Decoding signalling networks by mass spectrometry-based proteomics

Chunaram Choudhary and Matthias Mann*‡*

Abstract | Signalling networks regulate essentially all of the biology of cells and organisms in normal and disease states. Signalling is often studied using antibody-based techniques such as western blots. Large-scale 'precision proteomics' based on mass spectrometry now enables the system-wide characterization of signalling events at the levels of post-translational modifications, protein–protein interactions and changes in protein expression. This technology delivers accurate and unbiased information about the quantitative changes of thousands of proteins and their modifications in response to any perturbation. Current studies focus on phosphorylation, but acetylation, methylation, glycosylation and ubiquitylation are also becoming amenable to investigation. Large-scale proteomics-based signalling research will fundamentally change our understanding of signalling networks.

SH2 domain
(SRC homology 2 domain). An ~100 amino acid domain that recognizes phosphoTyr residues in a specific sequence context.

PTB domain
(PhosphoTyr-binding domain). Like the SH2 domain, the PTB domain binds to phosphoTyr, but usually binding specificity is determined by the sequence N-terminal to the phosphorylation site.

**The Novo Nordisk Foundation Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark. ‡Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany.*
e-mails: chuna.choudhary@cpr.ku.dk; mmann@biochem.mpg.de
doi:10.1038/nrm2900
Published online 12 May 2010

Cells are information processing devices that integrate myriad extracellular and intracellular signals to produce an optimal output. Many diseases, for example cancer, can be thought of as pathological alterations in signalling networks. Consequently, studying the nature and mechanisms of signalling events is a large and crucial part of biological and medical research. Extracellular ligands can activate plasma membrane-bound receptors, which transmit the signal through a phosphorylation cascade into the cell nucleus, ultimately leading to changes in the expression of specific genes. Signalling events also involve intricate networks, which encompass feedback loops, crosstalk with other signals and the integration of information relating to the internal state of the cell.

FIGURE 1 illustrates the principles of signalling networks using receptor Tyr kinases (RTKs) and tumour necrosis factor-α (TNFα) signalling as examples. In the case of RTKs, ligand binding typically induces dimerization of the receptors, leading to *trans*-phosphorylation of their intracellular region and activation of their kinase domains. These phosphorylation events can create docking sites for the recruitment of downstream signalling effectors, which contain specific modification interaction domains, such as SH2 domains and PTB domains, which recognize phosphorylated Tyr residues in specific sequence contexts[1]. Following these early events, the signal propagates through kinase cascades and adaptor protein interactions to various cellular compartments. The translocation of some activated kinases to the nucleus, and the subsequent phosphorylation of transcription factors, leads to changes in gene transcription. Regulation by many other signals follows the same general principles but there are also important differences. For example, in the case of the inflammatory cytokine TNFα, the initial signalling events involve activation of ubiquitin ligases and the attachment of ubiquitin to signal mediators. Downstream events involve both phosphorylation and ubiquitylation, as in the case of RTKs. Note that different signalling activators initiate multiple signalling pathways and share some of the same components, which is known as signalling crosstalk[2].

Signal propagation involves protein changes on three different levels: regulated protein post-translational modifications (PTMs), including but not limited to phosphorylation, acetylation, ubiquitylation, methylation and gylcosylation; protein–protein interactions, often owing to PTMs; and signal-induced protein expression changes (FIG. 1). All three levels are synchronized in a highly dynamic and often spatially segregated manner and may themselves lead to changes in protein activity, localization and association with small molecules such as phospholipids. A primary task of signalling research is therefore the measurement of PTMs, protein interactions and proteome dynamics. The principles of cell signalling have been worked out over decades using ingenious analytical approaches. One of these is the recognition of proteins and their modifications by specific antibodies: protein–protein interactions are often studied by co-immunoprecipitation followed by western blotting, whereas changes in PTMs are detected by modification-specific antibodies and protein expression changes are

