Experimental and Computational Tools for Analysis of Signaling Networks in Primary Cells

Erwin M. Schoof[1] and Rune Linding[1]

[1]Cellular Signal Integration Group (C-SIG), Center for Biological Sequence Analysis (CBS), Department of Systems Biology, Technical University of Denmark (DTU), Lyngby, Denmark

Cellular information processing in signaling networks forms the basis of responses to environmental stimuli. At any given time, cells receive multiple simultaneous input cues, which are processed and integrated to determine cellular responses such as migration, proliferation, apoptosis, or differentiation. Protein phosphorylation events play a major role in this process and are often involved in fundamental biological and cellular processes such as protein-protein interactions, enzyme activity, and immune responses. Determining which kinases phosphorylate specific phospho sites poses a challenge; this information is critical when trying to elucidate key proteins involved in specific cellular responses. Here, methods to generate high-quality quantitative phosphorylation data from cell lysates originating from primary cells, and how to analyze the generated data to construct quantitative signaling network models, are presented. These models can subsequently be used to guide follow-up in vitro/in vivo validation studies. *Curr. Protoc. Immunol.* 104:11.11.1-11.11.23. © 2014 by John Wiley & Sons, Inc.

Keywords: phosphorylation • mass spectrometry • network biology • primary cell signaling

CELL SIGNALING IN PRIMARY CELLS

Cellular responses to environmental stimuli are driven primarily by information processing in signaling networks. Cells receive multiple input cues simultaneously at any given time, and have to decide on appropriate cellular responses such as apoptosis, proliferation, differentiation, or migration (Manning et al., 2002; Jorgensen et al., 2009). Post-translational modifications (PTMs) are an important mechanism for cells to accomplish this, as they alter protein activity and interactions as required by a given cellular response. For example, PTMs can direct and modulate the binding of protein domains to a specific motif on a substrate protein (Pawson, 1995; Seet et al., 2006), thereby altering the kinetics of a protein-protein interaction. Although many types of PTMs exist, such as ubiquitination, acetylation, or methylation, phosphorylation events are among the most extensively used by the cell. They largely govern cellular information processing, and have been demonstrated to be involved in most fundamental biological and cellular processes such as protein-protein interactions, enzyme activity, and immune response (Miller and Berg, 2002; Cannons and Schwartzberg, 2004; Seet et al., 2006; Readinger et al., 2009).

When considering phosphorylation-based signaling within the immune system, it is well established that many immune responses are evoked through the activation of specific receptors on the cell surface by, for example, ligand or antigen binding. These, in turn, activate protein kinases to generate a cellular response through specific phosphorylation network dynamics. For example, Syk, Tec, Src, and protein kinase C (PKC) family kinases have extensively demonstrated to be involved in immune responses involving T-cell activation upon antigen presentation (Monks et al., 1998; Isakov and Altman,

Biochemistry of Cell Activation

2002; Miller and Berg, 2002). Due to its integral role in overall cellular functioning, dysregulation of phosphorylation-based signaling often causes severe changes to the cellular phenotype by evoking distinct alterations to normal cellular responses. This, combined with their ubiquitous nature, implicates them in many human diseases, and the modulation of their dynamics constitutes potential treatment targets (Shawver et al., 2002; Tan et al., 2009; Fedorov et al., 2010; Lemmon and Schlessinger, 2010).

To establish causal relationships between observed phosphorylation events and their effects on signaling networks, one must decipher not only the kinase-phosphosite relationships (i.e., which kinase(s) phosphorylate(s) which phosphorylation sites/substrates), and which phosphatases and phospho-binding domains (e.g., SH2, BRCT, or PTB domains) dephosphorylate and interact, respectively, with the observed phosphorylation sites. Additionally, insight must be gained into the biochemical effects that the modification of these sites exerts on the cellular signaling proteins and networks and, ultimately, how these alter cellular phenotypes or behavior (Cantley et al., 1991; Pawson and Hunter, 1994; Pawson, 1995; Pawson and Kofler, 2009; Brognard and Hunter, 2011; Creixell et al., 2012).

Here, methods to generate high-quality, quantitative phosphorylation data from cell lysates originating from primary cells, such as monocyte-derived immature dendritic cells, are described. The strategy for accomplishing this involves: (1) performing cell lysis, protein digestion, and peptide labeling (see Basic Protocol 1); (2) separating the peptides according to charge state, and allowing the fractions to be subsequently enriched for phosphopeptides separately (SCX fractionation; see Basic Protocol 2); (3) performing specific enrichment techniques that need to be deployed in order to boost the detection of phosphopeptides (see Basic Protocol 3); (4) purifying samples for MS analysis (see Basic Protocol 4); and (5) analyzing the generated data to construct quantitative signaling network models, which can be used to guide follow-up in vitro/in vivo validation studies (see Basic Protocol 5).

Several hurdles must be overcome when studying phosphorylation-based signaling (in primary cells). First, the intrinsically low signal-to-noise ratio of phosphorylation events due to their low abundance and low stoichiometry compared to non-phosphorylated peptides (Jin et al., 2010) represents a challenge that significantly increases the complexity of detecting these events. A second challenge is the transient nature of kinase-substrate interactions, which, due to the high off-rate (k_{off}) of a kinase-substrate interaction, often renders it infeasible to determine experimentally the substrates of a particular kinase using conventional affinity-based biochemistry methods such as tandem affinity purification (TAP) or immunoprecipitation (IP) MS (Burckstummer et al., 2006; Dyson et al., 2011). These approaches depend on stable interactions between the target proteins and the antibody to separate the antibody-bound proteins from the cell lysate. In this manner, one may be able to enrich for kinases and proteins bound to them, but this does not directly translate to the kinase phosphorylating these proteins, as they may purely exist in a scaffolding complex to bring the kinase in the appropriate cellular context for targeting other substrates. Similarly, in vitro kinase reactions do not reflect the cellular context, and thus the specificity in such assays and kinase peptide arrays do not accurately reflect cellular specificity and often leads to large amounts of false positives (Obenauer et al., 2003; Hjerrild et al., 2004). Kinases and substrates typically interact in a transient manner. This makes cellular (or so-called in vivo) kinase-substrate interactions challenging or impossible to capture by experimental methods alone (Linding et al., 2007).

While mass spectrometry (MS) is now able to identify and quantify thousands of phosphorylated residues from a single sample (Bodenmiller and Aebersold, 2010; Mohammed and Heck, 2011; Monetti et al., 2011; Munoz and Heck, 2011), thereby providing a

robust solution to the first aforementioned challenge (i.e., low signal-to-noise ratio), this technique often cannot solve the aforementioned second challenge (i.e., identifying the responsible kinases for these sites). Moreover, so-called Shokat kinases, which rely on a modified ATP binding pocket within the kinase domain in an attempt to utilize labeled ATP for identifying direct kinase substrates, in addition to their limited kinome-coverage, cannot be readily deployed in primary cells, as the cells need to be stably transfected to obtain the required kinase domain mutations (Shah and Shokat, 2003). This has led to a large knowledge gap between the identification of phosphorylation sites and their regulating kinases, information that is critical when attempting to elucidate kinase-substrate networks. It has thus been demonstrated that a combination of computational and experimental approaches is required. Computational approaches have been developed to address this issue, which, in combination with experimental techniques, can be deployed to decrease the knowledge gap (Linding et al., 2007; Miller et al., 2008; Szklarczyk et al., 2011).

GENERATING QUANTITATIVE PHOSPHO-PROTEOMICS DATA USING MASS SPECTROMETRY

While immunoblotting using phospho-specific antibodies was originally one of the most commonly used techniques to investigate phosphorylation events, the low-throughput nature of this approach, combined with its confined character (phosphopeptide-specific antibodies are required, biased by preconceived notions about which phosphorylation sites/proteins are important), non-linear dynamic range, and inaccurate quantitation, meant global quantitative approaches were desired. In the last decade, MS has been increasingly deployed, as it is able to routinely identify and quantify thousands of proteins in a single analysis, and is much more systematically biased (driven by protein stoichiometry and technical design of the instrument), allowing such biases to at least partially be corrected for (Callister et al., 2006; Prakash et al., 2007). Due to the low signal-to-noise ratio of phosphorylated peptides compared to the non-phosphorylated peptides, specific enrichment techniques need to be deployed in order to boost the detection of phosphopeptides. Several techniques exist for this, ranging from IP-based techniques using broad-spectrum phospho-specific antibodies (e.g., against phospho-tyrosine peptides or peptides with a simple motif, e.g., S/TQ for ATM/ATR kinases) to metal affinity-based approaches such as immobilized metal affinity chromatography (IMAC) or titanium dioxide (TiO$_2$; Kawahara et al., 1990; Tani and Suzuki, 1994; Posewitz and Tempst, 1999; Pandey et al., 2000; Jiang and Zuo, 2001; Tanl et al., 2002; Larsen et al., 2005; Rikova et al., 2007). These methods enable selective enrichment of phosphorylated peptides from a peptide pool, thereby making them more readily detectable for the mass spectrometer. An initial drawback of these approaches was the requirement of a relatively large amount of starting material. This has subsequently been overcome by steadily increasing enrichment efficiency as a result of technological developments, which currently makes it possible to identify several thousands of phosphopeptides from a few hundred micrograms of starting material (Engholm-Keller et al., 2012; Zhou et al., 2013). This, in turn, facilitates the investigation of the phosphorylation dynamics in biological systems where a limited number of cells are available, such as primary cells, cancer stem cells, or blood-circulating cells. Furthermore, sample fractionation techniques such as strong cation exchange (SCX) (Mohammed and Heck, 2011), hydrophilic interaction chromatography (HILIC) (McNulty and Annan, 2008), or electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) (Alpert, 2008) can spread the sample complexity across sample fractions, thereby facilitating greater phosphoproteome coverage by increasing the time available for the mass spectrometer to find unique peptides.

