V5: Progression through the cell-division cycle is regulated by the coordinated activities of **cyclin/cyclin-dependent kinases** (CDK) complexes.

One level of regulation of these cyclin-CDK complexes is provided by their binding to **CDK inhibitors (CKIs).**

There are two important families of CKIs:

(1) Members of **INK4 gene family** ($p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$, and $p19^{INK4d}$) bind to CDK4 and CDK6 and inhibit their kinase activities by interfering with their association with D-type cyclins.

(2) CKIs of the **Cip/Kip family** bind to both cyclin and CDK subunits and can modulate the activities of cyclin D-, E-, A-, and B-CDK complexes.

Besson et al. Develop Cell 14, 159 (2008)

Cip/Kip Proteins

Cip/Kip family members:

p21 $^{Cip1/Waf1/Sdi1}$ (**p21**, encoded by *cdkn1a*)
p27 ^{Kip1} (**p27**, encoded by *cdkn1b*)
p57 ^{Kip2} (**p57**, encoded by *cdkn1c*)

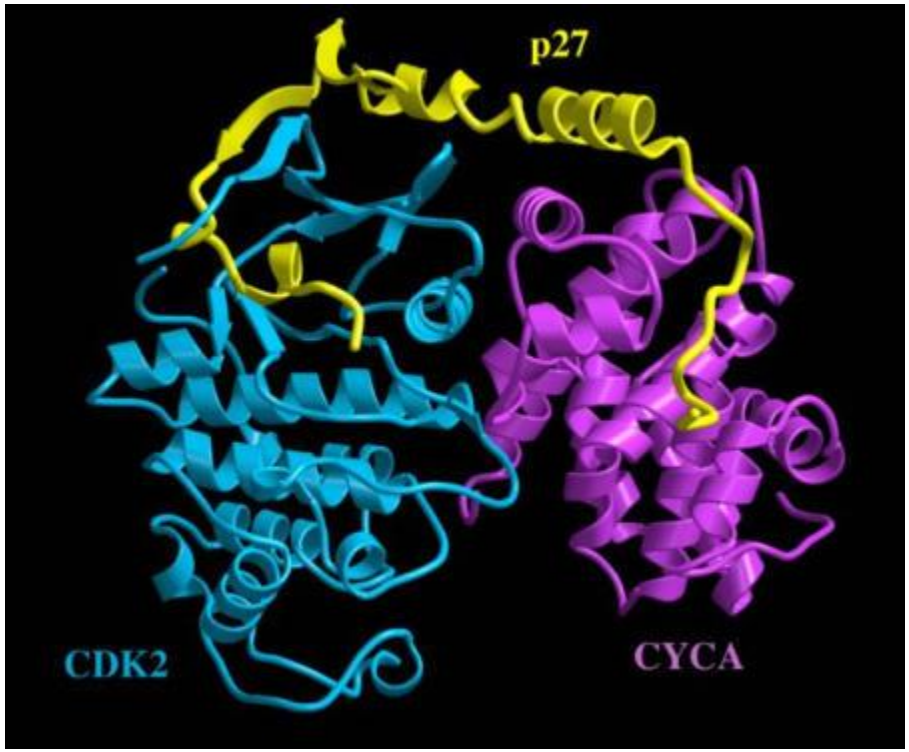
Cip/Kip family members have a general importance in **restraining proliferation** during development, differentiation, and response to cellular stresses.

However, each Cip/Kip has specific biological functions.

Different anti-proliferative signals tend to cause elevated expression of only a subset of the Cip/Kip proteins.

Besson et al. Develop Cell 14, 159 (2008)

Review (V5): Crystal structure



p27(Kip1)-CyclinA-Cdk2 Complex

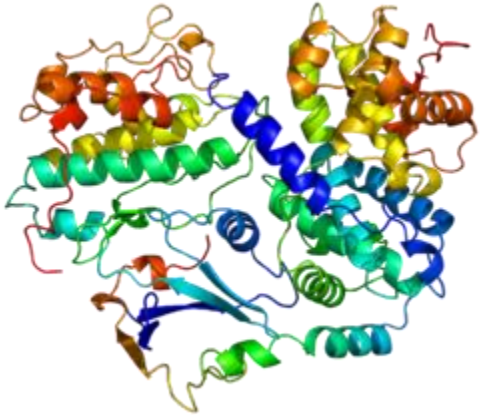
Nikola Pavletich
(crystallographer)

p27 (Kip1) is shown bound to the CyclinA-Cdk2 complex, provoking profound changes in the kinase active site and rendering it inactive (by blocking the ATP-binding site).

p27 also interacts with the secondary substrate recognition site on the cyclin.

Cip/Kip Proteins

p21 is an important transcriptional target of **p53** and mediates DNA-damage-induced **cell-cycle arrest** in G1 and G2.



p27 binds to and prevents the activation of cyclinE-CDK2 or cyclinD-CDK4 complexes, and thus controls the cell cycle progression at G1.

In contrast to p21 and p27, **p57** has a tissue-restricted expression pattern during **embryogenesis** and in the adult. p57 is the only CKI that is required for embryonic development.

The gene coding for p57 (*cdkn1c*) is **genetically imprinted** with preferred expression of the maternal allele.

Besson et al. Develop Cell 14, 159 (2008)
www.wikipedia.org

Cip/Kip Proteins

Initially, p21, p27, and p57 were considered as **tumor suppressors** based on their **ability to block cell proliferation**.

However, p21, p27, and p57 are also involved in the regulation of cellular processes beyond cell-cycle regulation, including transcription, apoptosis and migration.

These processes may be **oncogenic** under certain circumstances.

In tumors, inactivating mutations of the *cdkn1b* gene (p27) are extremely rare.

p27 is downregulated by other mechanisms, including proteolytic degradation, decreased transcription, cytoplasmic mislocalization, and by miRNAs.

Besson et al. Develop Cell 14, 159 (2008)

Cip/Kip Proteins

The Cip/Kip proteins are **intrinsically unstructured**.

They adopt specific tertiary conformations only after binding to other proteins.

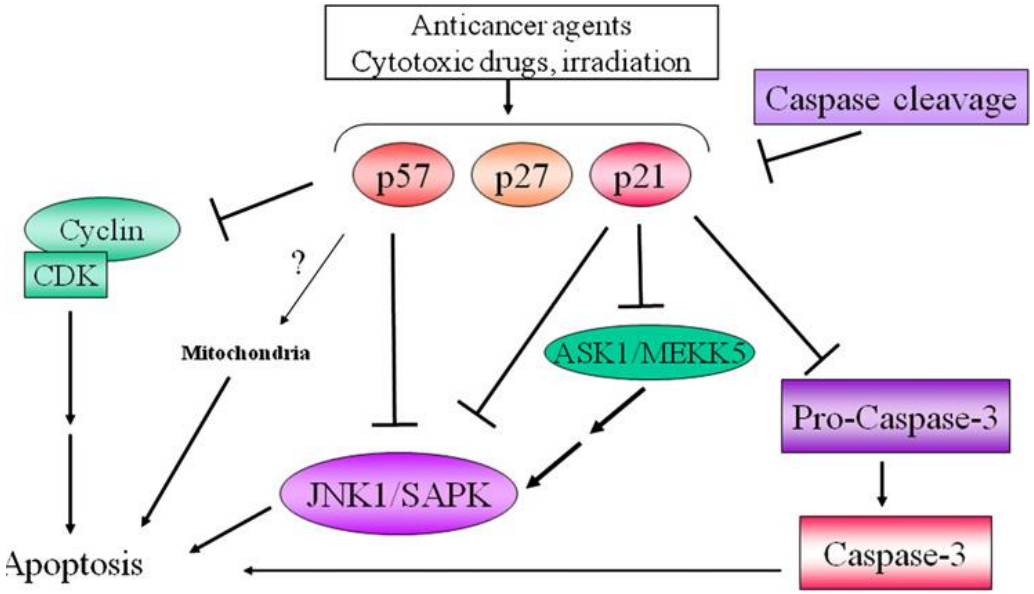
This may explain why CKIs are capable of interacting with a wide diversity of proteins to regulate various cellular functions.

The binding specificity of Cip/Kip proteins appears to be modulated by **phosphorylation** at specific residues, and by binding to other proteins.

Phosphorylation of Cip/Kip proteins also affects their **stability** and their **subcellular localization**.

Besson et al. Develop Cell 14, 159 (2008)

Cip/Kip Proteins and Apoptosis (cell death)



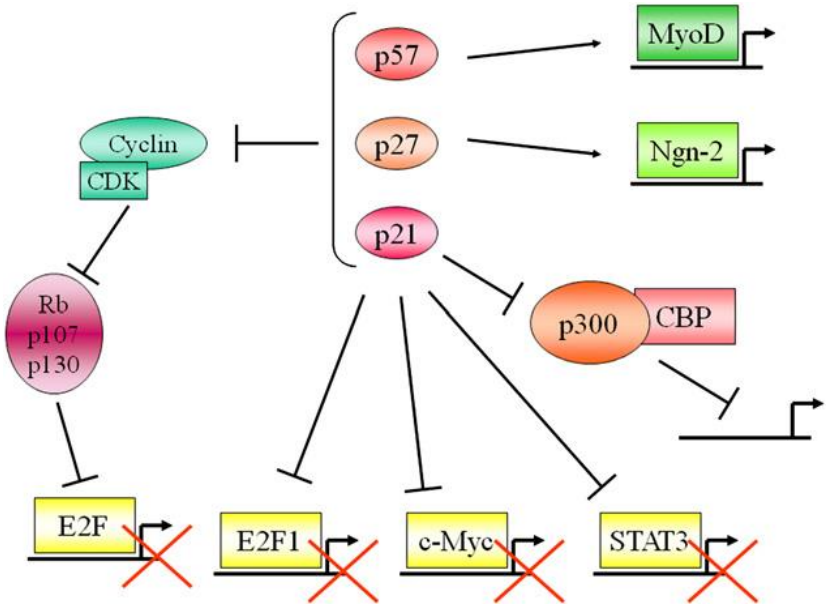
Cip/Kip proteins can **inhibit apoptosis** via the inhibition of cyclin-CDK complexes.

But p21 and p27 may also be cleaved by **caspases**. Their degradation promotes cyclin-CDK activation during the apoptotic process.

p21and p57 may also directly prevent the induction of apoptosis by interfering with activation of the **stress-signaling pathways**; for instance, both bind to and inhibit the activity of **JNK1/SAPK**, and p21 can also inhibit ASK1/MEKK5.

Besson et al. Develop Cell 14, 159 (2008)

Transcriptional Regulation by Cip/Kip Proteins



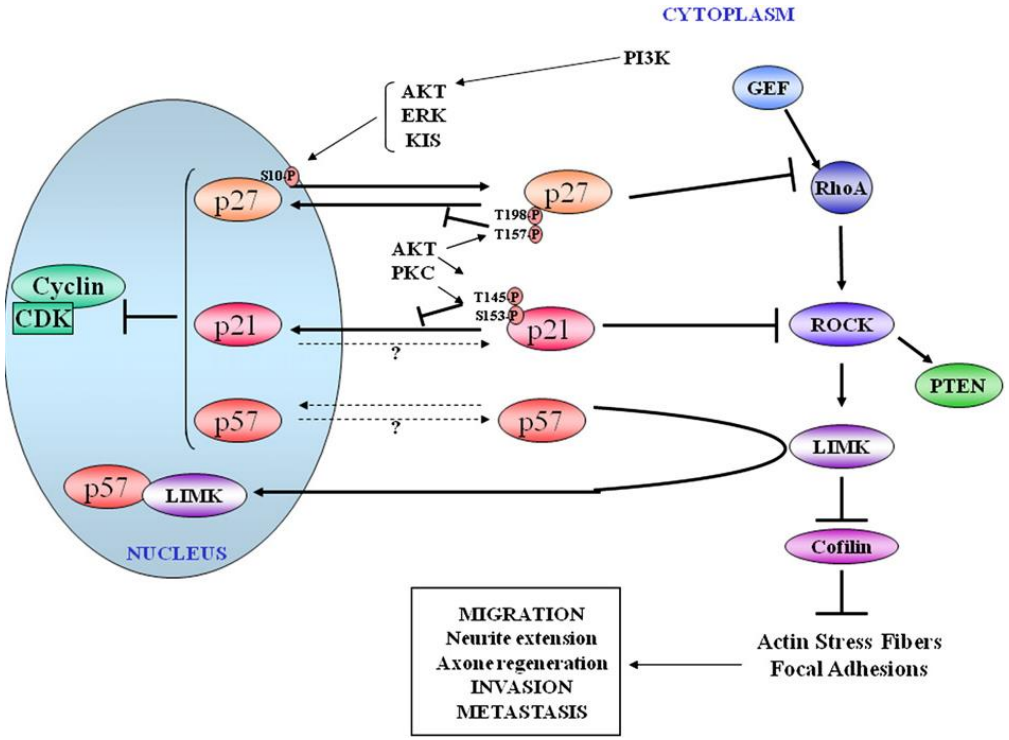
The CKIs p21, p27, and p57 can indirectly repress transcription mediated by the TF **E2F** via the inhibition of cyclin-CDK complexes.