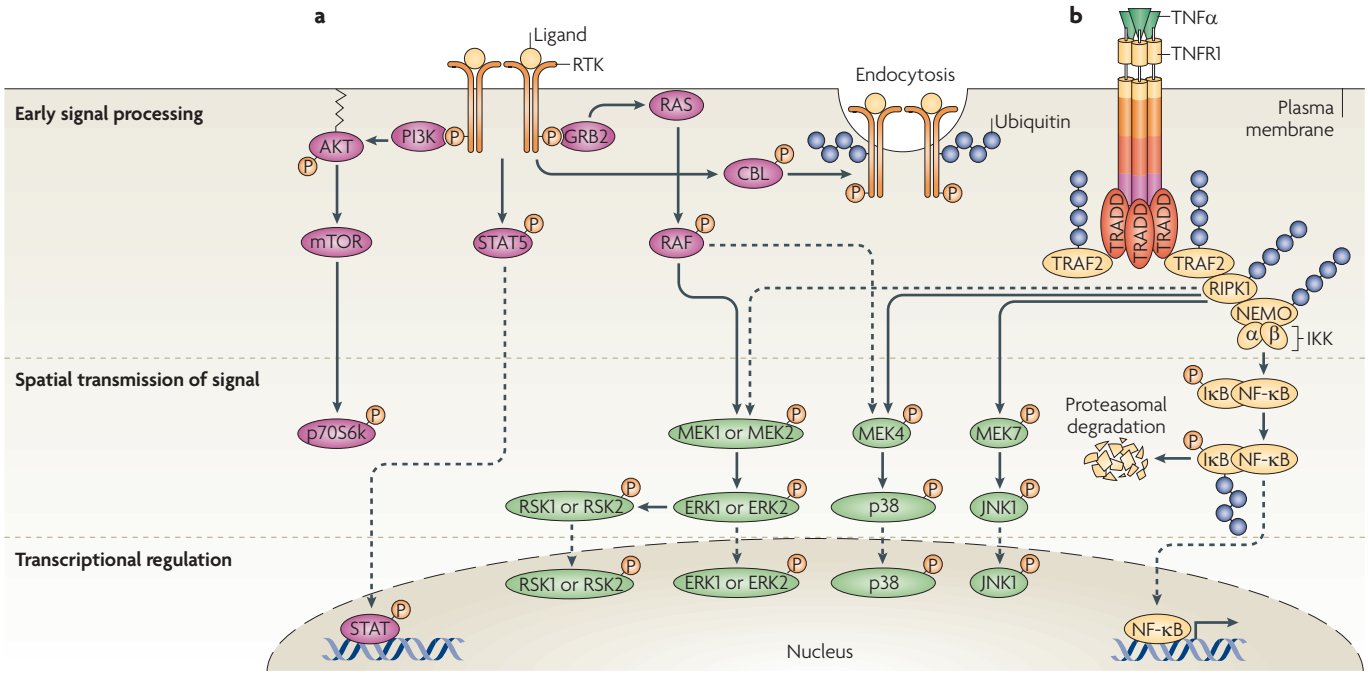

Figure 1 | **General principles of signalling pathways.** Signal processing and amplification by plasma membrane proximal events are followed by communication of the signal to different cellular compartments. This process involves the activation of multiple signal cascades by receptors, different protein post-translational modifications (PTMs), crosstalk between signalling pathways and feedback loops to ensure optimal signalling output. **a** | The binding of receptor Tyr kinases (RTKs) to their cognate ligands at the cell surface results in receptor dimerization and autophosphorylation. Phosphorylated Tyr residues then serve as docking sites to recruit signalling mediators, such as growth factor receptor-bound protein 2 (GRB2). Assembly of these signalling complexes activates multiple signalling cascades such as the phosphoinositide-3 kinase (PI3K)–AKT, Ras–Raf–extracellular signal-regulated kinase (ERK) mitogen-activated protein kinase (MAPK), and signal transducer and activator of transcription (STAT) pathways. Casitas B-lineage lymphoma (CBL)-mediated ubiquitylation of RTKs controls their endocytosis and the duration of receptor signalling. **b** | Binding of tumour necrosis factor-α (TNFα) to its receptor, TNFR1, induces trimerization of the receptor and recruitment of the adaptor protein TNFR1-associated death domain (TRADD), which functions as a hub to assemble a multiprotein signalling complex containing TNFR-associated factor 2 (TRAF2), receptor interacting Ser/Thr protein kinase 1 (RIPK1) and nuclear factor-κB (NF-κB) essential modulator (NEMO). This results in the activation of different signalling networks, such as the ERK MAPK, p38 MAPK and NF-κB pathways. Proteins in the MAPK signalling pathways are activated by both RTKs and TNFα, which allows cells to integrate multiple signals. Dashed lines indicate indirect activation of signalling pathways or translocation of proteins into the nucleus. IκB, inhibitor of NF-κB; IKK, inhibitor of NF-κB kinase; JNK1, Jun N-terminal kinase 1; MEK, MAPK ERK kinase; mTOR, mammalian target of rapamycin; p70S6K, p70 ribosomal S6 kinase-α; RSK, ribosomal protein S6 kinase-α.