Due to the highly dynamic nature of biological systems, phosphorylation-based signaling networks, phosphoproteomes, or proteomes should not be conceptualized, interpreted,

nor described as static entities. Gaining a deeper understanding of the dynamics within signaling networks and how it relates to cell phenotypes is one of the major current challenges in systems biology (Creixell et al., 2012). To this end, it is important to elucidate the cellular information flow-through ensembles of signaling network states, which can be accomplished by conducting, e.g., time-series experiments or dose-response studies. The number and scale of time-points will depend on the system and biological question at hand, but dynamic monitoring of the system will generally give much more in-depth biological insight into the cellular processes driving a given phenotype (Janes et al., 2005; Miller-Jensen et al., 2007; Kreeger et al., 2010). This also enables one to explore the multivariate nature of cellular signaling (Linding, 2010; Jensen and Janes, 2012), which is based on the notion that cells have to integrate many signaling cues simultaneously, the responses to which are often non-linearly related to each other. This enables cells to integrate the different stimuli and respond with appropriate quantitative phenotypic outcomes. One can, for example, stimulate a biological system with a combination of stimuli, i.e., chemical inhibitors, RNAi, antigens, or antibodies (Pedersen et al., 2010), simultaneously or in a time-staggered manner for more comprehensive signaling network models to be constructed (Saez-Rodriguez et al., 2009). These can subsequently guide efforts to formulate so-called network-drugs, which target specific signaling network states rather than individual proteins (Pawson and Linding, 2008; Erler and Linding, 2010; Creixell et al., 2012; Lee et al., 2012).

In cell culture, a quantitative tool can easily be introduced through isotopic labeling, commonly known as stable isotope labeling by amino acids in cell culture (SILAC) (Ong et al., 2002). The principle in SILAC is the incorporation of non-radioactive isotopes through several (typically four to seven) cell divisions to ensure full isotope incorporation. While this is a very powerful approach for several cell types, it is not a suitable option for primary cells, as they can only undergo a limited, pre-determined number of divisions in culture, if any at all. Rather, a post-culture labeling method where proteins/peptides are labeled after cell lysis is a more effective approach. Several techniques for this exist, the primary ones being isobaric tag for relative and absolute quantitation (iTRAQ), tandem mass tag (TMT), or stable isotope dimethyl labeling (Thompson et al., 2003; Ross et al., 2004; Boersema et al., 2009). These labeling strategies all work based on the principle of adding a small but detectable mass shift to the cellular peptides to be able to mix, process, and subsequently analyze them together. The mass shift introduced by the labeling can be detected by the mass spectrometer, and used to trace the sample origin of a given peptide (e.g., to a specific time point, treatment condition, cell type, etc.). This also enables direct comparison of the abundance of the differentially labeled peptides, thus strengthening the quantitation of peptides. For simplicity, this unit focuses on the dimethyl-labeling method. The only limitation of this method compared to iTRAQ or TMT is that while the latter can be used to simultaneously label and compare up to eight samples simultaneously, dimethyl labeling is limited to triplex analysis. This does suffice for many experimental setups, however, and is comparable to the widely used SILAC approach. In general, dimethyl labeling is recommended as an appropriate and powerful default experimental approach for the study of signaling in primary cells, while more complex analyses (comparing more than three samples) would benefit from iTRAQ or TMT approaches.

Finally, this protocol focuses mainly on TiO_2-based enrichment, which, given the relatively higher abundance of phosphorylated serine (pSer) and threonine (pThr) residues in comparison to tyrosine phosphorylation (pTyr), will produce a larger number of pSer/pThr identifications than pTyr (Olsen et al., 2006). TiO_2-based enrichment gives rise to a significant global phosphoproteome coverage, while a fraction of the pTyr events can still be captured using TiO_2. However, pTyr enrichment using pTyr-specific antibodies such as pTyr-100/1000 or 4G10 is highly recommended if one desires a specific

focus on tyrosine kinases or pTyr signaling (Rikova et al., 2007; Jorgensen et al., 2009). In this case, the appropriate experimental protocols that are supplied by the antibody manufacturers are recommended. Nevertheless, despite the fact that TiO_2-based enrichment does not target tyrosine signaling specifically, it is still capable of identifying some of these events (Olsen et al., 2006). It is also worth pointing out that due to the high inter-connectedness of kinase-substrate signaling networks, pSer/pThr signaling events can still give insight to a particular phenotype in cases where high pTyr involvement is expected, as they are likely to also be utilized by the cell as "down-stream" effectors to achieve a specific response (Samelson et al., 1986; Dustin, 2009).

CELL LYSIS, PROTEIN DIGESTION, AND DIMETHYL LABELING

The following protocol can in principle be applied to any type of primary cells of interest, and should be done immediately after the experimental aim has been achieved (e.g., stimulation, mixing of cells, drug/antigen exposure) and preferably in a time-point-dependent manner. The number of cells to start with depends on the availability, but this protocol is optimized for protein amounts ranging from 2 to 24 mg of protein, or ∼20 to 200 million cells. For an overview of the complete experimental workflow, see Figure 11.11.1; this protocol focuses on cell lysis, protein digestion, and peptide labeling.

Materials

Cell line(s) of interest
Phosphate-buffered saline (PBS; Sigma, cat. no. P5368), ice cold
Modified RIPA buffer (see recipe), ice cold
Acetone, HPLC-grade (Sigma, cat. no. 650501), −20°C
Denaturation buffer (see recipe)
Bradford reagent (Sigma, cat. no. B6916)
Dithiothreitol (DTT; Sigma, cat. no. 43815)
Chloroacetamide (CAA; Sigma, cat. no. 22790)
Lysyl endopeptidase (Lys-C; Wako, cat. no. 129-02541; 0.5 μg/μl stock solution made up in MilliQ water)
Triethyl ammonium bicarbonate (TEAB; Sigma, cat. no. T7408)
Trypsin (Sigma, cat. no. T6567; 0.5 μg/μl stock solution made up in 50 mM acetic acid)
Trifluoroacetic acid (TFA; Sigma, cat. no. T6508)
Acetic acid (Fisher Scientific, cat. no. A35-500)
Dimethyl labeling solution (see recipe)

15- or 50-ml tubes
Sonicator
Refrigerated centrifuge
Axial rotator
SepPak C18 columns (Waters, cat. no. WAT020515)
10-ml syringe (polypropylene)

Additional reagents and equipment for Bradford assay (Bradford, 1976)

Perform cell lysis and digestion

1. Remove the cell medium and wash cells two times with ice-cold PBS to remove any serum-containing medium. For adherent cells, pour out the medium, add ∼20 ml of PBS for a 15-cm dish (use more or less according to culture vessel used), briefly swirl by hand, and discard. Repeat this process two times. For non-adherent cells, spin down cells 3 min at $300 \times g$, 5°C, in a 15-ml tube, remove the supernatant, and

Biochemistry of Cell Activation

11.11.5
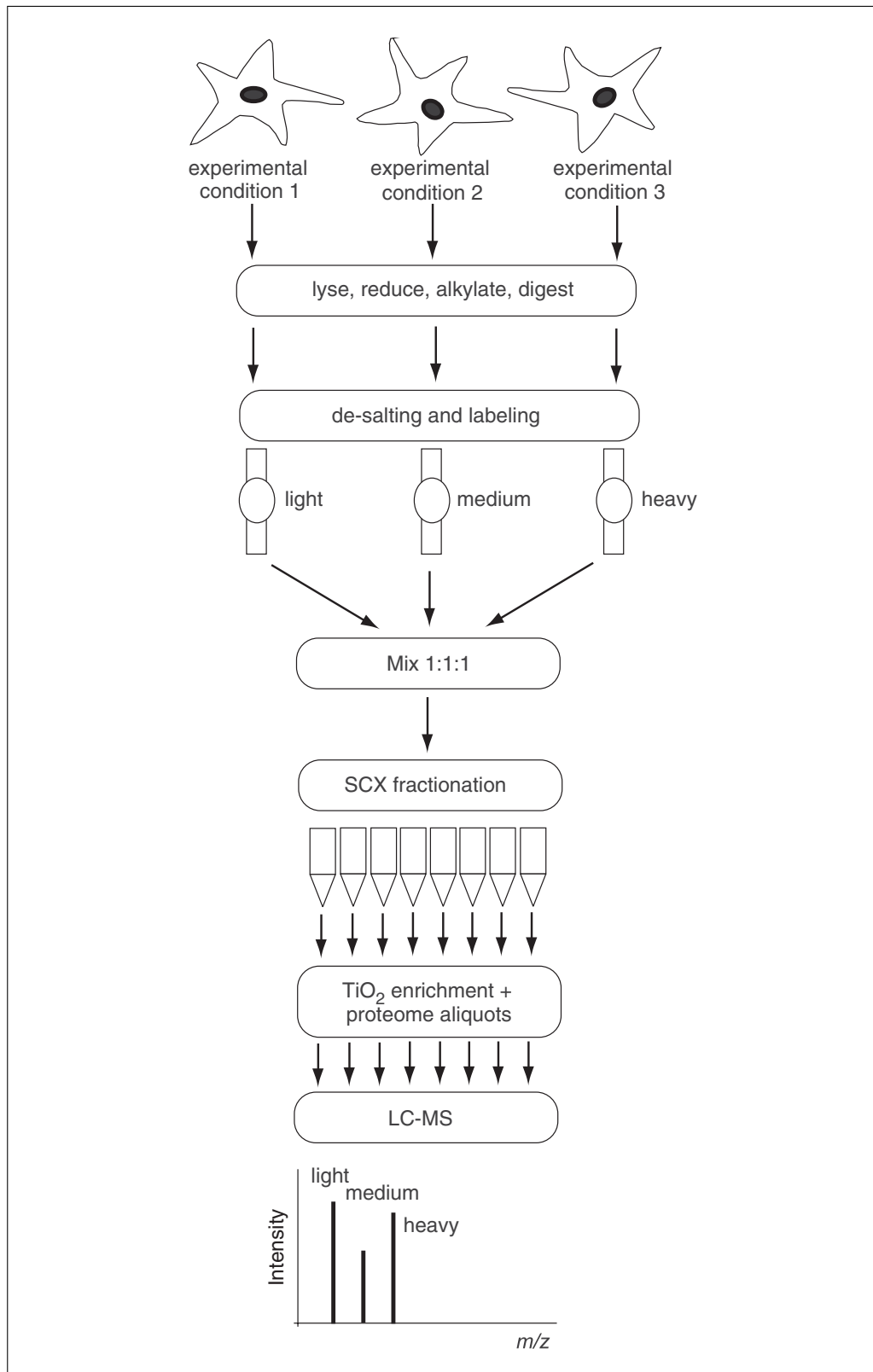
Figure 11.11.1 Experimental workflow overview, highlighting the key components of the sample preparation procedure.

add 10 ml of ice-cold PBS, pipetting up and down carefully. Repeat this process two times.