Blocking cyclin-CDK keeps low phospho-rylation levels of their targets in the Rb-family proteins (Rb/p110, p107, and p130) in which they block E2F.

Cip/Kip proteins also regulate TFs directly. For instance, **p57** and **p27** can interact with MyoD and Neurogenin-2 (Ngn-2), respectively, stabilizing them and promoting transcription of their target genes.

p21 also binds to E2F1, c-Myc, and STAT3 to inhibit their activities and derepresses p300/CBP targets by inhibiting the transcriptional repression domain of p300.

Besson et al. Develop Cell 14, 159 (2008)



In the nucleus, Cip/Kip proteins primarily function to restrict the activities of cyclin-CDK complexes.

Phosphorylation of **p27** on Ser-10 promotes its binding to the exportin CRM1 and **nuclear export**.

On the other hand, phosphorylations on T157 (by Akt) or T198 (by Akt or p90Rsk) promote binding to 14-3-3 proteins and prevent the reentry of p27 in the nucleus.

In the cytosol, p27 can bind to RhoA, preventing its activation by its GEFs (guanine-nucleotide exchange factors), leading to decreased actin stress fiber and focal-adhesion formation and resulting in several cell types in increased migration, invasion, and metastasis. PI3K-AKT induction of cytoplasmic localization of p27 is also involved in the inhibition of PTEN activation via p27-mediated inhibition of the RhoA-ROCK pathway.

p21 cytoplasmic localization is induced by phosphorylation on T145 and S153 by Akt and PKC, respectively. Cytoplasmic p21 can bind to ROCK, inhibiting its kinase activity, resulting in decreased actin stress fibers formation.
Cytoplasmic p57 can bind to LIMK and induce its translocation into the nucleus, resulting in loss of actin stress fibers.

Besson et al. Develop Cell 14, 159 (2008)

Cell Cycle and the Phosphoproteome

CELL CYCLE

Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis

Jesper V. Olsen,[1,2*] Michiel Vermeulen,[1,3*] Anna Santamaria,[4*] Chanchal Kumar,[1,5*] Martin L. Miller,[2,6] Lars J. Jensen,[2] Florian Gnad,[1] Jürgen Cox,[1] Thomas S. Jensen,[7] Erich A. Nigg,[4] Søren Brunak,[2,7] Matthias Mann[1,2†]

(Published 12 January 2010; Volume 3 Issue 104 ra3)

www.SCIENCESIGNALING.org 12 January 2010 Vol 3 Issue 104 ra3

Aim: Analyze all proteins that are modified by phosphorylation during different stages of the cell cycle of human HeLa cells.

Ion-exchange chromatography + HPLC + MS + sequencing led to the identification of 6695 proteins.
From this 6027 quantitative cell cycle profiles were obtained.

A total of 24,714 phosphorylation events were identified.
20,443 of them were assigned to a specific residue with high confidence.

Finding: about **70%** of all proteins get phosphorylated.

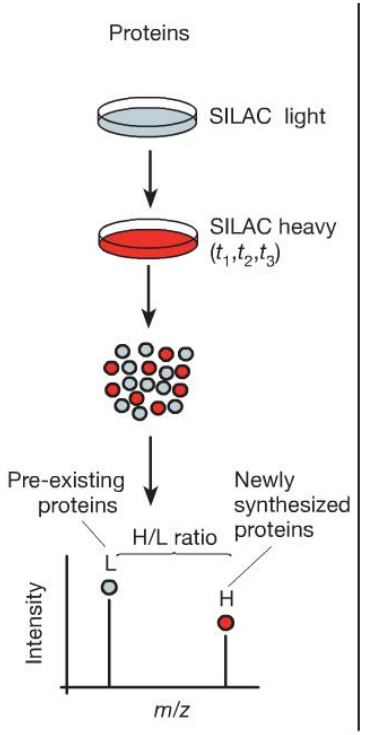
Review: protein quantification by SILAC

ARTICLE

doi:10.1038/nature10098

Global quantification of mammalian gene expression control

Björn Schwanhäusser[1], Dorothea Busse[1], Na Li[1], Gunnar Dittmar[1], Johannes Schuchhardt[2], Jana Wolf[1], Wei Chen[1] & Matthias Selbach[1]



SILAC: „**s**table **i**sotope **l**abelling by **a**mino acids in **c**ell culture“ means that cells are cultivated in a medium containing heavy stable-isotope versions of essential amino acids.

When non-labelled (i.e. light) cells are transferred to heavy SILAC growth medium, newly synthesized proteins incorporate the heavy label while pre-existing proteins remain in the light form.

Quantification protein turnover and levels.
Mouse fibroblasts are transferred to medium with heavy amino acids (SILAC)

.

Protein turnover is quantified by mass spectrometry and next-generation sequencing, respectively.

Schwanhäuser et al. Nature 473, 337 (2011)

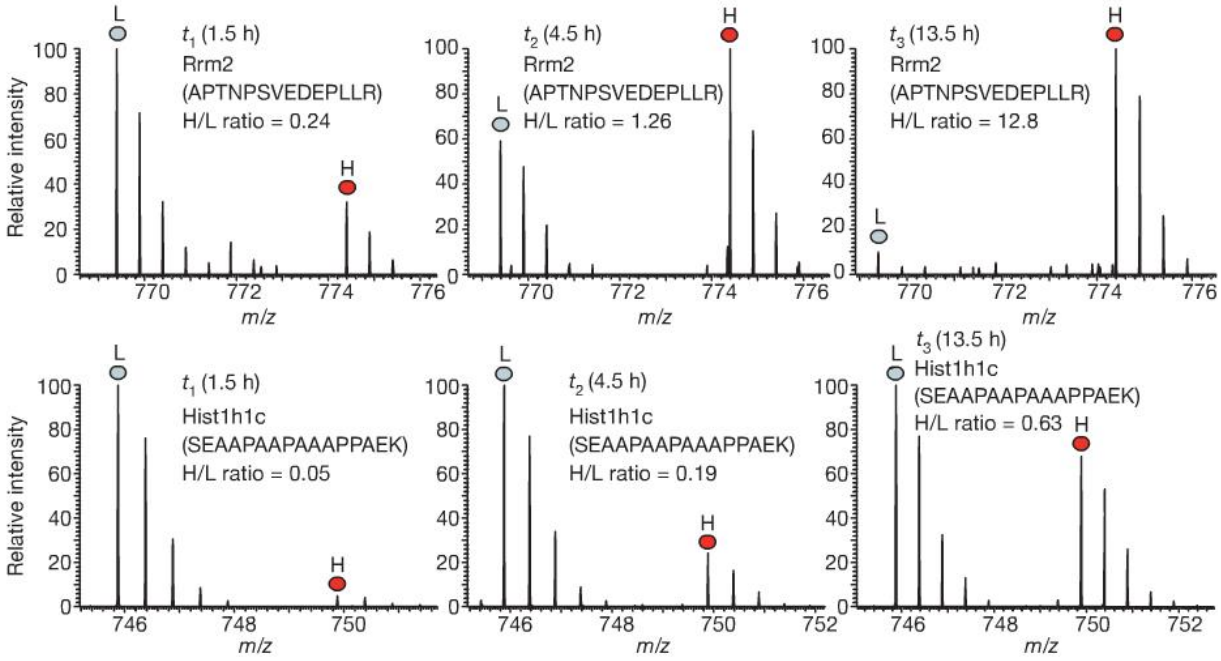
Rates of protein translation

Mass spectra of peptides for two proteins.

Top: **high-turnover protein**
Bottom: **low-turnover protein**.

Over time, the heavy to light (H/L) ratios increase.

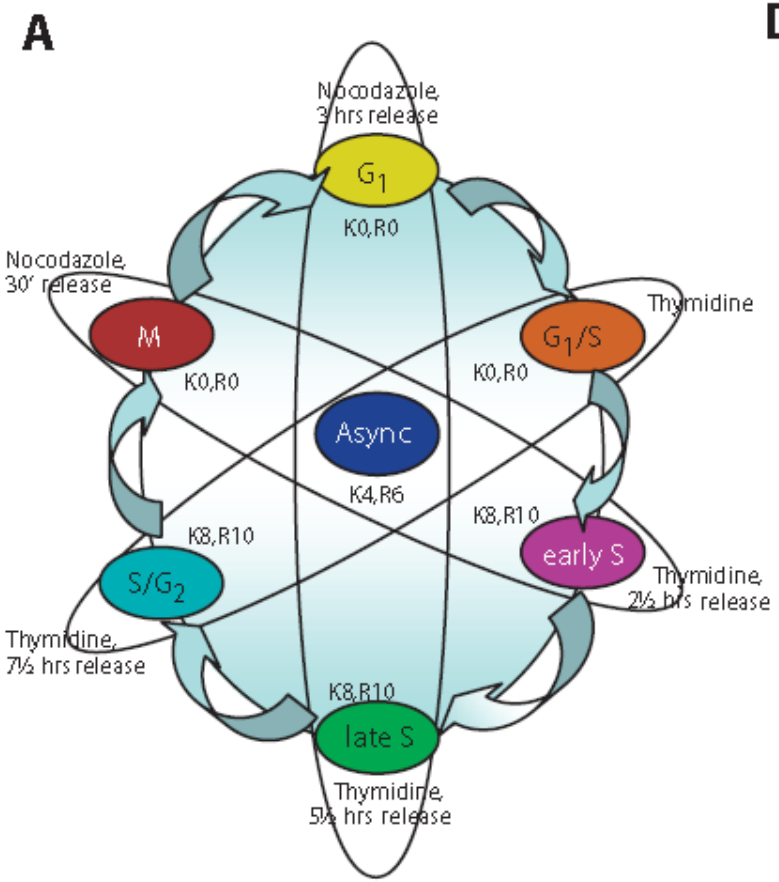
H-concentration of high-turnover protein saturates.
That of low-turnover protein still increases.



This example was introduced to illustrate the principles of SILAC and mass spectroscopy signals (peaks).
In the Olson et al. study, the authors used H and L forms to label different stages of the cell cycle.

Schwanhäuser et al. Nature 473, 337 (2011)

Quantitative proteomic analysis



HeLa S3 cells were SILAC-labeled with 3 different isotopic forms (light – medium –heavy) of arginine and lysine.

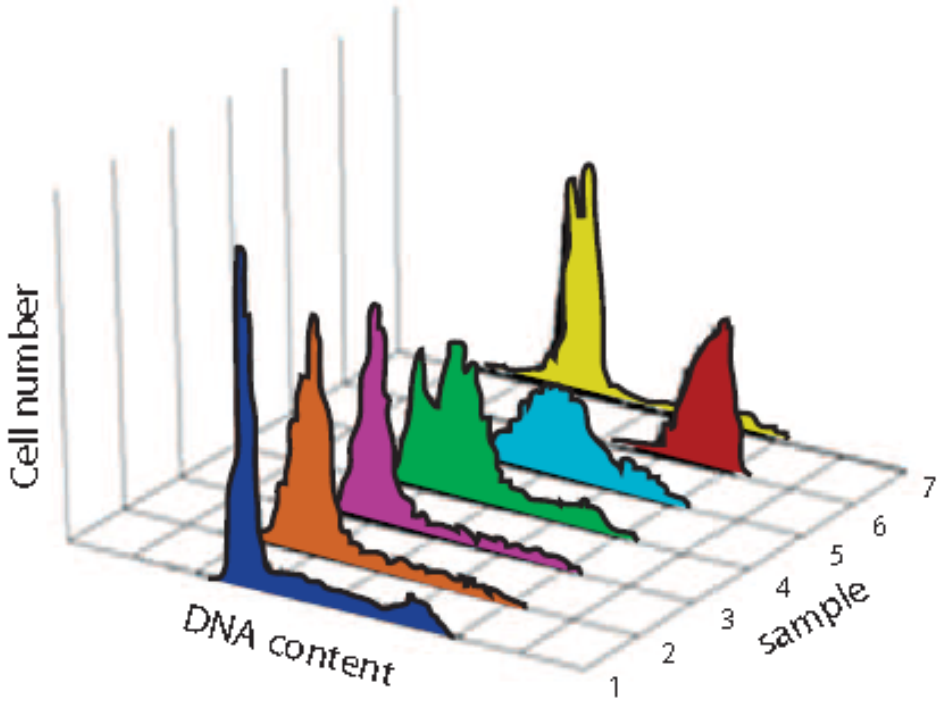
3 individual populations of heavy and light SILAC cells were synchronized with a **thymidine** block (analog of thymine, blocks entry into S phase). Cells were then collected at six different time points across the cell cycle after release from the thymidine arrest.