determined by the degree of staining on western blots. However, antibody recognition is not always specific and it is difficult to convert antibody signals into a truly quantitative format. Perhaps most importantly, antibodies are inherently directed to a small number of known signalling players and modifications.

In this Review, we discuss the modern quantitative mass spectrometry (MS) techniques that now complement antibody-based approaches. In contrast to most previous techniques, MS-based proteomics is not limited to specific sites or proteins of interest. In this sense, MS-based technology can be unbiased (meaning hypothesis free), global and systems-wide. It therefore represents a fundamentally different approach to studying cell signalling. We describe the currently dominant technologies in MS-based proteomics, discuss methods for enriching for peptides with specific PTMs and review the first applications of MS-based proteomics in the elucidation of global signalling networks.

Liquid chromatography
In high performance liquid chromatography (HPLC), the peptide mixture is separated in liquid phase based on hydrophobic interactions with the C_{18} stationary phase of the chromatography column (C_{18} is the length of the alky chains decorating the chromatographic beads). In proteomics, columns are typically very small (75 μm inner diameter) and flow rates very low (200 nl min^{-1}).

MS-based proteomics for cell signalling

The proteome of a cell, tissue or body fluid is the collection of all proteins and their modifications present in a given state. The generic workflow of modern MS-based proteomics can be decomposed into four layers (FIG. 2). The proteome is extracted and fractioned or enriched for specific proteins or PTMs; peptides are separated and ionized, mass spectra of the peptides and their fragments are obtained and the acquired data are analysed. These steps are discussed below.

MS and MS/MS analysis of peptides. Before peptides can be analysed by MS, proteins must be digested in gel[3] or in solution[4,5] (FIG. 2a). One-dimensional (1D) gel separation of the proteome is usually followed by separation of peptides by on-line liquid chromatography (LC) and subsequent electrospray ionization. This converts peptides to intact ions in the gas phase, ready for MS analysis (FIG. 2b). In contrast, in-solution digestion is followed by