2. Remove the PBS from the final washing step, add 1 to 2 ml ice-cold RIPA buffer per 10×10^6 cells; if working with adherent cells, scrape the plates, otherwise pipet the cells and lysis buffer up and down until full lysis is achieved. Subsequently, transfer lysate to a 15- or 50-ml tube on ice (depending on total lysate volume), and sonicate on ice three times, 10 sec each time.

3. Centrifuge 20 min, full speed ~4500 × g, 4°C.

4. Transfer supernatant to a clean 50-ml tube, and add ice-cold acetone (−20°C) to a final concentration of ~80% acetone. Place at −20°C and precipitate proteins overnight.

5. Centrifuge 5 min at 2000 × g, 4°C, to pellet the proteins, and discard the acetone by decanting, being careful not to disturb the protein pellet.

6. Add sufficient denaturation buffer to a final concentration of ~5 to 10 mg/ml, and leave for a few hours to overnight at room temperature on an axial rotator to completely dissolve the protein pellet.

Determine protein concentration

7. Determine exact protein concentration using a Bradford assay (Bradford, 1976), either in cuvette- or 96-well plate format.

8. Add 1:1000 (v/v) of 1 M DTT to achieve a final concentration of 1 mM, and incubate 1 hr at room temperature on an axial rotator.

9. Add 1:100 (v/v) of 500 mM CAA to achieve a final concentration of 5 mM, and incubate 1 hr at room temperature in the dark on an axial rotator.

10. Check that the pH is 8, and add 1 µg of lysyl endopeptidase (Lys-C) per 100 µg of protein (1:100). For larger amounts of protein (>10 mg), add 1 µg of Lys-C per 200 µg of protein (1:200). Incubate ~4 to 5 hr at room temperature on an axial rotator.

   If the pH needs to be adjusted, use a very low volume of 1 M NaOH or HCl.

11. Dilute sample(s) 1:4 with 50 mM TEAB in water to reduce (Thio)urea concentration, and check that pH is 8.0 to 8.5 (adjust, if necessary, with 1 M NaOH or HCl).

12. Add 1 µg of Trypsin per 100 µg of protein (1:100). For larger amounts of protein (>10 mg), add 1 µg of Trypsin per 200 µg of protein (1:200). Incubate overnight at room temperature on an axial rotator.

13. Add TFA to a final concentration of 2% (using 20% TFA stock solution) to de-activate any remaining Trypsin, and centrifuge the acidified peptide mixture 5 min at 2000 × g, 20ºC, to clarify and transfer the supernatant to a clean tube.

De-salt samples and perform dimethyl labeling (adapted from Boersema et al., 2009)

If dimethyl labeling is not to be performed, skip step 19, the other steps must be performed for desalting purposes.

14. For each sample that is to be labeled, prepare a SepPak column by attaching a 10-ml syringe to it, after having removed the plunger.

15. Add 5 ml of 100% acetonitrile to each syringe, and allow it to run through the SepPak column by gravity.

Biochemistry of
Cell Activation

11.11.7

If necessary and if no vacuum manifold is available, minimal pressure can be applied by replacing the plunger into the top of the syringe, but never push the plunger down beyond the rubber part sitting at the top of the syringe as this may put too much pressure on the column. The whole SepPak/dimethyl labeling process takes between 2 and 4 hr for optimal results.

16. Wash the SepPak column two times with 4 ml of 0.6% acetic acid solution each time, again allowing gravity to pull solution through the column.

17. Load equal amounts of each sample (as previously determined by the Bradford assay) onto their respective SepPak columns and allow gravity flowthrough; depending on the sample volume this can take a while.

18. Wash the SepPak column with 5 ml of 0.6% acetic acid solution.

19. Flush each SepPak column with 1 ml of its respective labeling reagent, repeating this procedure five times to ensure complete labeling.

    Again, this process could take a while and should take at least 10 min to ensure complete labeling.

20. Wash the SepPak column with 5 ml of 0.6% acetic acid solution.

21. Elute the labeled peptides from the SepPak column two times with 2 ml of 80% acetonitrile plus 0.6% acetic acid, each time.

22. Mix the differentially labeled samples, and proceed to Basic Protocol 2 for SCX fractionation, or Basic Protocol 3 if no SCX fractionation will be done (recommended for protein amounts <2 mg) and the sample will be directly enriched for phosphopeptides.

    In this case, if one is interested in analyzing proteome samples, other fractionation techniques such as gel-based fractionation (Schirle et al., 2003), HILIC fractionation (McNulty and Annan, 2008), or Offgel fractionation (Michel et al., 2003; Hörth et al., 2006) can be deployed to gain better proteome coverage.

SCX FRACTIONATION

To spread sample complexity over several fractions, protein samples >2 mg are recommended to be subjected to SCX fractionation. This will separate the peptides according to the charge state, and allow the fractions to be subsequently enriched for phosphopeptides separately, thereby gaining a better phosphoproteome coverage. This protocol covers sample injection, running the gradient, and subsequent pooling of fractions.

This protocol has been adapted from Olsen and Macek (2009).

Materials

Sample
Acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
SCX buffer A (see recipe)
SCX buffer B (see recipe)
Loading buffer: 1%TFA and 2% acetonitrile in MS H_2O

HPLC/FPLC system (e.g., GE Healthcare AktaMicro)
1-ml SCX column or equivalent (e.g., Resource S 1ml; GE Healthcare Resources)
2-ml microcentrifuge tubes

1. Load sample into the LC system as per manufacturer's instructions.

2. Load the peptides onto an equilibrated 1-ml SCX column as per manufacturer's instructions, and elute the peptides into clean 2-ml microcentrifuge tubes over a

30-min period using the following gradient: 1% to 30% SCX buffer B gradient, followed by 5 column volumes of 100% SCX solvent B, and ending the gradient with 5 column volumes of 100% SCX buffer A to equilibrate the column.

> *Make sure to collect all of the sample, including the flow-through and final equilibration fractions. Fractionation on the Resource S 1 ml column should be carried out at a flowrate of 1 ml/min.*

3. Pool some of the fractions according to their chromatographic peaks, while obtaining about eleven pools of fractions, which can be individually enriched for phosphopeptides.

> *The flow-through (early fractions) consists mainly of multiply phosphorylated peptides and will not bind to the SCX column; it is therefore recommended to pool, and sequentially enrich this pooled fraction for phosphopeptides at least three times.*

4. If desired, dispense proteome samples into aliquots to be able to compare the phosphoproteome with the proteome.

> *While exact amounts depend on the chromatography and amount of sample loaded, pipetting 5 to 10 µl from each pooled fraction is generally sufficient, and one should aim to have about six samples in total for MS analysis.*

5. Acidify and reduce the acetonitrile concentration of the proteome samples with 100 µl of loading buffer, and keep for several hours at 4°C until the StageTipping stage; process as soon as possible.

TITANIUM DIOXIDE PHOSPHOPEPTIDE ENRICHMENT

This protocol covers the TiO_2-based enrichment procedure that enriches samples for phosphorylated peptides. These steps should be carried out at room temperature.

This protocol has been adapted from Thingholm et al. (2006) and Olsen and Macek (2009).

NOTE: This protocol is intended for SCX-fractionated samples. For non-fractionated samples, adjust step 3 to sequentially incubate the single sample three to five times separately to enrich for the majority of the phosphopeptides.

Materials

TiO_2 beads (GL Sciences, cat. no. 5020-75010)
TiO_2 loading solution (see recipe)
SCX samples (see Basic Protocol 2)
SCX buffer B (see recipe)
TiO_2 washing solution 1 (see recipe)
TiO_2 washing solution 2 (see recipe)
Acidification buffer (see recipe)
TiO_2 elution buffer 1 (see recipe)
TiO_2 elution buffer 2 (see recipe)

Automated sample shaker (e.g., Eppendorf Thermomixer)
End-over-end rotator
Centrifuge
C8 StageTips (Thermo Fisher, cat. no. SP321)
10-ml luer-lock syringes and StageTip adaptor (Millian, cat. no. HAM-31330)
96-well PCR plates
Vacuum centrifuge with microplate rotor (e.g., Thermo Savant SC250)
Litmus paper
Vortex

1. Make up the TiO_2-bead slurry solution by mixing ~1.5 mg TiO_2 beads per sample with 6 µl of TiO_2 loading solution, and put on an automated sample shaker (e.g., Eppendorf Thermomixer at 1400 rpm) for 15 min at room temperature. For example, when analyzing fifteen SCX fractions, mix 25 mg TiO_2 beads with 100 µl of TiO_2 loading solution.