2 samples were collected after a **cell cycle arrest** with **nocodazole** and release. (Nocodazole interferes with polymerization of microtubules.)

Cells were lysed and mixed in equal amounts using an asynchronously growing cell population as the internal standard to allow normalization between experiments.
3 independent experiments were performed to cover six cell cycle stages.

FACS profiles of individual HeLa populations

	% Cells		
	G_1	S	G_2/M
1. Asynchronous	64	27	9
2. Thymidine block	50	46	4
3. Thymidine block + release 2½ h	36	60	4
4. Thymidine block + release 5½ h	23	70	7
5. Thymidine block + release 7½ h	15	70	15
6. Nocodazole block + release ½ h	1	11	88
7. Nocodazole block + release 3 h	82	12	6



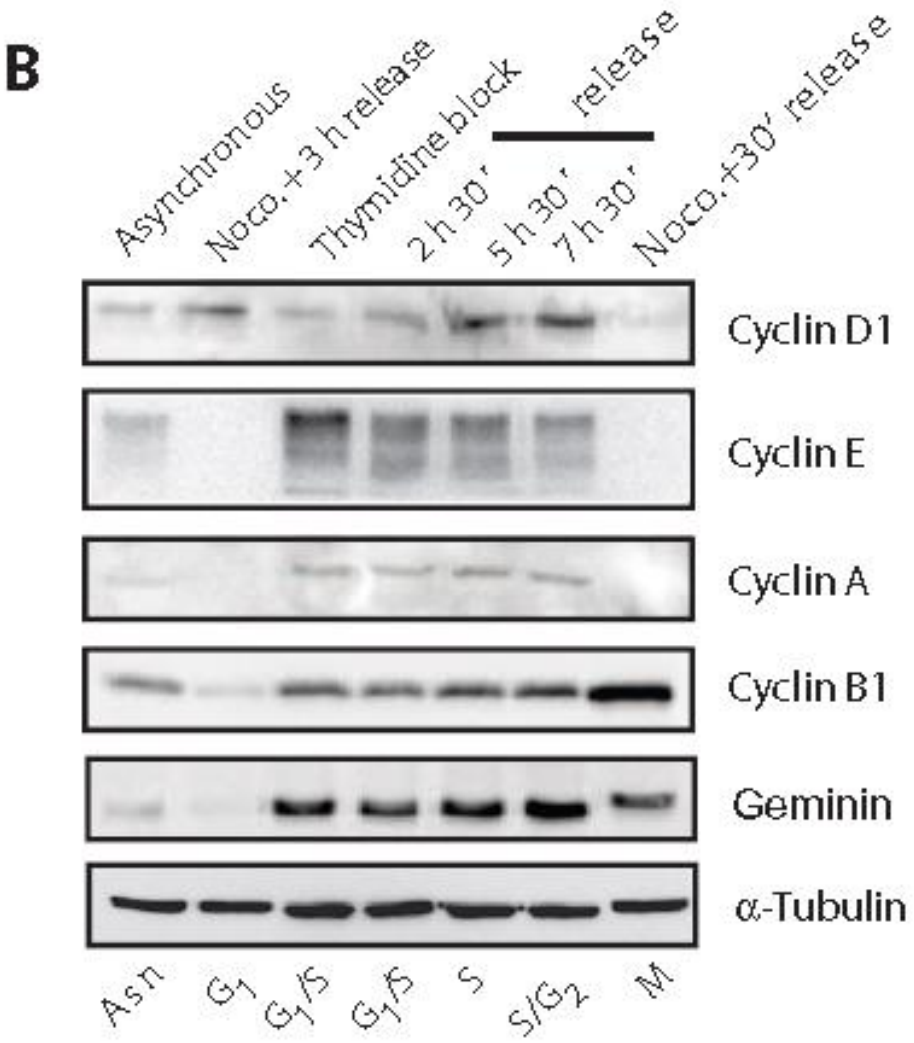
Cells were fixed and collected by centrifugation.

Then the DNA content of the cells was determined with propidium iodide.

This is the basis for classifying the state along the cell cycle.

Olsen Science Signaling 3 (2010)

Quantification of cell cycle markers

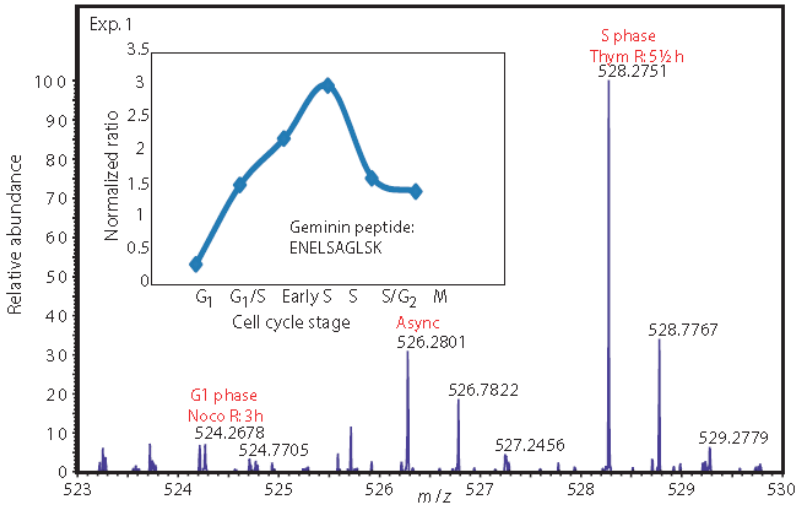


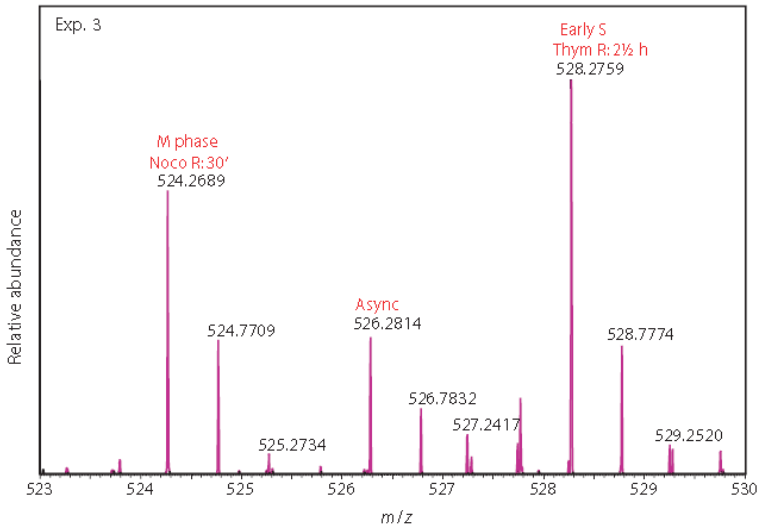
Immunoblot analysis of known cell cycle marker proteins in the different cell populations.

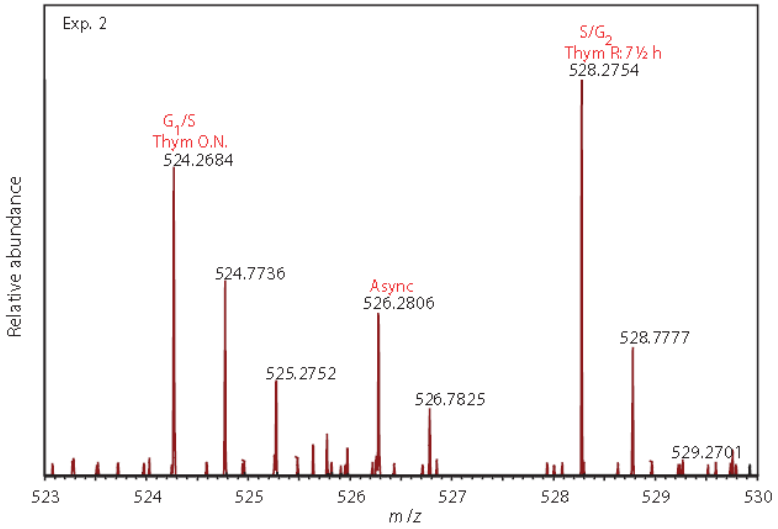
The abundance of a fifth of the proteome changed by at least fourfold throughout the cell cycle (difference between lowest and highest abundance).

Because a **fourfold change** also best accounted for the dynamics of already described cell cycle components, this ratio was used as a threshold for subsequent analysis.

Monitoring of protein abundance by MS







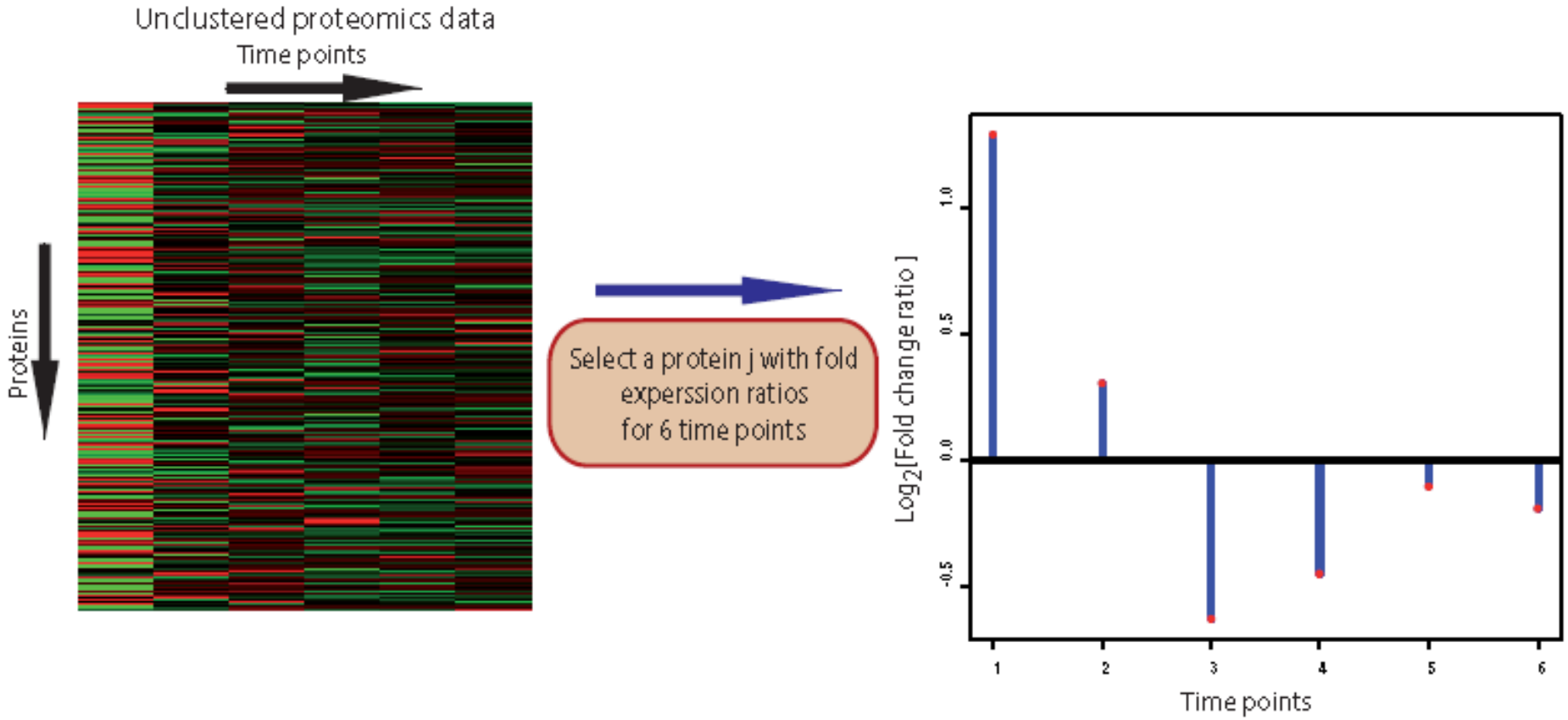
Representative MS data showing how the abundance of the proteins was monitored in three experiments (Exp. 1, Exp. 2, Exp. 3) to obtain information from the 6 stages of the cell cycle.

The data show the MS analysis of a tryptic SILAC peptide triplet derived from the cell cycle marker protein Geminin.

Relative peptide abundance changes were normalized to the medium SILAC peptide derived from the asynchronously grown cells in all three experiments. The inset shows the combined six-time profile of Geminin over the cell cycle.