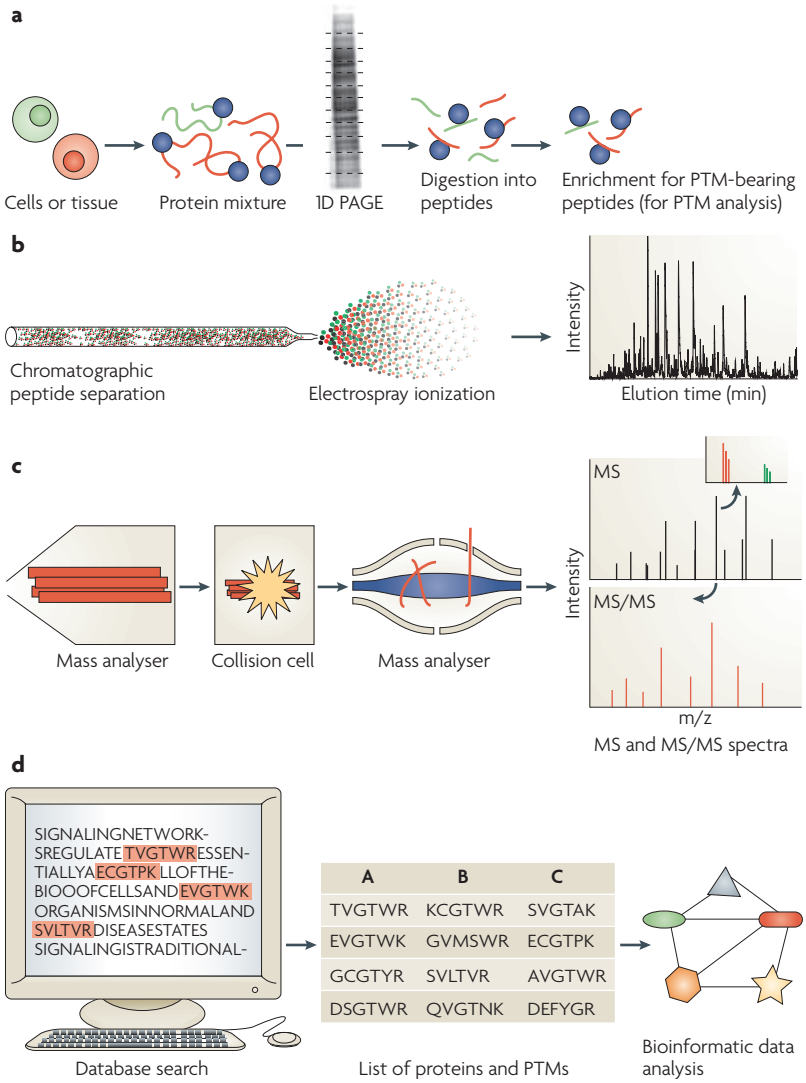

Figure 2 | **Typical work flow for proteome and PTM analysis using shotgun proteomics.** **a** | Proteins extracted from organs, tissues or cells are separated by one-dimensional polyacrylamide gel electrophoresis (1D PAGE) and 'in-gel digested' into peptides using proteases such as trypsin. The peptides containing specific post-translational modifications (PTMs) can be enriched using different approaches (see TABLE 1). Non-modified peptides are used to identify and quantify total cellular proteins. **b** | Purified peptides are separated on a miniaturized reverse phase chromatography column with an organic solvent gradient. Peptides eluting from the column are ionized by electrospray at the tip of the column, directly in front of the mass spectrometer (known as on-line coupling). **c** | The electrosprayed ions are transferred into the vacuum of the mass spectrometer. In the mass spectrometry (MS) mode, all ions are moved to the orbitrap mass analyser, where they are measured at high resolution (top mass spectrum). The first mass analyser then selects a particular peptide ion and fragments it in a collision cell. The inset in the MS panel indicates the stable isotope labelling by amino acids in cell culture (SILAC) ratio of one of the peptides. The MS/MS spectrum can be obtained in the ion trap mass analyser at low resolution or in the orbitrap at high resolution. For modified peptides, the peptide mass will be shifted by the mass of the modification, as will all fragments containing the modification, allowing the unambiguous placement of the PTM on the sequence. **d** | The mass and list of fragment masses for each peptide are scanned against protein sequence databases, resulting in a list of identified peptides and proteins. These lists of proteins and their quantitative changes are the basis for biological discovery.

an additional peptide separation step — usually based on ion exchange chromatography — before on-line LC MS/MS (see below).

Three pieces of information are needed for MS analysis of each peptide: its mass, its ion intensity and a list of its fragments. The mass and fragments identify the peptide, whereas the intensity is used for quantification. To obtain these data the mass spectrometer is used in two different modes. In the MS mode, a spectrum of all peptides eluting from the column at any given time is acquired, yielding the mass and intensity. The mass spectrometer then isolates each peptide species and fragments them by imparting enough energy to break chemical bonds. The mass spectrum of these peptide fragments is called the tandem or MS/MS spectrum. In modern mass spectrometers, ions can be moved in milliseconds between different parts of the instrument. MS and MS/MS can be carried out in the same mass analyser or, in hybrid mass spectrometers, in different mass analysers in the same instrument.