2. Add 6 µl of the TiO_2 slurry to each sample, keeping the beads well suspended in the slurry in between sample loading, by briefly vortexing the slurry prior to transferring 6 µl to each sample. Incubate 30 min with end-over-end rotation at room temperature.

3. Centrifuge sample tubes 5 min at 2000 × g, room temperature, to pellet the TiO_2 beads, and for the most concentrated fractions (the flow-through and single-peak fractions, based on chromatography), transfer the supernatant to a clean tube and re-incubate with an additional 6 µl of TiO_2 slurry for 30 min. For all other fractions, aspirate off the supernatant, resuspend pellet in 100 µl of SCX buffer B, and transfer to a clean microcentrifuge tube. Keep at 4°C while the other samples are incubating, repeating this process until the flow-through has been enriched three to five times, each time storing the beads in a clean microcentrifuge tube for MS sample preparation.

4. Centrifuge all samples 5 min at 800 × g, room temperature, and aspirate supernatant.

5. Resuspend beads in 100 µl TiO_2 washing solution 1.

6. Centrifuge all samples 5 min at 800 × g, room temperature, and aspirate supernatant.

7. Resuspend samples in 50 µl TiO_2 washing solution 2, and transfer each sample to a separate C8 StageTip, pipetting sample onto the top of the pipet tip in order for the beads to collect on top of the C8 filter.

8. Flick the sample down into the StageTip using a wrist motion, and push the TiO_2 washing solution 2 through the filter using a syringe, leaving only the TiO_2 beads behind.

9. Pipet 40 µl of acidification buffer into one well for each sample of a 96-well PCR plate, as this improves phosphopeptide stability. Elute the phosphopeptides into the PCR plate (one well per C8 StageTip) using one application of 20 µl TiO_2 elution buffer 1, and one application of 20 µl TiO_2 elution buffer 2.

10. Vacuum centrifuge samples for ~55 min (time is dependent on the model of vacuum centrifuge used), without heat, until the total volume for each sample is ~20 µl. While waiting for this step to complete, one can prepare the C18 StageTips for final peptide purification before MS analysis according to Basic Protocol 4 (up to step 4).

11. Add 20 µl of acidification buffer, and check that pH <2 using litmus paper. In case of high pH (due to, e.g., insufficient ammonia removal during SpeedVac), add an additional 20 µl of acidification buffer until the pH is <2.

12. Cover the PCR plate, briefly vortex (not too vigorously), and centrifuge 1 min (without vacuum) to get the entire sample down into the well.

MASS SPECTROMETRY SAMPLE PREPARATION

Following completion of Basic Protocols 1 through 3, the samples are ready to be purified for MS analysis using C18 StageTips (Rappsilber et al., 2007).

Materials

Methanol, HPLC-grade (Sigma, cat. no. 34860)
Buffer B (see recipe)

Sample buffer (see recipe)
Samples (see Basic Protocol 1, 2, or 3)
Buffer A (see recipe)
Loading buffer (see recipe)

C18 StageTips (Thermo Fisher, cat. no. SP301)
10-ml luer-lock syringes and StageTip adaptors (Millian, cat. no. HAM-31330)
Vacuum centrifuge (e.g., Thermo Savant SC250)
Mass spectrometer with nanospray source (e.g., Thermo Fisher Q Exactive or
    Orbitrap Fusion)

1. Clearly label each C18 StageTip for the sample that is to be loaded onto it.

2. Prime the StageTips with 20 μl methanol, flicking the StageTip down using a wrist
   motion to get the liquid down into the C18 filter, and subsequently slowly pushing
   through the liquid using a syringe. Always ensure that a small amount of liquid
   remains on top of the filter to keep it from drying out.

   *If many samples are to be prepared, one can opt to use a microcentrifuge for spinning the
   liquid through the C18 filter. In this case, place the StageTip into a pipet adaptor placed
   inside an empty 2-ml microcentrifuge tube. Spin the tips at ~800–1000 × g to allow the
   liquid to spin through in ~30 sec.*

3. Push 20 μl of buffer B through the StageTips.

4. Wash StageTips two times with 20 μl sample buffer, each time.

5. Slowly push the previously prepared (phospho-) peptide samples through the
   StageTips.

6. Wash the StageTips two times with 20 μl buffer A, each time.

   *At this stage, the samples can be stored at 4°C, as long as the C18 filter remains covered
   in buffer A. For phosphopeptide samples, the StageTips should not be stored for longer
   than 1 to 2 weeks, whereas proteome samples can be stored for weeks. Long-term storage
   (several months) of both types of samples can be done at −80°C.*

7. Just before MS analysis, elute the purified StageTips two times with 20 μl buffer B,
   each time.

8. Vacuum centrifuge the eluted peptides for ~15 min (time is dependent on exact model
   of vacuum centrifuge) until ~5 μl total volume remains, then add 5 μl loading buffer
   and mix the sample well by pipetting up and down. Briefly vortex and spin down for
   1 min to collect the entire sample in the bottom of the well.

9. Run 5 μl of each sample on a mass spectrometer with nanospray source according to
   the manufacturer's instructions.

   *For example, run 2-hr gradients on 15-cm columns, and 4-hr gradients on 50-cm columns
   to gain optimal (phospho-) proteome coverage.*

ANALYZING PHOSPHORYLATION DATA AND CONSTRUCTING QUANTITATIVE NETWORK MODELS

After generating the MS data, the raw spectra have to be searched against a protein
database in order to match them against possible peptides from which the observed
proteins and phosphorylation sites can be identified. Several search algorithms exist,
some of the most popular being MaxQuant, ProteomeDiscoverer/SEQUEST, and Mascot
(Link et al., 1999; Perkins et al., 1999; Cox and Mann, 2008; Cox et al., 2011). While
these algorithms can all identify and quantitate peptides and proteins, they have different
accuracies and specific requirements, of which an extensive discussion is beyond the

scope of this unit. Here, focus is on MaxQuant, as it is a relatively user-friendly tool that enables custom confidence thresholds to be set, is actively maintained (Cox and Mann, 2008), and is free-of-charge. Moreover, if samples are not labeled, MaxQuant allows one to conduct label-free quantitation. However, this approach suffers from lower accuracy than labeling-based quantitation as it compares peptide abundances between samples, which have been prepared and analyzed separately. Therefore, they are likely to be affected by (slightly) different sample preparation and analysis conditions, which gives rise to artificial experimental artifacts that will influence the data. If possible, one should therefore opt for labeled approaches, but it may nevertheless still prove useful in specific cases where labeled approaches are impossible (Cox and Mann, 2008).

As briefly introduced earlier, an important aim when constructing quantitative phosphorylation networks is to derive crucial kinase-substrate interactions, which may be involved in the phenotype that one is investigating. An increased or decreased activity of one or several kinase(s) involved in this phenotype will likely be manifested in modulated phosphorylation sites, which may show higher or lower abundance. Interpreting phosphorylation site modulation allows for the determination of kinases that are differentially active between experimental conditions. However, a common pitfall that must be taken into account is the importance of distinguishing whether an increase of phosphorylation site abundance is due to the substrate protein having been phosphorylated more (thereby indicating an increased level of kinase activity), or whether the substrate protein was more abundant, thereby explaining the increased levels of the observed phosphopeptide(s). In the case of trying to determine dysregulated kinase-substrate networks, the latter would give rise to false conclusions and should be avoided where possible. This can be controlled for by comparing the phosphorylation levels to the protein levels, which is why it is critical to, in addition to the phosphoproteomic samples, analyze the proteome samples as mentioned in Basic Protocol 1, step 22. This allows for the normalization of the phosphorylation levels to their respective protein abundance, thereby more accurately acting as a proxy for kinase activity (Wu et al., 2011a; see Fig. 11.11.3).