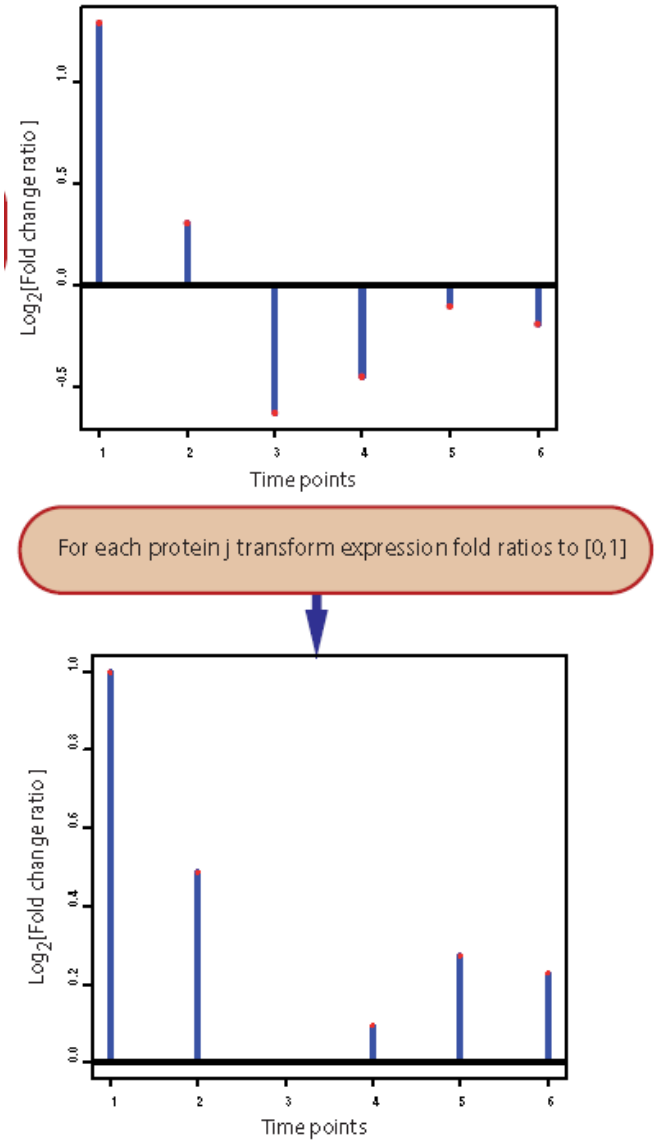
Olsen Science Signaling 3 (2010)

Bioinformatics Workflow (1)



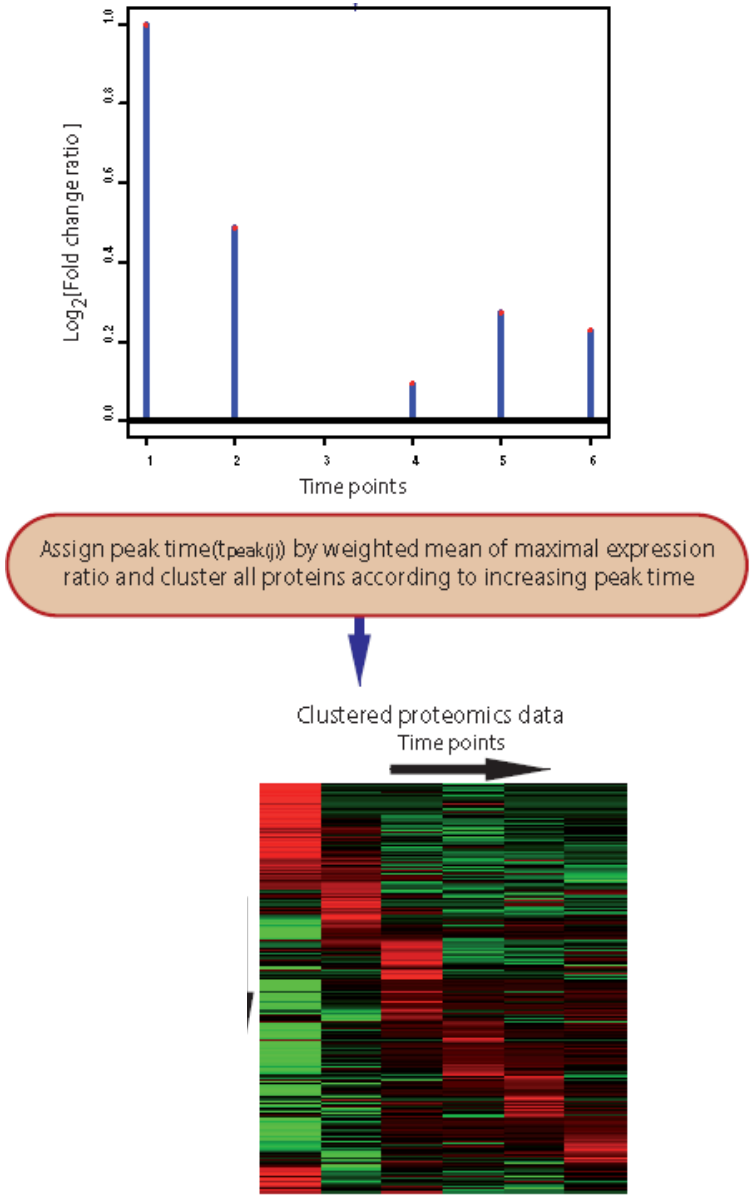
Olsen Science Signaling 3 (2010)

Bioinformatics Workflow (2)



Olsen Science Signaling 3 (2010)

Bioinformatics Workflow (3)

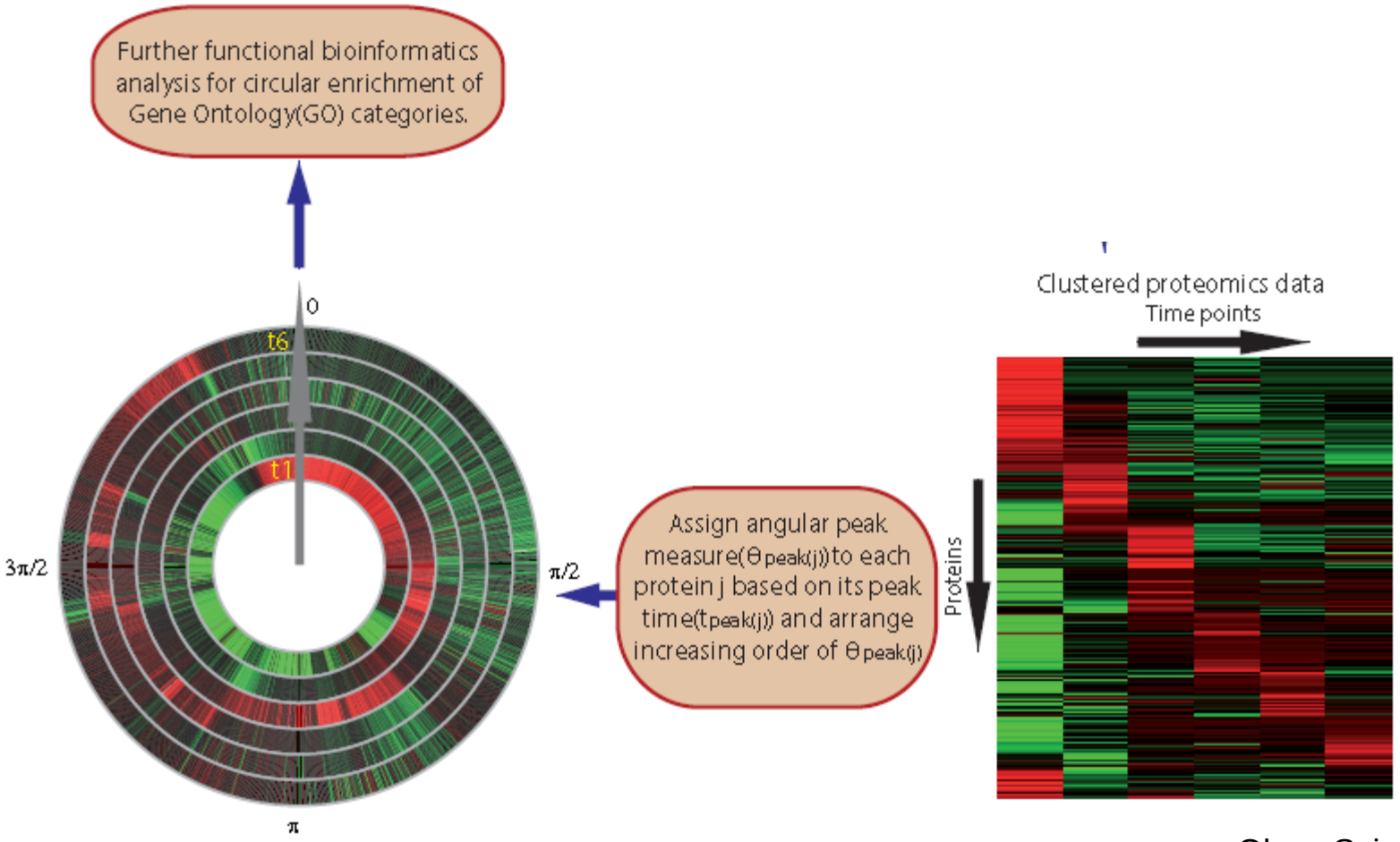


For each protein a peak time index was calculated by weighted mean of its maximal expression at time point t_i w.r.t its adjacent time points t_{i-1} and t_{i+1}.

The proteins were then clustered according to their increasing peak time indices.

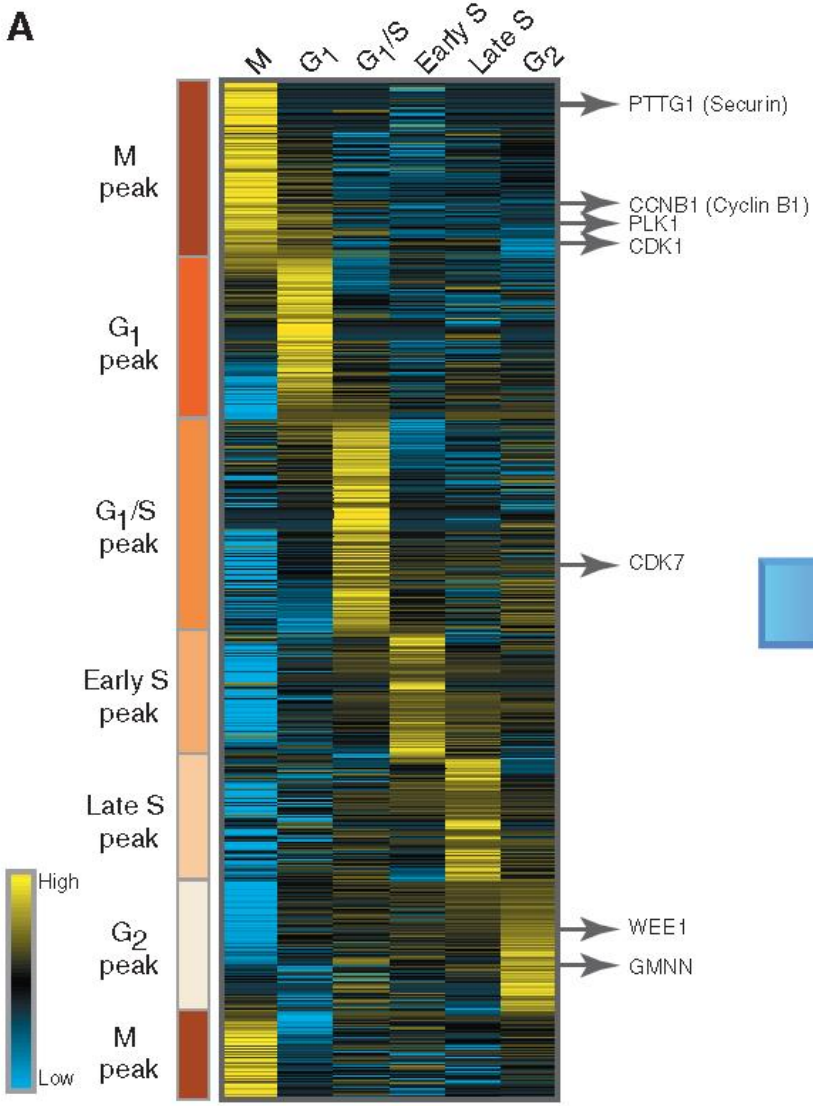
Olsen Science Signaling 3 (2010)

Bioinformatics Workflow (4)



Olsen Science Signaling 3 (2010)