Among many different types of mass spectrometers[6], two configurations are often used in proteomics: quadrupole time-of-flight (TOF) instruments and hybrid linear ion trap–orbitrap instruments. In TOF instruments, peptide ions are separated in time by their arrival at the detector. In the orbitrap mass analyser, the frequency of peptide ions oscillating in the trap is measured and the mass spectrum is obtained by Fourier transformation. MS resolution is an important parameter, regardless of which instrument is being used, because at any given time during the gradient the many peptides co-eluting from the chromatographic column need to be distinguished in the mass spectrum. TOF instruments now have resolution in excess of 10,000 and the orbitrap is routinely used at a resolution of 60,000 (MS resolution is a unit-less quantity). This is a big advance over the ubiquitous ion traps with typical resolution below 1,000, which are therefore generally restricted to the analysis of simple mixtures. Mass accuracy can be in the low parts-per-million range for TOF instruments and even lower for the orbitrap, greatly improving the percentage of peptides that can be identified[7] (BOX 1).

There are also different ways to fragment peptide ions. Most commonly they are collided with a low pressure of an inert gas (collision-induced dissociation (CID); see REF. 8 for an introduction to MS-based peptide sequencing). In ion trap mass spectrometers, peptide ions are resonantly excited by an electric field, leading to an increase in internal energy and fragmentation primarily through cleavage at the peptide bonds. Ion trap fragmentation spectra can be uninformative if the lowest energy pathway only involves the loss of a small chemical group, such as a water molecule or a labile PTM. For this reason, ion traps are usually operated in more complicated modes, with additional fragmentation of these uninformative ions[9]. In the ion trap, peptide fragments are always measured at low resolution. In quadrupole TOF instruments, the quadrupole serves as a mass filter that only passes the ions of a particular mass-to-charge (m/z) ratio. These ions are fragmented in a collision chamber and a TOF spectrum of the fragments is subsequently obtained. These instruments can produce more informative MS/MS spectra because uninformative fragment ions frequently fragment

Box 1 | **Precision proteomics**

A key advance in mass spectrometry (MS)-based proteomics in the past few years has been a huge increase in the quality of the data. The three decade-old technology of two-dimensional gel electrophoresis typically resolved only the products of a few hundred genes at best, was low throughput and had a low dynamic range and, thus, was never a serious foundation for proteomics. Even the analysis of complex peptide mixtures, the current mainstay technology in MS-based proteomics, was previously associated with ion trap mass spectrometers that had a mass spectrometric resolution of only a few hundred and, consequently, could not distinguish co-eluting peptides of similar mass in complex peptide mixtures. This low quality of data resulted in low identification rates (often only a few percent[119]) and precluded accurate quantification. The situation has changed dramatically. First, modern quadrupole time-of-flight instruments achieved medium resolution and mass accuracy in MS and MS/MS spectra. Next, the hybrid linear ion trap–orbitrap instruments made high resolution and high accuracy mass measurements routine without sacrificing robustness, speed or sensitivity[13,120–122]. However, the peptide fragmentation spectra on these instruments were still measured at relatively low resolution in the ion trap part of the instrument. This is called a 'high-low' strategy because MS spectra are taken at high resolution and MS/MS spectra at low resolution. Recently, the introduction of ion sources with much higher ion transmission from the atmosphere to the vacuum of the mass spectrometer, combined with the higher energy coalitional dissociation (HCD) fragmentation method, has made it possible to routinely obtain extremely high mass accuracy in both the survey spectrum and the fragmentation spectrum[10]. This finally allows a routine 'high-high' strategy in shotgun proteomics[123] without loss of analysis depth, and even ensures near certain identification of, for example, peptides defining new genes and unexpected modifications.

Electrospray ionization
An ionization method developed by J. Fenn, for which he shared the 2002 Nobel Prize in chemistry. A liquid is passed through a charged needle, producing electrosprayed droplets that contain the peptides. On evaporation of the solvent, intact and protonated peptides (or other analyte molecules) are left in the gas phase.

Mass analyser
A part of a mass spectrometer that measures mass to charge (m/z) ratios of ions (for example, ionized peptides). Multiplying the m/z value by the charge and subtracting the weight of the charging entity (typically two protons) yields the mass of the peptide. A mass spectrometer can contain several mass analysers of the same or different types, and ions can be moved between these analysers at will.

Fourier transformation
A mathematical operation that transforms one complex-valued function of a real variable (typically a frequency spectrum) into another domain. In Fourier transformation MS, the frequencies associated with ions moving in a trap are mass dependent and this signal is transformed by Fourier transformation into a mass spectrum.