As mentioned above, inferring kinase activity from phosphorylation levels requires computational analyses, which, based on sequence-motif information of the sequence window around a given phosphorylation site combined with the signaling network context of the kinase-substrate interaction, can predict likely kinases to have phosphorylated observed phosphorylation sites. Several approaches have been published over the years, including GPS (Xue et al., 2008), KinasePhos (Wong et al., 2007), NetPhosK (Miller and Blom, 2009), and Scansite (Obenauer et al., 2003), but here a methodology using NetPhorest (Miller et al., 2008) and NetworKIN (Linding et al., 2007), which are developed in-house and have now been combined into a framework known as KinomeXplorer (Horn et al., unpub. observ.), is described. The main reasons for using these two algorithms are (1) they are kept up-to-date on a regular basis, thereby including the latest knowledge in the field, (2) they have been benchmarked intensively to provide their users with accurate modeling capabilities (Miller et al., 2008), (3) they generate probabilities for their predictions, thus allowing probabilistic integration with other types of data and use of confidence thresholds to filter results, and (4) they provide the user with a convenient Web interface, enabling analysis of large datasets in a semi-automated fashion. Additionally, NetPhorest and NetworKIN will generate predictions for other phospho-binding domains interacting with observed phosphorylation sites, enabling more comprehensive modeling to be conducted. Even though these algorithms do not have complete kinome coverage (222 out of 538 at the time of writing), they have the highest coverage compared to alternatives, and additional kinases will be included as the required data becomes available. Below, a computational workflow that allows for the construction of quantitative phosphorylation signaling networks, potentially highlighting kinases of interest in the biological phenomenon that is being investigated, is described. As the above-mentioned
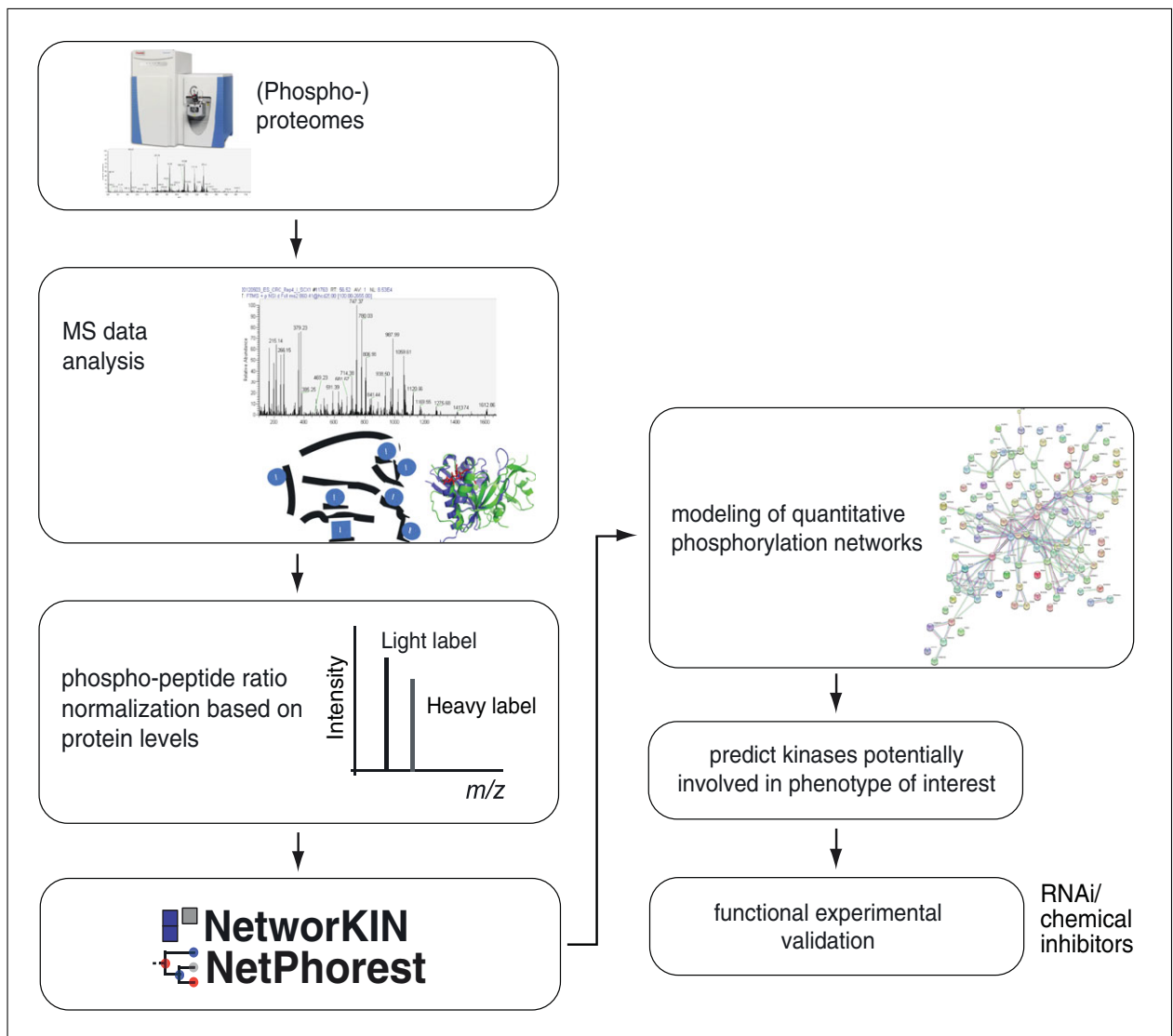
Figure 11.11.2 Modeling workflow overview, detailing the different steps required for constructing quantitative network models that can be used to guide follow-up functional validation in the laboratory.

database-searching software packages provide extensive documentation, here, data analysis steps once the raw MS data has been searched are described and the user has the lists of identified proteins and phosphorylation sites with their corresponding quantitative ratios. See Figure 11.11.2 for an overview of the modeling workflow.

Materials

Desktop computer with Internet access
Mass spectrometry spectral matching software (e.g., MaxQuant or Proteome Discoverer/SEQUEST)
R statistical software
Visual network editing software (e.g., Gephi.org or Cytoscape.org)

1. Conduct a database search for protein and phosphopeptide identification and quantification. In a larger project, it is very useful to rely on a fixed database and release version, e.g., ENSEMBL, to enable easy sequence tracking and mapping. Set the false discovery rate (FDR) to 1% to minimize false-positive protein and phosphorylation site identifications (Elias and Gygi, 2007).

Biochemistry of Cell Activation

11.11.13

2. Filter out the identified phosphorylation sites with a localization probability <0.75, as for these peptides, the exact position of the phosphorylated residue cannot be assigned with reasonable accuracy (Beausoleil et al., 2006; Taus et al., 2011).

3. Transform the protein/phosphorylation site ratios to $\log_2$. This balances out the positive and negative ratios, as down-regulated proteins/phosphorylation sites would otherwise have ratios between 0 and 1, whereas up-regulated proteins/phosphorylation sites would have ratios from 1 to ∞.

   Log_2 transformation ensures a more accurate and direct comparison between up- and down-regulated peptides.

4. Statistically test protein/phosphorylation site ratios for significance. Using the R statistical software package, use the two-sided, unpaired Mann-Whitney Wilcoxon test. Ratios that have a p value of <0.05 can be considered as significantly up/down modulated, whereas ratios with a p value >0.85 should be considered as being non-modulated (Jorgensen et al., 2009).

5. Where possible, normalize modulated phosphorylation site ratios with their respective parent protein ratios. Parent protein ratios ideally are determined from peptides originating from the same protein that cannot contain a PTM (i.e., peptides without serine, threonine, tyrosine, and methionine residues). If at least three unique peptides are observed for a given protein, its ratio can be determined by taking the mean of all unique peptide ratios. This is to ensure that an observed increase in phosphorylation is due to an increase in kinase activity, rather than an increase in substrate protein. This can be accomplished by dividing the phosphorylation site ratio by the protein ratio, as this normalizes the phosphorylation abundance compared to the protein abundance and filters out any phosphorylation site modulation only due to increased protein abundance or degradation (see Fig. 11.11.3). Additionally, protein phosphorylation stoichiometry should be investigated to gain a better perspective of the phosphorylation dynamics (Wu et al., 2011b). This is again to ensure that observed phosphopeptide regulation is due to altered kinase or phosphatase activity, rather than altered protein expression levels or protein degradation.

6. Once the significantly modulated phosphorylation sites have been accurately determined, the NetworKIN and NetPhorest algorithms can be accessed via the portal KinomeXplorer.info to predict the modulating kinases. For this, it is required to know the protein sequence and the absolute location of the phosphorylation site within the protein, which can be extracted from the database search results. This information can be submitted to the KinomeXplorer Website (*http://www.kinomeXplorer.info*), which will generate all possible predicted kinases for the submitted phosphorylation sites. Due to the probabilistic nature of the framework, confidence filtering of the results can be done and one should only include predictions with a score >1. Additionally, as there will generally be multiple predicted kinases for a particular phosphorylation site, only the top scoring kinase and kinases having a probability within 30% of the top scoring kinase should be included for further analysis.

7. To more accurately model the phosphorylation networks, it must be determined which kinases have been experimentally observed in the MS experiment. This can be achieved by, e.g., using the protein identification lists and a filtering method, either through a scripting language (e.g., Python or Perl) or the VLOOKUP function in Excel. At the time of writing, the KinomeXplorer framework has not included this functionality, but this will be implemented shortly. By filtering the kinase predictions to only include kinases that were experimentally observed in the cell type(s) that was analyzed, more in vivo/in vitro relevance can be extended to the in silico predictions. In cases where the phosphoproteome is sequenced enough (i.e., coverage of a
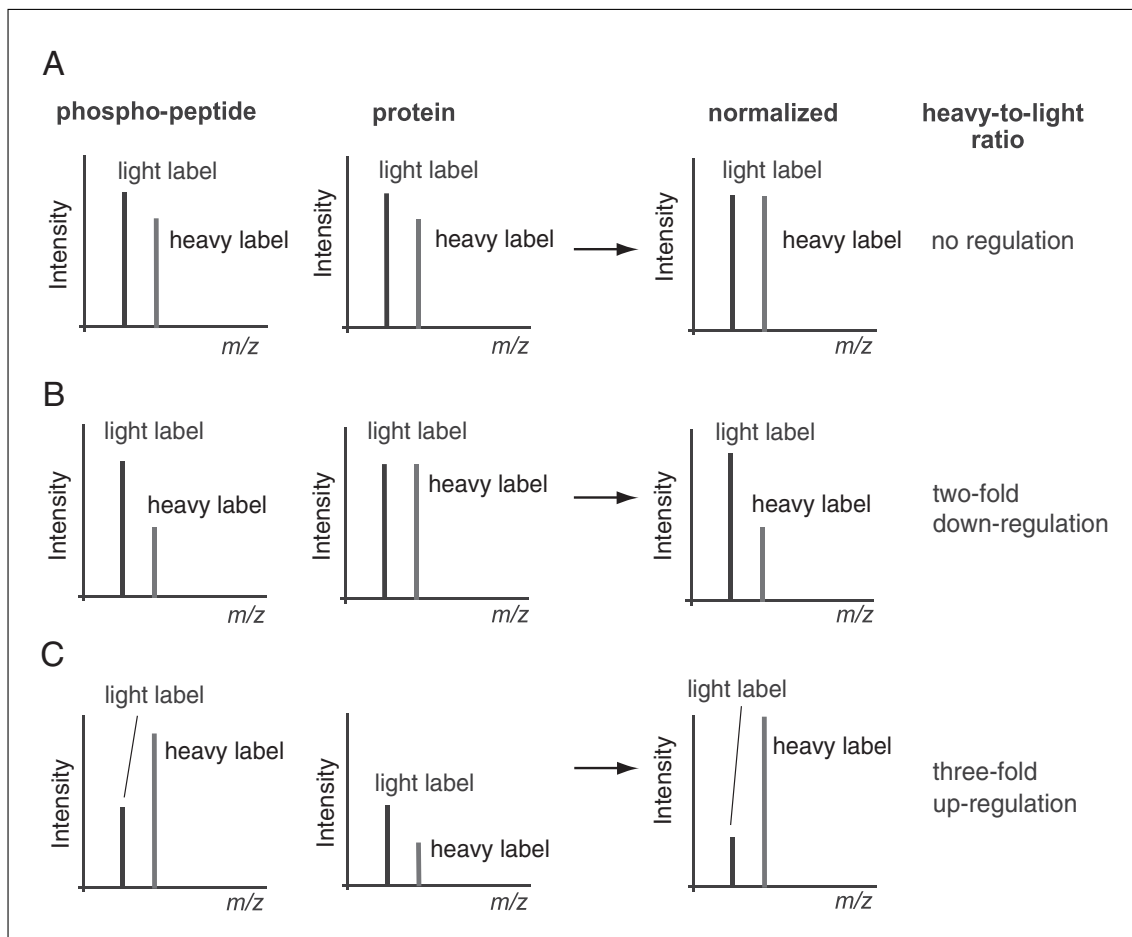
Figure 11.11.3 Phosphorylation site ratio normalization based on protein abundance corrects for protein abundance affecting phosphorylation site abundance, rather than regulated kinase activity. The examples shown are: (**A**) a phosphopeptide that is down-regulated in the heavy labeled sample, whose parent protein is also down-regulated, should be considered as non-regulated. (**B**) A phosphopeptide that is down-regulated in the heavy labeled sample, whose parent protein shows no regulation, should be considered as down-regulated. (**C**) A phosphopeptide that is up-regulated in the heavy sample, whose parent protein is down-regulated, should be considered as an increased up-regulated peptide.