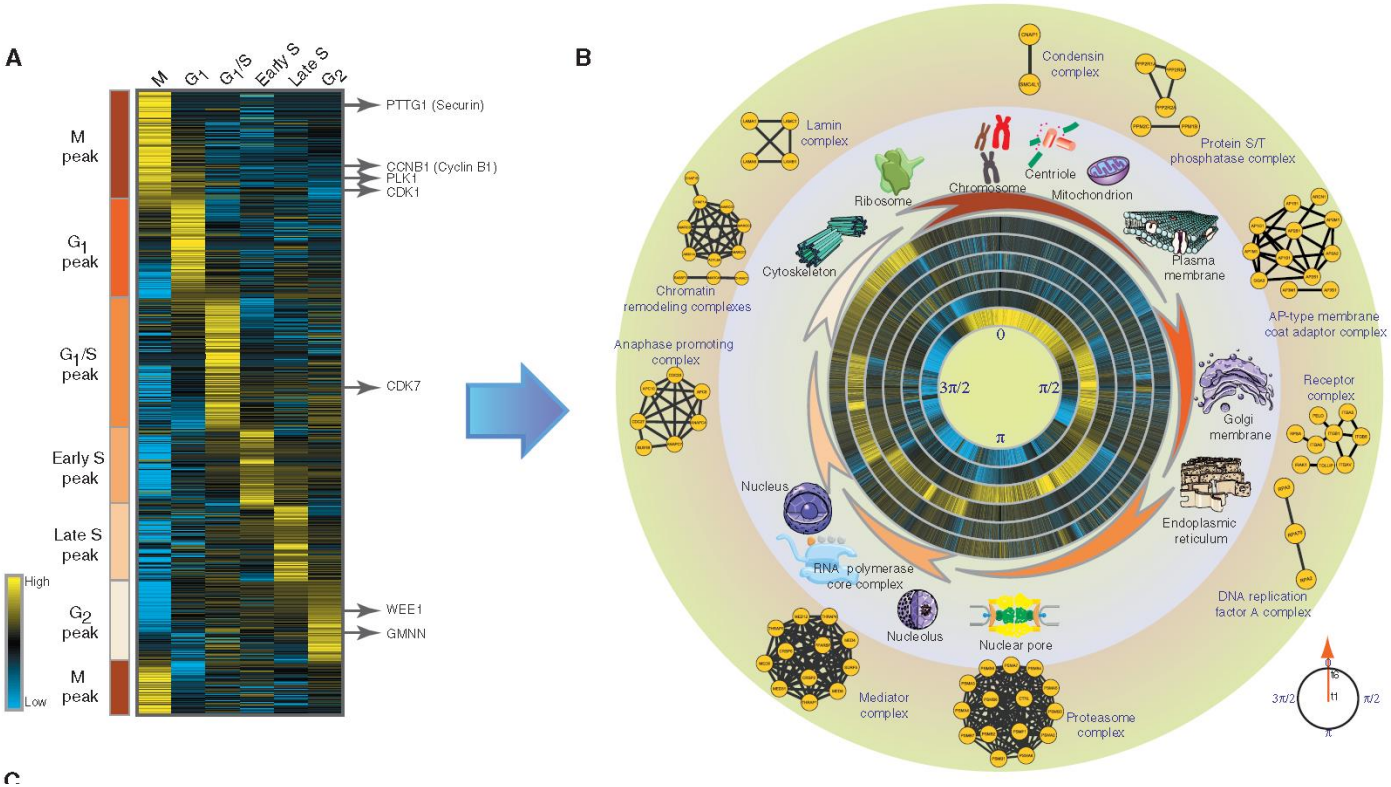
Dynamics of the proteome during the cell cycle



Proteins whose abundance changed at least fourfold during the cell cycle were clustered in all cell cycle stages by calculating a time peak index by weighted mean of the ratio of maximal abundance.

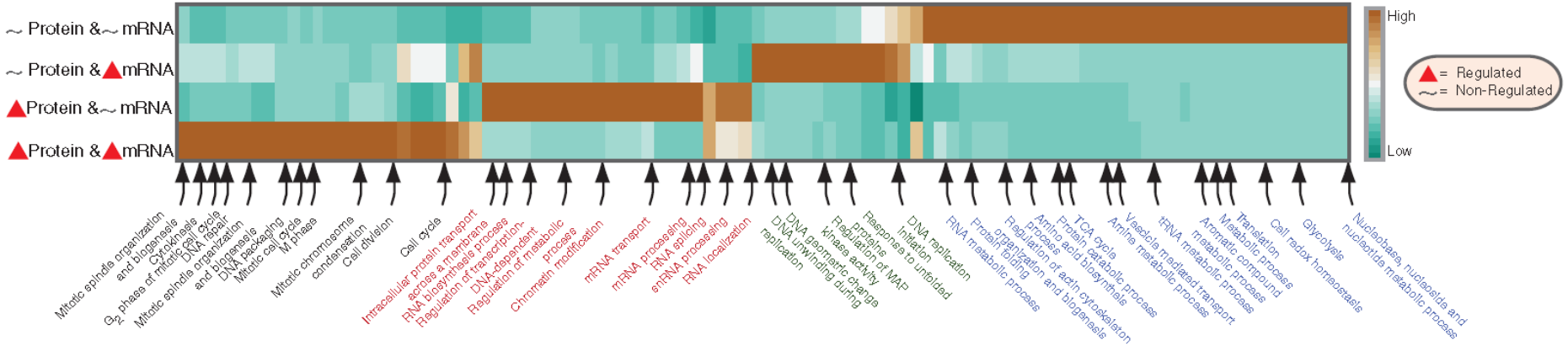
For each cell cycle stage, there are clear patterns of up- and down-regulation.

Determine protein peaks



(B) A circularized representation of the data shown in (A) was used to determine the angle in the cell cycle where the abundance of particular proteins peaks. Coordinately regulated protein complexes and organellar proteins at each cell cycle stage are indicated around the circle.

Comparison of mRNA and protein dynamics



Comparison of mRNA and protein dynamics during the cell cycle. Measured protein dynamics were correlated to published mRNA data.

Proteins were grouped on the y axis in four categories from top to bottom:

- unchanging mRNA and protein
- changing mRNA and unchanging protein
- unchanging mRNA and changing protein
- and changing mRNA and changing protein.

The x axis shows clustered gene ontology (GO) biological process terms enriched in at least one of the above four categories. High and low represent statistical over- or underrepresentation, respectively.

Absolute phosphorylation site stochiometry

Now we want to derive the phosphorylation state of protein residues during the cell cycle.
We need to substract out the changes of protein abundance.
-> we want to know (1) and (2) below

(1) Proportion of phosphorylated to unphosphorylated peptide in Light SILAC state: $\frac{N_L^{PHOS}}{N_L^{NonP}} = a$

N_L^{PHOS} is the total copy number of a given phosphopeptide in the light SILAC state, and N_L^{NonP} is the total copy number the corresponding unphosphorylated peptide in the light SILAC state

(2) Proportion of phosphorylated to unphosphorylated peptide in Heavy SILAC state: $\frac{N_H^{PHOS}}{N_H^{NonP}} = b$

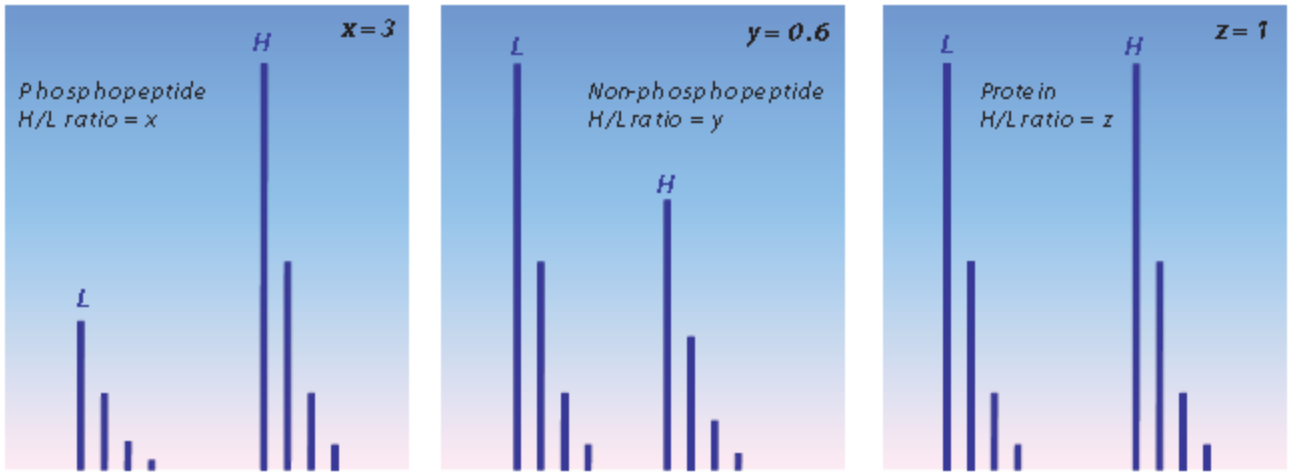
N_H^{PHOS} is the total copy number of a given phosphopeptide in the heavy SILAC state, and
N_H^{NonP} is the total copy number the corresponding unphosphorylated peptide in the heavy SILAC state

(3) We expect that $\frac{N_H^{PHOS} + N_H^{NonP}}{N_H^{PROTEIN}} = \frac{N_L^{PHOS} + N_L^{NonP}}{N_L^{PROTEIN}}$

$N_L^{PROTEIN}$ is the total copy number of the phosphoprotein in the light SILAC state, and
$N_H^{PROTEIN}$ is the total copy number the phosphoprotein in the heavy SILAC state

Olsen Science Signaling 3 (2010)

Available experimental data



To determine phosphorylation sites that show dynamic profiles due to changes in phosphorylation state rather than due to changes in protein abundance, we consider the measured phosphopeptide H/L ratios.

From the experiment we have:

- the SILAC ratio *x* for phosphopeptide
- the SILAC ratio *y* for non-phosphopeptide (the unphosphorylated version of the phosphopeptide),
- and protein ratio *z* (the total amount of the protein in both phosphorylated and nonphosphorylated forms).

Absolute phosphorylation site stochiometry

From the MS data we know:

(4) Relative phosphopeptide ratio = $\frac{N_H^{PHOS}}{N_L^{PHOS}} = x$

(5) Relative unphosphorylated peptide ratio = $\frac{N_H^{NonP}}{N_L^{NonP}} = y$

(6) Relative total phosphoprotein ratio = $\frac{N_H^{PROTEIN}}{N_L^{PROTEIN}} = z$

If we know x, y and z then we can solve equations 1 and 2 by substituting in equations 3:

(1) Occupancy rate in Light SILAC state: $\frac{N_L^{PHOS}}{N_L^{NonP}} = a = \frac{z-y}{x-z}$ z

(2) Occupancy rate in Heavy SILAC state: $\frac{N_H^{PHOS}}{N_H^{NonP}} = b = \frac{x \cdot (z-y)}{y \cdot (x-z)}$

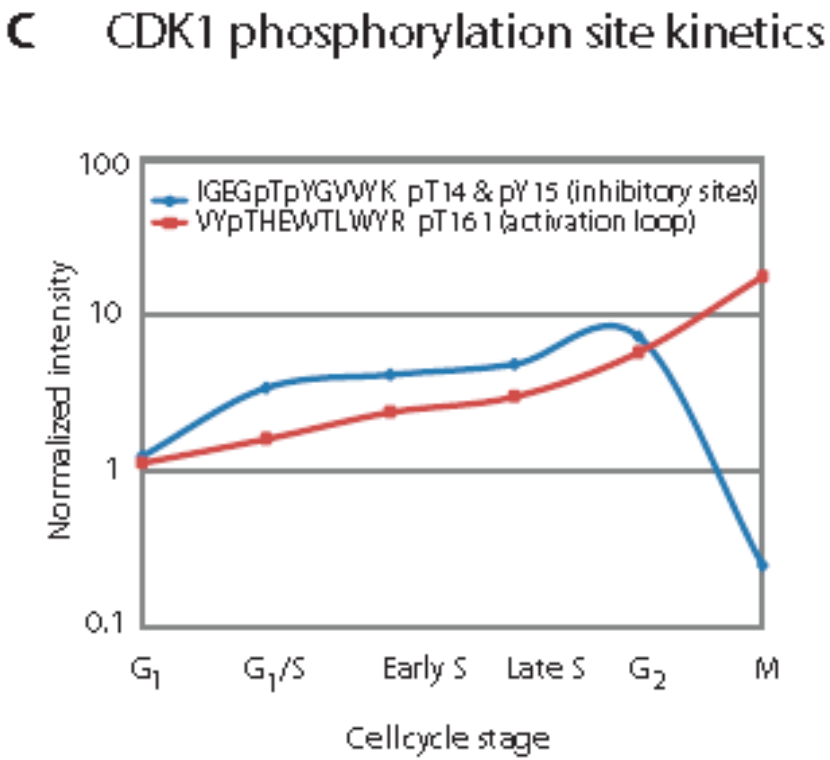
We expect that $N_L^{PHOS} + N_L^{NonP} = N_H^{PHOS} + N_H^{NonP} = 100\% = 1$

and can therefore calculate the phosphorylation site occupancy in the Light and Heavy SILAC state as:

(3) Light SILAC occupancy: a/(1+a) and Heavy SILAC occupancy: b/(1+b)

Olsen Science Signaling 3 (2010)

Example: Dynamic phosphorylation of CDK1

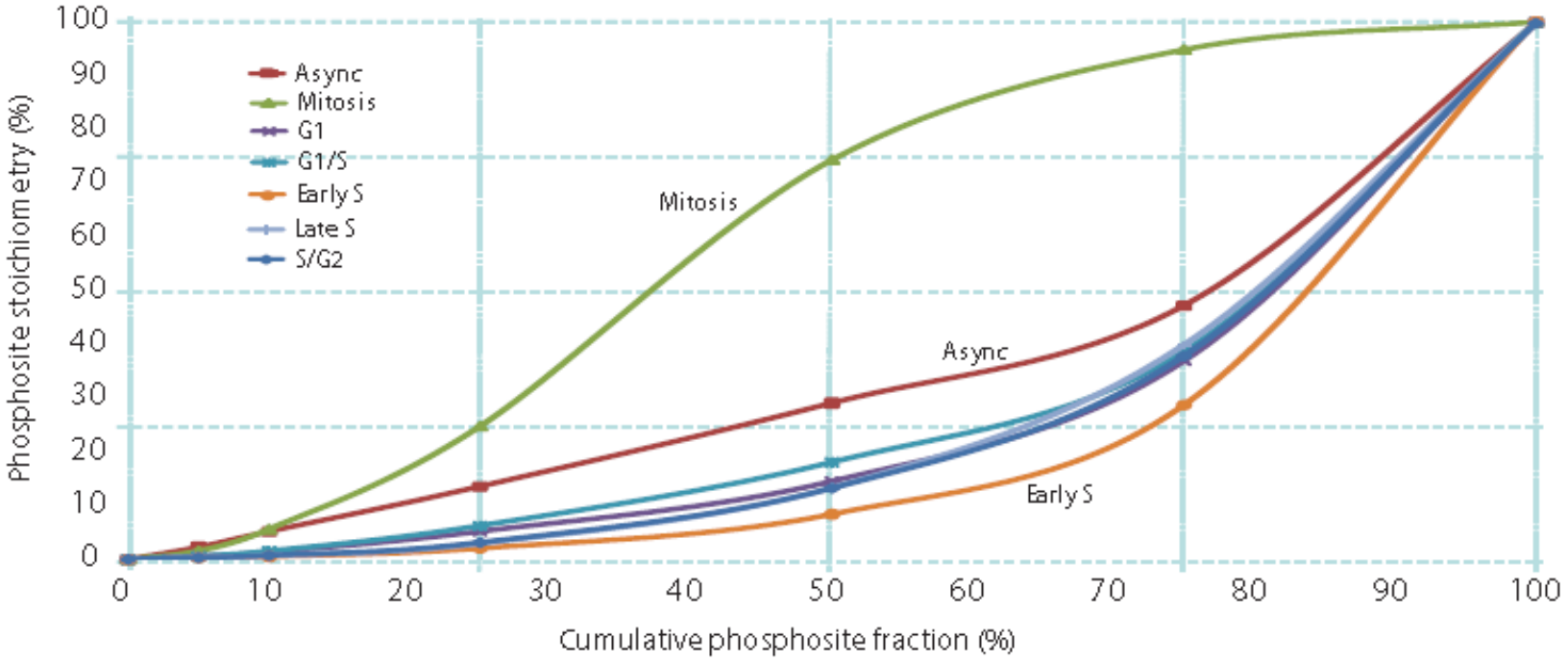


Dynamic profile of two CDK1 phosphopeptides during the cell cycle.

The activating site T161 peaks in mitosis, whereas phosphorylation of the inhibitory sites T14 and Y15 is decreased in mitosis

Olsen Science Signaling 3 (2010)

Total phosphosite occupancy in different stages of cell cycle

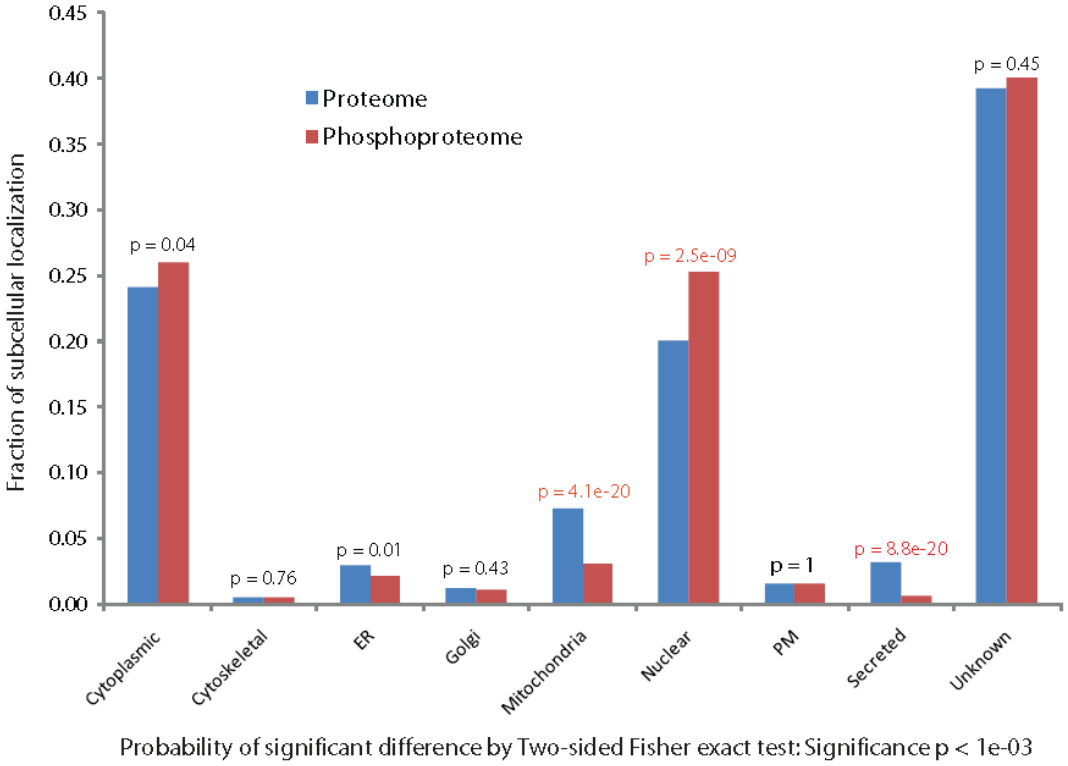


Fifty percent of all mitotic phosphorylation sites have occupancy of 75% or more.

Olsen Science Signaling 3 (2010)

Differential phosphorylation

Gene ontology (GO) analysis of protein and phosphoproteins subcellular localization. All proteins identified by MS were clustered according to their GO annotation for sub-cellular localization (Blue bars). The same clustering was done for all phosphoproteins (Red bars).

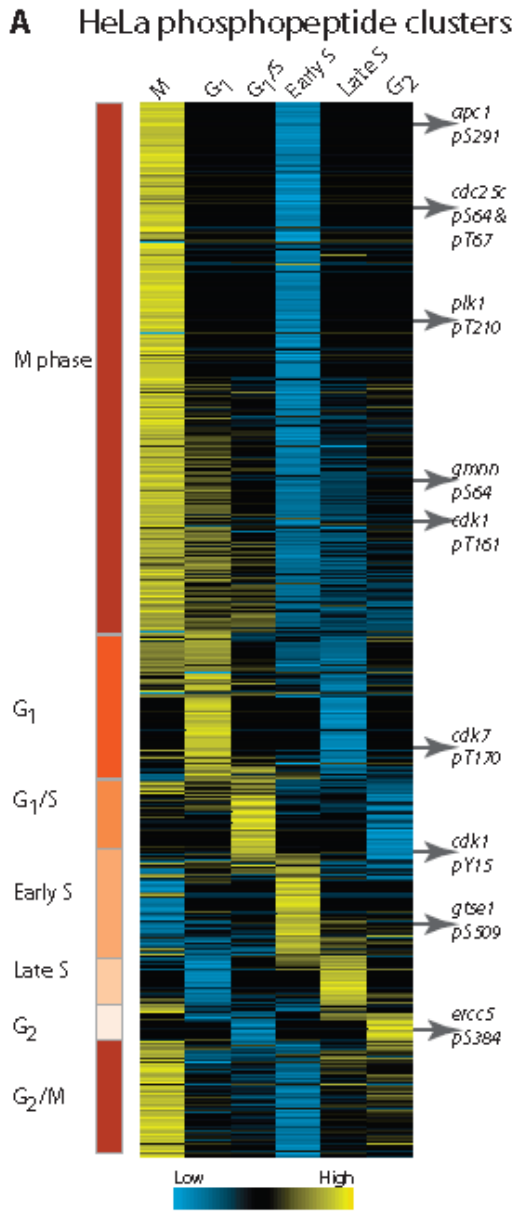


y-axis : percentage of the indicated sub-cellular fractions from the total.

Compared to the proteome distribution, phosphorylated proteins are over-represented in the nucleus and under-represented amongst mitochondrial and secreted proteins.

Olsen Science Signaling 3 (2010)

Dynamics of the Phosphoproteome



Dynamics of the phosphoproteome during the cell cycle.

Clustering of regulated phosphorylation sites in all cell cycle stages.

More than half of all identified regulated phosphorylation sites peak in mitosis.

Olsen Science Signaling 3 (2010)

Who phosphorylates? -> NetPhorest algorithm

COMPUTATIONAL BIOLOGY

Linear Motif Atlas for Phosphorylation-Dependent Signaling

Martin Lee Miller,[1,2*] Lars Juhl Jensen,[2,3*] Francesca Diella,[3] Claus Jørgensen,[4] Michele Tinti,[5] Lei Li,[6] Marilyn Hsiung,[4] Sirlester A. Parker,[7] Jennifer Bordeaux,[7] Thomas Sicheritz-Ponten,[1] Marina Olhovsky,[4] Adrian Pasculescu,[4] Jes Alexander,[8] Stefan Knapp,[9] Nikolaj Blom,[1] Peer Bork,[2,10] Shawn Li,[6] Gianni Cesareni,[5] Tony Pawson,[4] Benjamin E. Turk,[7] Michael B. Yaffe,[8†] Søren Brunak,[1,2†] Rune Linding[4,8,11†]

(Published 2 September 2008)