MS resolution
This value is defined as the width of the peak at half height divided by the mass of the peak and is therefore a dimensionless number. High resolution distinguishes co-eluting peptides with similar mass, a prerquisite for unambiguous identification and quantification of peptides.

Chemical derivatization
A chemical method used to transform one chemical compound into a derivative. In proteomics, side chains of amino acids can be chemically modified, which can be used for enriching these peptides from complex mixtures or for quantification of the modified peptide in MS.

further to yield informative ones. The recent introduction of higher energy collisional dissociation (HCD) has made a similar fragmentation mode to that in quadrupole TOF available to linear ion trap–orbitrap instruments, without diminishing sensitivity[10,11]. In HCD, the MS/MS spectrum is analysed with high mass accuracy in the orbitrap analyser.

Two further fragmentation techniques rely on a completely different physical mechanism. In electron capture dissociation (ECD) and electron transfer dissociation (ETD), peptides obtain excess energy from an electron, which neutralizes one of the peptide's positive charges[12,13]. Fragmentation is much faster and more direct than in CID, and often cleaves the backbone of the peptide without cleaving the most labile bonds first. Thus, ECD and ETD are attractive for analysing peptides with labile PTMs, such as the glycosylation of proteins by O-linked β-N-acetylglucosamine (O-GlcNac).

Making MS quantitative. One of the attractions of MS over classical methods in signalling research is its potential for highly accurate quantitative results for thousands of proteins and PTMs. To achieve highest accuracy, the two proteomes to be compared are differentially isotopically labelled. In metabolic labelling, cells incorporate the isotopic label as part of normal biosynthesis[14]. Stable isotope labelling by amino acids in cell culture (SILAC) is one example of metabolic labelling and generally uses Arg and Lys labelled with the stable ^{13}C isotope and/or ^{15}N isotope — after tryptic digestion peptides then contain a labelled amino acid at their carboxy terminus. The known molecular weight difference between the 'light' (normal) and 'heavy' (labelled) amino acid that is used during the growth of the two cell populations allows the proteomes to be distinguished. After mixing the light and heavy cells, they can be fractionated, or otherwise manipulated, without introducing quantitative errors between them[15,16]. Tryptic peptides appear as heavy and light pairs separated by a defined mass offset (for example, heavy $^{13}C_6$- and $^{15}N_4$-labelled Arg has a mass that is 10.00827 daltons higher than light Arg, and heavy $^{13}C_6$- and $^{15}N_2$-labelled Lys has a mass that is 8.0142 daltons higher than light Lys; note that these isotopes are stable and not radioactive). Triple labelling of Lys and Arg is also straightforward using $^{13}C_6$-labelled Arg and D_4-labelled Lys as additional labels that do not introduce overlap between the natural isotope distributions of the SILAC forms. Triple labelling allows the direct comparison of three proteomes, which is useful for three state comparisons and in time-series measurements. Because of its simplicity and accuracy, SILAC has become a method of choice in MS-based signalling research and its use has recently been extended beyond cell culture systems to small mammals[17] and even to human tumour tissues[18].

The most popular chemical labelling technique is isobaric tag for relative and absolute quantification (iTRAQ). Like many other chemical labelling techniques, iTRAQ attaches a chemical group to primary amino groups (the N-terminus and Lys side chains)[19]; however, in iTRAQ, the differential labelling is only apparent in the fragmentation spectra (that is, after the peptides have been fragmented during MS/MS). When working with iTRAQ it is important to exclude 'co-fragmented' peptides that have similar mass and elute at the same time from quantitative analysis, as this would skew observed ratios[20,21]. Many other chemical labelling techniques have also been described. For example, labelling by heavy or light dimethyl groups[22] is very economical and therefore also allows chemical labelling in cases where large amounts of starting material are required[23]. Note that some level of side reaction is unavoidable in chemical derivatization and this may interfere with unbiased PTM analysis[24].

The attraction of label-free quantification, in which peptide signals between different experimental conditions are compared directly, is that no experimental manipulation of the sample is necessary. However, it has been very difficult to control for overall differences in the signals between runs, which, unlike in the case of isotopic labelling, are not automatically taken into account. Several successful applications of label-free

analysis have nevertheless been reported[25–28]. Recent advances in computational methods now allow the sophisticated normalization of signals between runs, and we predict that the role of label-free quantification will increase in the future, especially when precise determination of ratios is not required.