representative subset of the kinome), this principle can be extended to only include kinase predictions from kinases for which a so-called regulatory phosphorylation site has been observed. This is based on the principle that many kinases have a regulatory loop containing a specific residue that is required to be phosphorylated for the kinase to be catalytically activated (or inactivated) (Jorgensen et al., 2009). This will, in the near future, be a built-in function of KinomeXplorer, which will help automate the data processing steps. The regulatory phosphorylation sites, which have been annotated from the literature, can be extracted from public resources such as PhosphoSitePlus (Hornbeck et al., 2012), but it should be noted that deploying this extent of filtering stringency requires considerable depth of kinome coverage in the phosphoproteome data and may not always be feasible. Furthermore, despite on-going efforts, knowledge about these regulatory phosphorylation sites is somewhat limited, so their filtering cannot be applied at a kinome-wide scale.

8. Once the set of kinase predictions has been filtered to include only experimentally supported predicted kinases and their observed substrates, insight into enriched kinase activity and altered signaling networks can be gained. For the former, it can be investigated whether a specific group of kinases is predicted to be more active in one

Biochemistry of Cell Activation

11.11.15

experimental condition than another, by dividing the total number of phosphorylation sites a particular kinase is predicted for by the total number of phosphorylation sites modulated in the same fashion (up/down). This allows for inter-kinase and inter-experimental comparisons, and can elucidate key kinases, which may display different activity levels. To extend this enrichment analysis to a more global (e.g., disease- or condition-specific) level, enrichment should be calculated compared to kinase enrichment in a large collection of known phosphorylation sites such as phospho.ELM (Dinkel et al., 2011) or PhosphoSitePlus (Hornbeck et al., 2012), as this normalizes experimental kinase activity enrichment to a global activity profile (Van Hoof et al., 2009).

9. For a more visual representation and potential mechanistic insight into the signaling network dynamics, an overview of the kinase-substrate interactions can be obtained by importing the filtered predictions into a visual network editor such as Cytoscape (Shannon et al., 2003) or Gephi (Bastian and Heymann, 2009). Here, specific color coding can be deployed to distinguish between up- and down-regulated kinase-substrate interactions, which can often pinpoint specific kinases that become differentially regulated under given experimental conditions or time-points. If one is mainly interested in kinase-kinase networks, it is useful to draw up the networks of kinases that are predicted to phosphorylate each other, together with the observed substrates they are predicted to phosphorylate. This may allow for the elucidation of a core kinase-substrate network, driven by the interaction of several kinases, which may be involved in the phenotype under investigation.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2A; for suppliers, see APPENDIX 5.

Acidification buffer

1% trifluoroacetic acid (TFA; Sigma, cat. no. T6508)
5% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
MilliQ H_2O
Store up to 1 week at room temperature

Buffer A

0.1% formic acid, HPLC-grade (Fisher Scientific, cat. no. A117-50)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 1 month at room temperature

Buffer B

80% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
0.1% formic acid, HPLC-grade (Fisher Scientific, cat. no. A117-50)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 1 month at room temperature

Denaturation buffer

6 M urea (Sigma, cat. no. 15604)
2 M thiourea (Sigma, cat. no. T7875)
10 mM HEPES, pH 8 (Sigma, cat. no. H4034)
Prepare fresh or store upto 6 months at −80°C
Never heat >25°C

Dimethyl labeling solution (Boersema et al., 2009)

Volumes based on one sample that is to be labeled (adjust as necessary):

4.5 ml 50 mM sodium phosphate buffer, pH 7.5 (mix 1 ml of 50 mM NaH_2PO_4 with 3.5 ml of 50 mM Na_2HPO_4)

250 µl 4% (v/v) formaldehyde in MilliQ H_2O (CH_2O for light, CD_2O for medium, $^{13}CD_2O$ for heavy)

250 µl 0.6 M cyanoborohydride in MilliQ H_2O ($NaBH_3CN$ for light or $NaBD_3CN$ for medium/heavy labels)

Store for maximum 24 hr at 4°C

Formaldehyde (CH_2O) (37% (v/v), Sigma, cat. no. 252549)

Formaldehyde (CD_2O) (20%, 98% D, Isotec, cat. no. 492620)

Formaldehyde ($^{13}CD_2O$) (20%, 99% 13C, 98% D, Isotec, cat. no. 596388)

Sodium cyanoborohydride ($NaBH_3CN$) (Fluka, cat. no. 71435)

Sodium cyanoborodeuteride ($NaBD_3CN$) (96% D, Isotec, cat. no. 190020)

Sodium dihydrogen phosphate (NaH_2PO_4) (Merck, cat. no. 1.06346)

Di-sodium hydrogen phosphate (Na_2HPO_4) (Merck, cat. no. 1.06580)

Loading buffer

1% trifluoroacetic acid (TFA; Sigma, cat. no. T6508)
2% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 2 weeks at room temperature

Modified RIPA buffer

50 mM Tris·Cl, pH 7.5 (Sigma, cat. no. T3253)
150 mM NaCl (Sigma, cat. no. S7653)
1% NP40/IgePal (Sigma, cat. no. I8896)
0.5% Na-deoxycholate (Sigma, cat. no. D6750)
1 mM EDTA (Sigma, cat. no. E1644)
β-glycerophosphate (5 mM final concentration) (Sigma, cat. no. G9422), add fresh
NaF (5 mM final concentration) (Sigma, cat. no. S7920), add fresh
Na-orthovanadate (activated; Gordon et al., 1991; 1 mM final concentration) (Sigma, cat. no. 450243), add fresh
Roche complete protease inhibitor cocktail (one tablet added fresh per 10 ml RIPA buffer) (Roche, cat. no. 05 892 791 001)
Store up to 6 months at −20°C

Sample buffer

3% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
1% trifluoroacetic acid (TFA; Sigma, cat. no. T6508)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 2 weeks at room temperature

SCX buffer A

5 mM potassium dihydrogen phosphate (Sigma, cat. no. P9791)
30% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
70% H_2O, HPLC-grade (Sigma, cat. no. 39253)
pH 2.7 with TFA
Store up to 1 month at room temperature

SCX buffer B

5 mM potassium dihydrogen phosphate (Sigma, cat. no. P9791)
350 mM potassium chloride (Millipore, cat. no. 1.04936.0500)
30% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
70% H_2O, HPLC-grade (Sigma, cat. no. 39253)
pH 2.7 with TFA
Store up to 1 month at room temperature

TiO_2 elution buffer 1

5% ammonia solution (Emsure, cat. no. 1.05432.1000)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 3 days at room temperature

TiO_2 elution buffer 2

10% ammonia solution (Emsure, cat. no. 1.05432.1000)
25% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 3 days at room temperature

TiO_2 loading solution

20 mg/ml 2,5-dihydroxybenzoic acid (Sigma, cat. no. 85707)
5% trifluoroacetic acid (TFA; Sigma, cat. no. T6508)
30% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Store up to 1 month at 4°C

TiO_2 washing solution 1

40% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
0.25% acetic acid, HPLC-grade (Fisher Scientific, cat. no. A35-500)
0.5% trifluoroacetic acid (TFA; Sigma, cat. no. T6508)
H_2O, HPLC-grade (Sigma, cat. no. 39253)
Prepare fresh

TiO_2 washing solution 2

80% acetonitrile, HPLC-grade (Sigma, cat. no. 34851N)
0.5% acetic acid, HPLC-grade (Fisher Scientific, cat. no. A35-500)
Store up to 1 week at room temperature

COMMENTARY

Background Information

The techniques described in this unit enable a biological system under investigation to be modeled from the phosphorylation-based signaling perspective, potentially highlighting key proteins involved in a phenotype of interest. By utilizing experimental data as input for computational modeling, more in-depth insight about the signaling networks can be obtained, the results of which can be used to drive follow-up validation studies. In any data analysis approach involving computational predictions, it is of vital importance to experimentally validate (some of) the predictions, as this helps ensure that the predictions are biologically relevant and helps to guide threshold settings. Any key kinases determined in Basic Protocol 5, steps 8 and 9, should be used as input for guiding subsequent experimental validation studies, where the exact role of these kinases in a given phenotype should be functionally assessed by, e.g., RNAi or chemical inhibitor experiments. This can give conclusive evidence of whether or not a kinase or group of kinases are required for a specific phenotype, disease progression, or drug resistance development (Bakal et al., 2008; Jorgensen et al., 2009; Lee et al., 2012). Preferably, this is also done in a time-staggered manner

to monitor the cellular responses to a perturbation/stimulation or combination thereof, as this will elucidate a more complete picture of the altered signaling dynamics within the cell and enables one to tweak the resulting model to higher accuracy. In complex diseases such as cancer, but also in immune response–dependent signaling, this can give more insight into potential treatment strategies, as better understanding of the signaling networks is obtained. By integrating computational and experimental approaches, the strengths of both techniques can be combined, facilitating some limitations of either technique to be (partially) overcome. Finally, the modeling capacities generated by the KinomeXplorer (and underlying NetworKIN and NetPhorest algorithms) framework will grow with the availability of additional kinase-substrate recognition and kinase-substrate interaction data, enabling one to extend established kinase-substrate network models to a kinome-wide level.