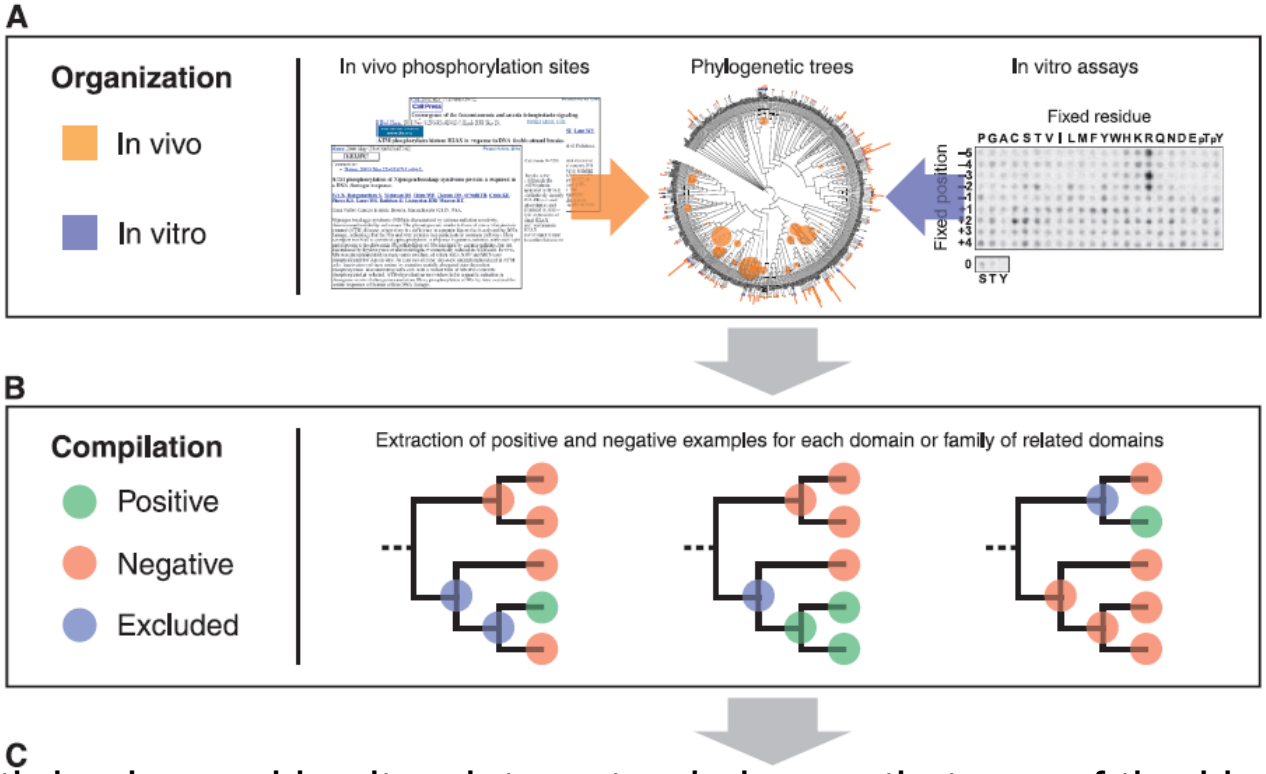
Systematic and quantitative analysis of protein phosphorylation is revealing dynamic regulatory networks underlying cellular responses to environmental cues. However, matching these sites to the kinases that phosphorylate them and the phosphorylation-dependent binding domains that may subsequently bind to them remains a challenge. NetPhorest is an atlas of consensus sequence motifs that covers 179 kinases and 104 phosphorylation-dependent binding domains [Src homology 2 (SH2), phosphotyrosine binding (PTB), BRCA1 C-terminal (BRCT), WW, and 14–3–3]. The atlas reveals new aspects of signaling systems, including the observation that tyrosine kinases mutated in cancer have lower specificity than their non-oncogenic relatives. The resource is maintained by an automated pipeline, which uses phylogenetic trees to structure the currently available in vivo and in vitro data to derive probabilistic sequence models of linear motifs. The atlas is available as a community resource (http://netphorest.info).

Miller Science Signaling 1 (2008)

NetPhorest algorithm

Analyze *in vivo* protein-phosphorylation sites that are linked to at least one kinase [Phospho.ELM] or phospho-binding domain [DOMINO].

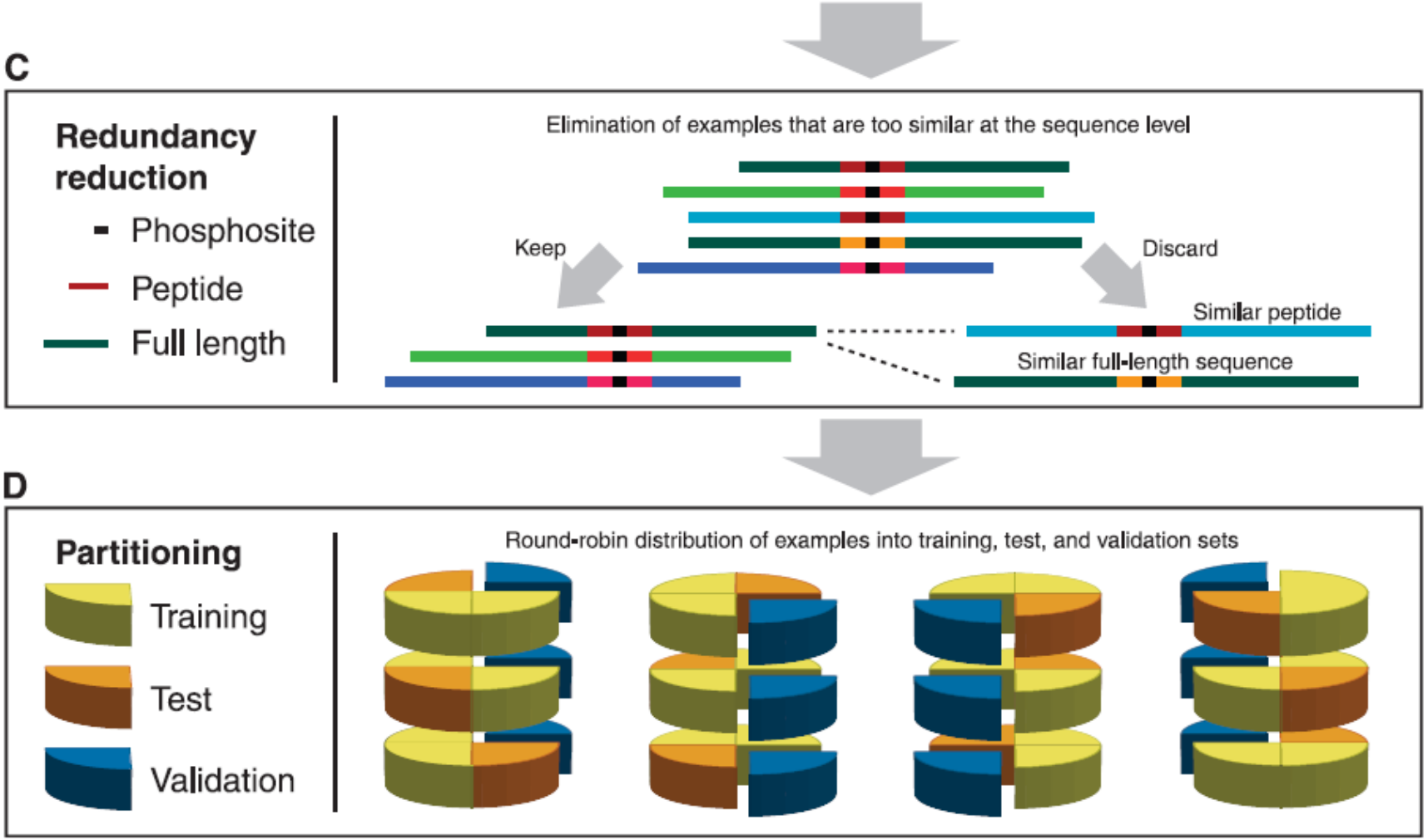
Analyze *in vitro* assays that interrogate kinase specificity by degenerate peptide libraries



map both in vivo and in vitro data onto phylogenetic trees of the kinase and phospho-binding domains, which capture how similar the domains are to one another and thereby how likely they are to have similar substrate specificities.

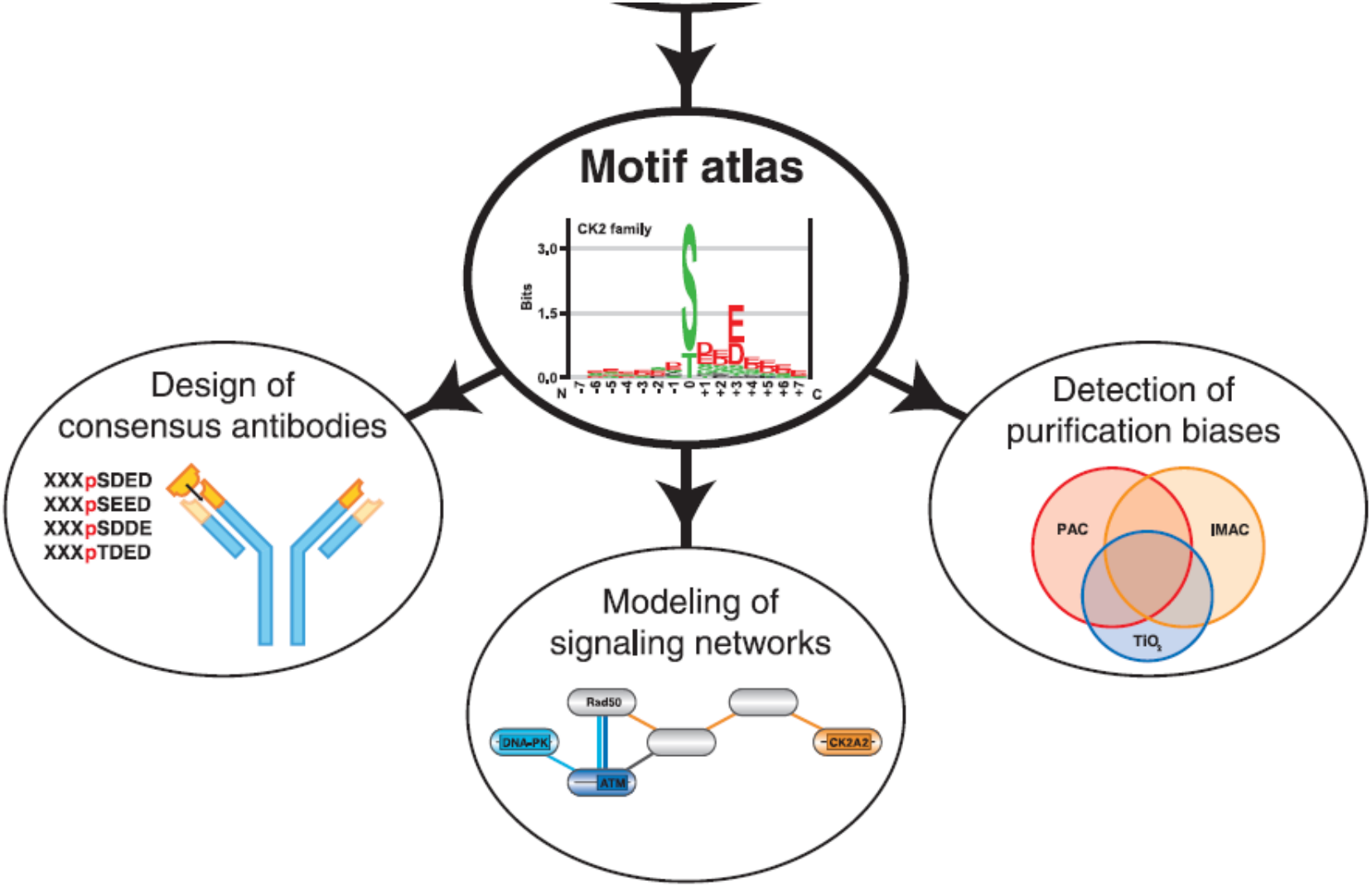
Miller Science Signaling 1 (2008)

NetPhorest algorithm



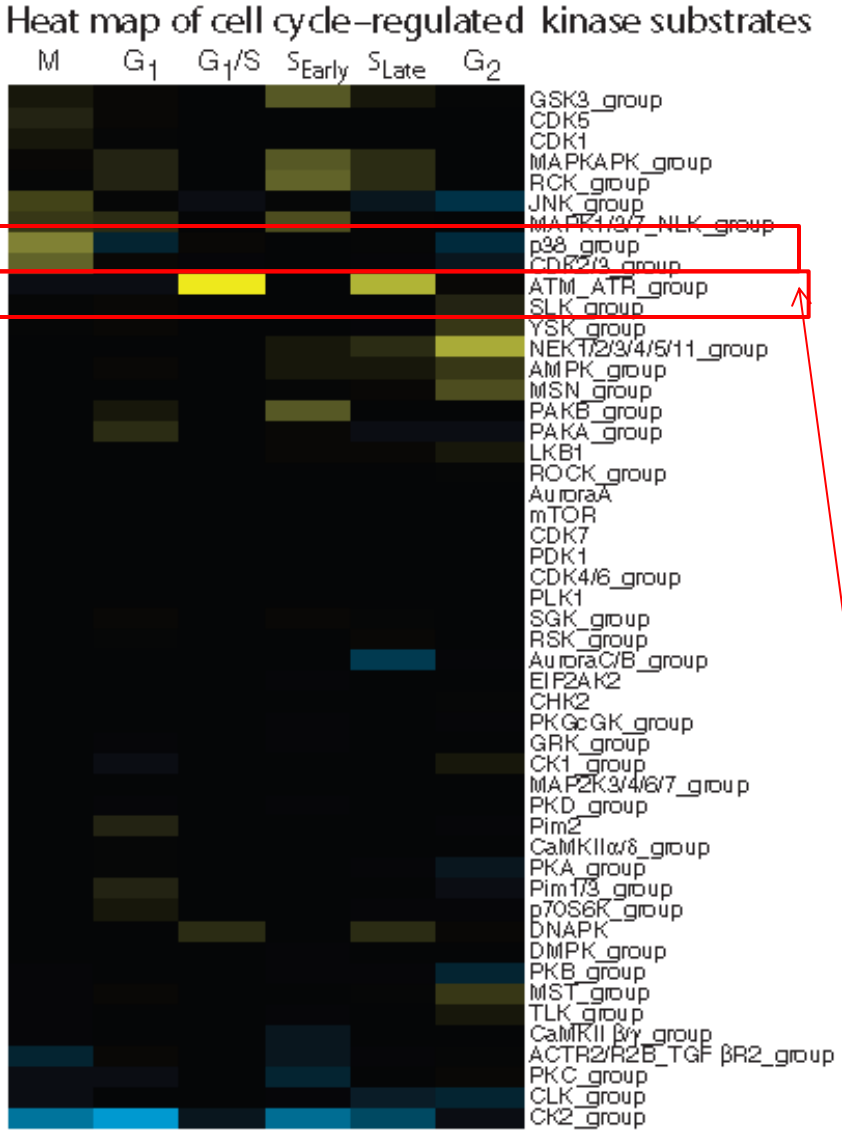
Miller Science Signaling 1 (2008)

NetPhorest algorithm: applications



Miller Science Signaling 1 (2008)

Cell Cycle-regulated kinase substrates



The NetPhorest algorithm was used to predict kinase-substrate relationships of all serine and threonine phosphorylated proteins.

The heat map shows over- (yellow) and underrepresentation (blue) of predicted kinase substrates during different stages of the cell cycle compared to a background of phosphorylation sites that did not change with the cell cycle.

Predicted CDK2 and CDK3 substrates were most highly phosphorylated in M phase.

ATM_ATR substrates are high in S phase.

PPI network of DDR kinase substrates

Substrates of the DNA damage response (DDR) kinases ATM, the related kinase ATR, and DNA-dependent protein kinase (DNA-PK) are significantly overrepresented in S phase.
This is likely due to coupling between DNA replication and repair..

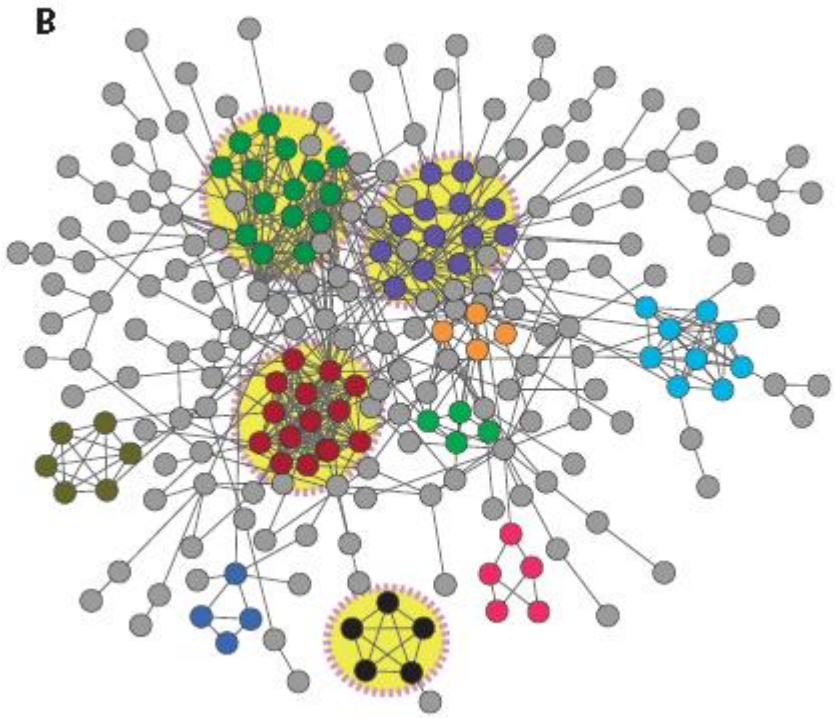


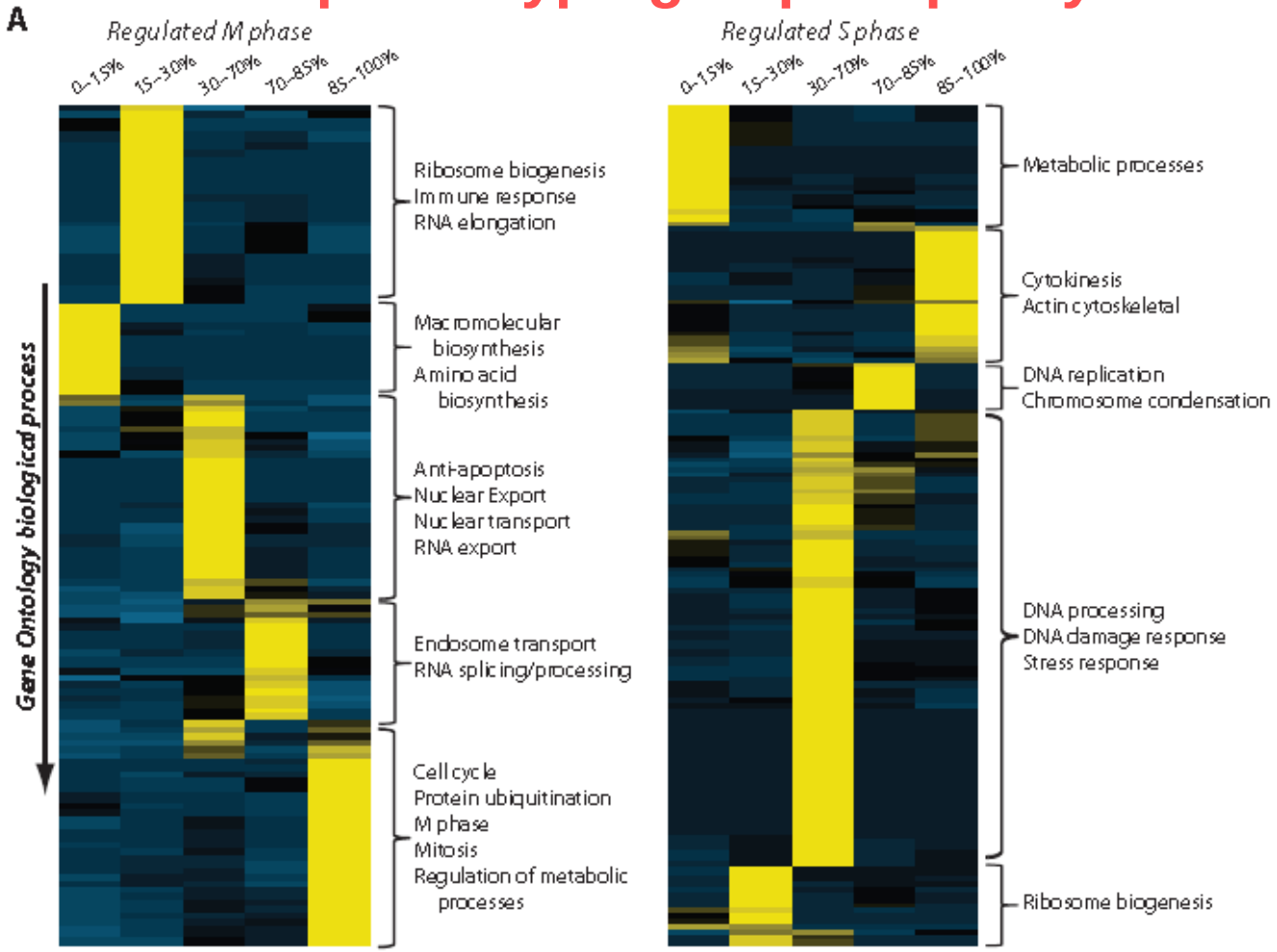
Figure shows the protein-protein interaction network of DDR kinase substrates.

The network was extracted from the STRING database

The color-coded nodes belong to 10 significant protein clusters.

Olsen Science Signaling 3 (2010)

Proteomic phenotyping of phosphorylation site stochiometry



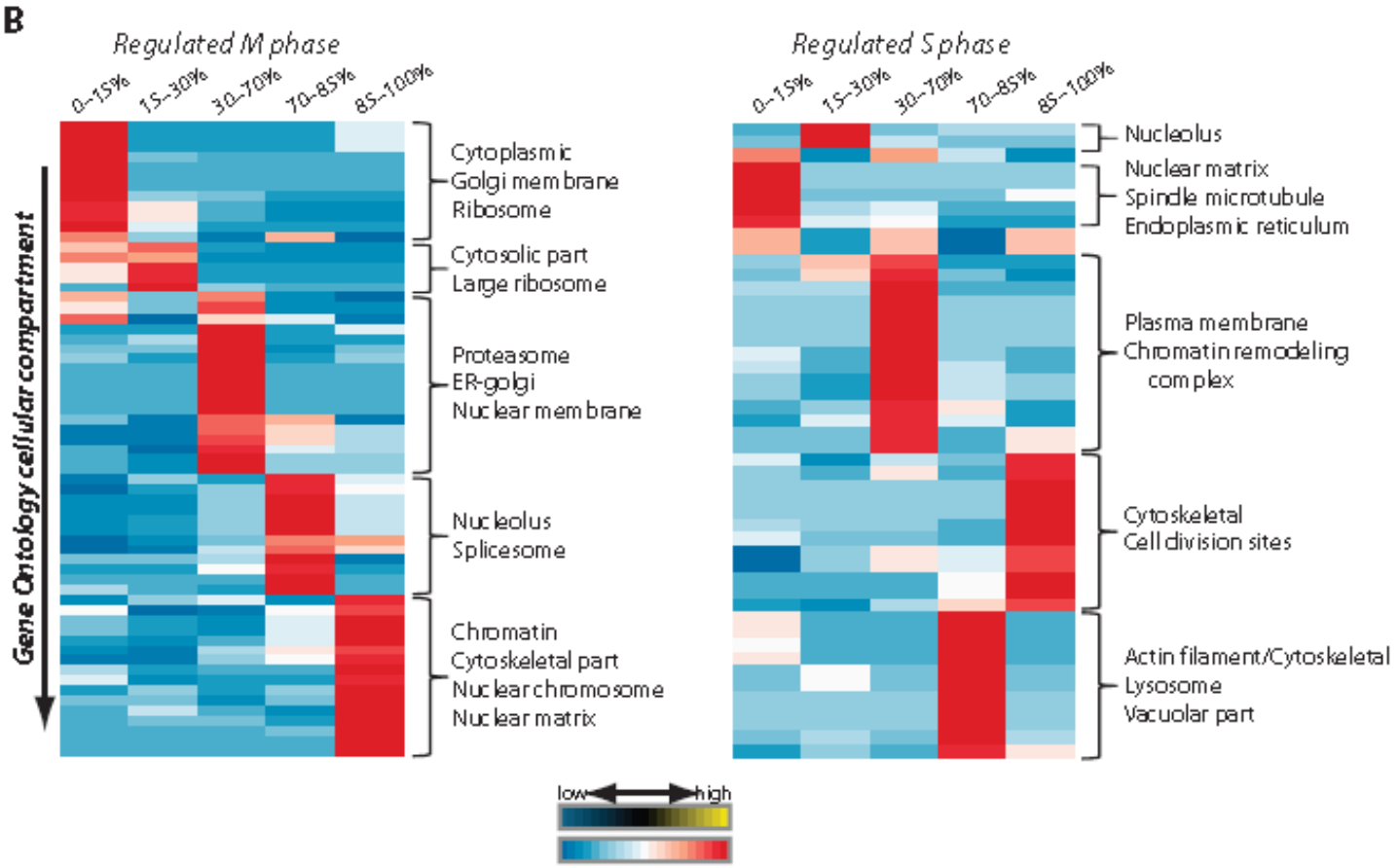
Phenotypic phosphoproteome comparison organized by GO biological process for mitotic (left) and S phase (right) cells.

Proteins involved in metabolic processes have high-occupancy phosphorylation sites during mitosis, but low-occupancy sites during S phase

color scale: yellow, high overrepresentation; dark blue, high underrepresentation.
The phospho proteins were divided into five quantiles on the basis of their maximum phosphorylation-site occupancy and analyzed for GO category (biological process and cellular compartment) enrichment by hypergeometric testing.

Olsen Science Signaling 3 (2010)

GO cellular compartment analysis



Proteomic phenotype analysis of GO cellular compartment level.

Olsen Science Signaling 3 (2010)

Summary

Phosphorylation of protein residues is an important mechanism to regulate protein structure, protein activity, protein localization, and protein interactions.

About 70% of all cellular proteins are phosphorylated to some extent.

Phosphorylation is a dynamic state variable during the cell cycle.

Phosphorylation levels are controlled by the ca. 518 different human kinases as well as by phosphatases.

-> these are important potential drug targets (problem is achieving specificity)