Critical Parameters

The most critical parameter of the methods described include conducting different experimental steps in a swift yet cautious manner, due to the labile nature of phosphorylated peptides. Stationary waiting stages should be kept to a minimum, as it is critical to have the enriched samples analyzed as quickly as possible. Additionally, due to the moderately volatile nature of many of the buffers (mainly acetonitrile- and ammonia-containing ones), preparing fresh buffers is imperative and they should be replaced as indicated. Additionally, all of the reagents utilized should be of HPLC quality to reduce contamination of the instrument, and likewise gloves should be worn throughout the protocol to minimize keratin contamination of the sample.

Troubleshooting

In the case of inadequate quantitative data generated, it should be investigated whether this could be attributed to inefficient labeling. The simplest way of checking for this is to run a small aliquot of the labeled samples individually, and to search for unlabeled peptides (Boersema et al., 2009). If this is the case, repeat Basic Protocol 1 until full labeling is achieved. In the case of low phosphoproteome coverage, several possible causes can be identified, and pinpointing the exact one(s) becomes a challenge. Generally, it is vital to ensure a quick lysis procedure with ice-cold

buffers and adequate protease and phosphatase inhibitors as described above, to ensure the preservation of the phosphorylated proteins. Additionally, it is important to monitor pH levels as indicated in the protocol, and to ensure that vacuum centrifugation is done correctly to eliminate organic solvents in the sample. Finally, in the case of lack of specificity during the enrichment (i.e., a large number of unphosphorylated peptides being detected), make sure the washing steps are carried out accurately, removing as much of the supernatant as possible without disturbing the pellet.

Anticipated Results

Depending on the amount of starting material, the number of unique phosphorylation sites that can be identified should range between hundreds and tens of thousands. Using this protocol in house, the authors identify between ∼1000 to 4000 unique phosphorylation sites with <2 mg of starting material without SCX fractionation, and ∼20,000 phosphorylation sites with 24 mg of starting material. Results will vary, however, depending mainly on the biological system under investigation, instrument performance, and experience.

Time Considerations

Lysing of cells requires 1 hr and acetone precipitation should be done overnight. Dissolving protein in denaturation buffer requires between a few hours and an overnight incubation. Reduction and alkylation require 1 day and an overnight digestion. Dimethyl labeling, SCX fractionation, and phospho-enrichment require 1 day. Mass spectrometry depends on the number of samples and gradient times. Data analysis is dependent on the number of samples and computer performance; it will require anywhere between days and weeks.

Acknowledgments

Biochemistry of Cell Activation

11.11.19

Literature Cited

Alpert, A.J. 2008. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal. Chem.* 80:62-76.

Bakal, C., Linding, R., Llense, F., Heffern, E., Martin-Blanco, E., Pawson, T., and Perrimon, N. 2008. Phosphorylation networks regulating JNK activity in diverse genetic backgrounds. *Science* 322:453-456.

Bastian, M. and Heymann, S. 2009. Gephi: An open source software for exploring and manipulating network. International AAAI Conference on Weblogs and Social Media. *https://gephi.org*

Beausoleil, S.A., Villén, J., Gerber, S.A., Rush, J., and Gygi, S.P. 2006. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* 24:1285-1292.

Bodenmiller, B. and Aebersold, R. 2010. Quantitative analysis of protein phosphorylation on a system-wide scale by mass spectrometry-based proteomics. *Methods Enzymol.* 470:317-334.

Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A.J. 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* 4:484-494.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.

Brognard, J. and Hunter, T. 2011. Protein kinase signaling networks in cancer. *Curr. Opin. Genet. Dev.* 21:4-11.

Burckstummer, T., Bennett, K.L., Preradovic, A., Schütze, G., Hantschel, O., Superti-Furga, G., and Bauch, A. 2006. An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat. Methods* 3:1013-1019.

Callister, S.J., Barry, R.C., Adkins, J.N., Johnson, E.T., Qian, W.J., Webb-Robertson, B.J., Smith, R.D., and Lipton, M.S. 2006. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* 5:277-286.

Cannons, J.L. and Schwartzberg, P.L. 2004. Fine-tuning lymphocyte regulation: What's new with tyrosine kinases and phosphatases? *Curr. Opin. Immunol.* 16:296-303.

Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R., and Soltoff, S. 1991. Oncogenes and signal transduction. *Cell* 64:281-302.

Cox, J. and Mann, M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26:1367-1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. 2011. Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 10:1794-1805.

Creixell, P., Schoof, E.M., Erler, J.T., and Linding, R. 2012. Navigating cancer network attractors for tumor-specific therapy. *Nat. Biotechnol.* 30:842-848.

Dinkel, H., Chica, C., Via, A., Gould, C.M., Jensen, L.J., Gibson, T.J., and Diella, F. 2011. Phospho.ELM: A database of phosphorylation sites—Update 2011. *Nucleic Acids Res.* 39:D261-D267.

Dustin, M.L. 2009. The cellular context of T cell signaling. *Immunity* 30:482-492.

Dyson, M.R., Zheng, Y., Zhang, C., Colwill, K., Pershad, K., Kay, B.K., Pawson, T., and McCafferty, J. 2011. Mapping protein interactions by combining antibody affinity maturation and mass spectrometry. *Anal. Biochem.* 417:25-35.

Elias, J.E. and Gygi, S.P. 2007. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* 4:207-214.

Engholm-Keller, K., Birck, P., Størling, J., Pociot, F., Mandrup-Poulsen, T., and Larsen, M.R. 2012. TiSH—A robust and sensitive global phosphoproteomics strategy employing a combination of TiO(2), SIMAC, and HILIC. *J. Proteomics* 75:5749-5761.

Erler, J.T. and Linding, R. 2010. Network-based drugs and biomarkers. *J. Pathol.* 220:290-296.

Fedorov, O., Muller, S., and Knapp, S. 2010. The (un)targeted cancer kinome. *Nat. Chem. Biol.* 6:166-169.

Gordon, J.A. 1991. Use of vanadate as protein-phosphotyrosine phosphatase inhibitor. *Methods Enzymol.* 201:477-482.

Hjerrild, M., Stensballe, A., Rasmussen, T.E., Kofoed, C.B., Blom, N., Sicheritz-Ponten, T., Larsen, M.R., Brunak, S., Jensen, O.N., and Gammeltoft, S. 2004. Identification of phosphorylation sites in protein kinase A substrates using artificial neural networks and mass spectrometry. *J. Proteome Res.* 3:426-433.

Hornbeck, P.V., Kornhauser, J.M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. 2012. PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40:D261-D270.

Hörth, P., Miller, C.A., Preckel, T., and Wenz, C. 2006. Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Mol. Cell. Proteomics* 5:1968-1974.

Isakov, N. and Altman, A. 2002. Protein kinase C(theta) in T cell activation. *Annu. Rev. Immunol.* 20:761-794.

Janes, K.A., Albeck, J.G., Gaudet, S., Sorger, P.K., Lauffenburger, D.A., and Yaffe, M.B. 2005. A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. *Science* 310:1646-1653.

Jensen, K.J. and Janes, K.A. 2012. Modeling the latent dimensions of multivariate signaling datasets. *Phys. Biol.* 9:045004.

Jiang, Z.T. and Zuo, Y.M. 2001. Synthesis of porous titania microspheres for HPLC packings by polymerization-induced colloid aggregation (PICA). *Anal. Chem.* 73:686-688.

Jin, L.L., Tong, J., Prakash, A., Peterman, S.M., St-Germain, J.R., Taylor, P., Trudel, S., and Moran, M.F. 2010. Measurement of protein phosphorylation stoichiometry by selected reaction monitoring mass spectrometry. *J. Proteome Res.* 9:2752-2761.

Jorgensen, C., Sherman, A., Chen, G.I., Pasculescu, A., Poliakov, A., Hsiung, M., Larsen, B., Wilkinson, D.G., Linding, R., and Pawson, T. 2009. Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science* 326:1502-1509.

Kawahara, M., Nakamura, H., and Nakajima, T. 1990. Titania and zirconia: Possible new ceramic microparticulates for high-performance liquid chromatography. *J. Chromatogr. A* 515:149-158.

Kreeger, P.K., Wang, Y., Haigis, K.M., and Lauffenburger, D.A. 2010. Integration of multiple signaling pathway activities resolves K-RAS/N-RAS mutation paradox in colon epithelial cell response to inflammatory cytokine stimulation. *Integr. Biol. (Camb.)* 2:202-208.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jørgensen, T.J. 2005. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* 4:873-886.

Lee, M.J., Ye, A.S., Gardino, A.K., Heijink, A.M., Sorger, P.K., MacBeath, G., and Yaffe, M.B. 2012. Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* 149:780-794.

Lemmon, M.A. and Schlessinger, J. 2010. Cell signaling by receptor tyrosine kinases. *Cell* 141:1117-1134.

Linding, R. 2010. Multivariate signal integration. *Nat. Rev. Mol. Cell. Biol.* 11:391.

Linding, R., Jensen, L.J., Ostheimer, G.J., van Vugt, M.A., Jørgensen, C., Miron, I.M., Diella, F., Colwill, K., Taylor, L., Elder, K., Metalnikov, P., Nguyen, V., Pasculescu, A., Jin, J., Park, J.G., Samson, L.D., Woodgett, J.R., Russell, R.B., Bork, P., Yaffe, M.B., and Pawson, T. 2007. Systematic discovery of in vivo phosphorylation networks. *Cell* 129:1415-1426.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R. 3rd. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* 17:676-682.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. 2002. The protein kinase complement of the human genome. *Science* 298:1912-1934.

McNulty, D.E. and Annan, R.S. 2008. Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol. Cell. Proteomics* 7:971-980.

Michel, P.E., Reymond, F., Arnaud, I.L., Josserand, J., Girault, H.H., and Rossier, J.S. 2003. Protein fractionation in a multicompartment device using off-gel isoelectric focusing. *Electrophoresis* 24:3-11.

Miller, A.T. and Berg, L.J. 2002. New insights into the regulation and functions of Tec family tyrosine kinases in the immune system. *Curr. Opin. Immunol.* 14:331-340.

Miller, M.L. and Blom, N. 2009. Kinase-specific prediction of protein phosphorylation sites. *Methods Mol. Biol.* 527:299-310.

Miller, M.L., Jensen, L.J., Diella, F., Jørgensen, C., Tinti, M., Li, L., Hsiung, M., Parker, S.A., Bordeaux, J., Sicheritz-Ponten, T., Olhovsky, M., Pasculescu, A., Alexander, J., Knapp, S., Blom, N., Bork, P., Li, S., Cesareni, G., Pawson, T., Turk, B.E., Yaffe, M.B., Brunak, S., and Linding, R. 2008. Linear motif atlas for phosphorylation-dependent signaling. *Sci. Signal.* 1:ra2.

Miller-Jensen, K., Janes, K.A., Brugge, J.S., and Lauffenburger, D.A. 2007. Common effector processing mediates cell-specific responses to stimuli. *Nature* 448:604-608.

Mohammed, S. and Heck, A. Jr. 2011. Strong cation exchange (SCX) based analytical methods for the targeted analysis of protein post-translational modifications. *Curr. Opin. Biotechnol.* 22:9-16.

Monetti, M., Nagaraj, N., Sharma, K., and Mann, M. 2011. Large-scale phosphosite quantification in tissues by a spike-in SILAC method. *Nat. Methods* 8:655-658.

Monks, C.R., Freiberg, B.A., Kupfer, H., Sciaky, N., and Kupfer, A. 1998. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395:82-86.

Munoz, J. and Heck, A.J. 2011. Quantitative proteome and phosphoproteome analysis of human pluripotent stem cells. *Methods Mol. Biol.* 767:297-312.

Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. 2003. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* 31:3635-3641.

Olsen, J.V. and Macek, B. 2009. High accuracy mass spectrometry in large-scale analysis of protein phosphorylation. *Methods Mol. Biol.* 492:131-142.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127:635-648.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1:376-386.

Biochemistry of Cell Activation

11.11.21

Pandey, A., Podtelejnikov, A.V., Blagoev, B., Bustelo, X.R., Mann, M., and Lodish, H.F. 2000. Analysis of receptor signaling pathways by mass spectrometry: Identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. *Proc. Natl. Acad. Sci. U.S.A.* 97:179-184.

Pawson, T. 1995. Protein modules and signaling networks. *Nature* 373:573-580.

Pawson, T. and Hunter, T. 1994. Signal transduction and growth control in normal and cancer cells. *Curr. Opin. Genet. Dev.* 4:1-4.

Pawson, T. and Kofler, M. 2009. Kinome signaling through regulated protein-protein interactions in normal and cancer cells. *Curr. Opin. Cell. Biol.* 21:147-153.

Pawson, T. and Linding, R. 2008. Network medicine. *FEBS Lett.* 582:1266-1270.

Pedersen, M.W., Jacobsen, H.J., Koefoed, K., Hey, A., Pyke, C., Haurum, J.S., and Kragh, M. 2010. Sym004: A novel synergistic anti-epidermal growth factor receptor antibody mixture with superior anticancer efficacy. *Cancer Res.* 70:588-597.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551-3567.

Posewitz, M.C. and Tempst, P. 1999. Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* 71:2883-2892.

Prakash, A., Piening, B., Whiteaker, J., Zhang, H., Shaffer, S.A., Martin, D., Hohmann, L., Cooke, K., Olson, J.M., Hansen, S., Flory, M.R., Lee, H., Watts, J., Goodlett, D.R., Aebersold, R., Paulovich, A., and Schwikowski, B. 2007. Assessing bias in experiment design for large scale mass spectrometry-based quantitative proteomics. *Mol. Cell. Proteomics* 6:1741-1748.

Rappsilber, J., Mann, M., and Ishihama, Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2:1896-1906.

Readinger, J.A., Mueller, K.L., Venegas, A.M., Horai, R., and Schwartzberg, P.L. 2009. Tec kinases regulate T-lymphocyte development and function: New insights into the roles of Itk and Rlk/Txk. *Immunol. Rev.* 228:93-114.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., Hu, Y., Tan, Z., Stokes, M., Sullivan, L., Mitchell, J., Wetzel, R., Macneill, J., Ren, J.M., Yuan, J., Bakalarski, C.E., Villen, J., Kornhauser, J.M., Smith, B., Li, D., Zhou, X., Gygi, S.P., Gu, T.L., Polakiewicz, R.D., Rush, J., and Comb, M.J. 2007. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131:1190-1203.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-

Jones, M., He, F., Jacobson, A., and Pappin, D.J. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3:1154-1169.

Saez-Rodriguez, J., Alexopoulos, L.G., Epperlein, J., Samaga, R., Lauffenburger, D.A., Klamt, S., and Sorger, P.K. 2009. Discrete logic modelling as a means to link protein signaling networks with functional analysis of mammalian signal transduction. *Mol. Syst. Biol.* 5:331.

Samelson, L.E., Patel, M.D., Weissman, A.M., Harford, J.B., and Klausner, R.D. 1986. Antigen activation of murine T cells induces tyrosine phosphorylation of a polypeptide associated with the T cell antigen receptor. *Cell* 46:1083-1090.

Schirle, M., Heurtier, M.A., and Kuster, B. 2003. Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol. Cell. Proteomics* 2:1297-1305.

Seet, B.T., Dikic, I., Zhou, M.M., and Pawson, T. 2006. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell. Biol.* 7:473-483.

Shah, K. and Shokat, K.M. 2003. A chemical genetic approach for the identification of direct substrates of protein kinases. *Methods Mol. Biol.* 233:253-271.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. 2003. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13:2498-2504.

Shawver, L.K., Slamon, D., and Ullrich, A. 2002. Smart drugs: Tyrosine kinase inhibitors in cancer therapy. *Cancer Cell* 1:117-123.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L.J., and von Mering, C. 2011. The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* 39:D561-D568.

Tan, C.S., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M.O., Jørgensen, C., Bader, G.D., Aebersold, R., Pawson, T., and Linding, R. 2009. Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci. Signal.* 2:ra39.

Tani, K. and Suzuki, Y. 1994. Syntheses of spherical silica and titania from alkoxides on a laboratory scale. *Chromatographia* 38:291-294.

Tanl, K., Sumizawa, T., Watanabe, M., Tachibana, M. Koizumi, H., and Kiba, T. 2002. Evaluation of titania as an ion-exchanger and as a ligand-exchanger in HPLC. *Chromatographia* 55:33-37.

Taus, T., Köcher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., and Mechtler, K. 2011. Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* 10:5354-5362.

Thingholm, T.E., Jørgensen, T.J., Jensen, O.N., and Larsen, M.R. 2006. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat. Protoc.* 1:1929-1935.

Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K., and Hamon, C. 2003. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 75:1895-1904.

Van Hoof, D., Muñoz, J., Braam, S.R., Pinkse, M.W., Linding, R., Heck, A.J., Mummery, C.L., and Krijgsveld, J. 2009. Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell* 5:214-226.

Wong, Y.H., Lee, T.Y., Liang, H.K., Huang, C.M., Wang, T.Y., Yang, Y.H., Chu, C.H., Huang, H.D., Ko, M.T., and Hwang, J.K. 2007. KinasePhos 2.0: A web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Res.* 35:W588-W594.

Wu, R., Dephoure, N., Haas, W., Huttlin, E.L., Zhai, B., Sowa, M.E., and Gygi, S.P. 2011a. Correct interpretation of comprehensive phosphorylation dynamics requires normalization by protein expression changes. *Mol. Cell. Proteomics* 10:M111.009654.

Wu, R., Haas, W., Dephoure, N., Huttlin, E.L., Zhai, B., Sowa, M.E., and Gygi, S.P. 2011b. A large-scale method to measure absolute protein phosphorylation stoichiometries. *Nat. Methods* 8:677-683.

Xue, Y., Ren, J., Gao, X., Jin, C., Wen, L., and Yao, X. 2008. GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. *Mol. Cell. Proteomics* 7:1598-1608.

Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A.J.R., Zou, H., and Mohammed, S. 2013. Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat. Protoc.* 8:461-480.

Internet Resources

http://www.kinomexplorer.info
Internet portal to access the integrated NetPhorest and NetworKIN frameworks.

http://www.networkin.info
Internet portal to access the original NetworKIN framework.

http://www.netphorest.info
Internet portal to access the original NetPhorest framework.

**Biochemistry of
Cell Activation**

11.11.23