In some cases the absolute number of protein molecules needs to be determined; for example, the protein copy number per cell for systems biology applications. The total MS signal — the added signal intensity of each of the peptides identifying the protein — is a good estimate of the total amount of protein in the cell[29,30]. For highest accuracy in absolute quantification, however, synthesis of isotopically labelled standards is necessary. Ideally, a known amount of a labelled protein is mixed with the sample to eliminate errors introduced by the digestion step[31–33].

Towards complete proteome analysis. The analysis of entire proteomes is challenging because of the large number of peptides that need to be measured and because of the large difference in their abundances — the 'dynamic range' of the proteome. Until recently, proteome experiments typically identified only a few hundred proteins. Technological advances in all areas of proteomics, especially in computational proteomics, have now enabled the analysis of the first complete proteome, that of the budding yeast model organism[29], as judged by independent tagging of each yeast open reading frame[34,35]. Mammalian proteomes are more complex and a single cell line seems to express the protein products of more than 10,000 genes. Current depth of analysis is about 7,000 proteins with a 48 hour measurement time[36]. Thus, comprehensive expression proteomics seems to be within reach, at least in the sense of measuring one protein form for each expressed gene. Greater depth of analysis leads to higher sequence coverage of the identified proteins, which, in turn, increases the chances of distinguishing between isoforms of the same gene by identifying isoform-specific peptides.

Unbiased versus targeted analysis. Sometimes only a few proteins or modified peptides are of interest and sequencing can be restricted to a subset of previously specified peptides[37]. Techniques termed single reaction monitoring (SRM) and multiple reaction monitoring (MRM) have been used for many years in small molecule MS on triple quadrupole mass spectrometers (two quadrupole analysers in a row separated by a collision chamber). The first mass analyser is set to the peptide mass of interest, the collision cell is used to fragment the peptide and the second quadrupole is set to a specific fragment that is characteristic of the peptide (SRM). To increase specificity of this low resolution method, more than one fragment is monitored (MRM). Multiplexing allows several hundred different peptides or phosphopeptides to be measured in single runs[30,38,39]. The attractions of MRM over hypothesis-free proteomics are potentially higher sensitivity and throughput because the mass spectrometer is tuned to only a few peptides and fragmentation events. However, the specificity and false-positive rates in the MRM-based experiments still need to be rigorously established, especially in complex mixtures[40]. In yeast, this method identified 90% of targeted proteins, including a protein with an estimated 41 copies per cell[41]. The same study also monitored expression changes of 46 proteins targeted in the glycolytic pathway on nutrient change.

Bioinformatic analysis of proteomics data. Analysis of proteomics results can be divided into separate areas[42]. The first deals with the analysis of the raw MS data itself and encompasses extraction of peptide signals, identification of the peptides with a search engine and quantification of the proteins. Important points are the certainty of identification, which is now usually set to a false discovery rate (FDR) of 1% but in some studies is as high as 50%. This value can easily be determined by searching a 'nonsense database' that contains the same protein sequences as the normal database, but in which all the sequences are reversed or scrambled. If a given score cut-off leads to 1% hits in this database, the FDR is 1% (REF. 43).

The significance of quantification is often based on a simple cut-off value for convenience, but ideally should be set at a specific p-value[7]. This takes account of the fact that fold changes of as little as a few percent can be significant if the protein is measured with many peptides each covered with many MS scans, whereas for low abundance proteins even a measured factor of two may not be significant. PTM quantification is usually based on single peptides and is therefore more challenging than protein quantification.

After extraction of identified and quantified protein lists, bioinformatics analysis usually involves standard tools, many of which were developed for microarray data. Typical analyses entail gene ontology enrichment, which checks for over-representation of cellular compartments, functions or pathway mapping[44,45]. Given the multidimensional nature of proteomics experiments, integration and mining of these data is a central challenge for the field.

Studying PTMs on a large scale

Expression proteomics uncovers the ultimate end result of any perturbation on the gene expression programme, but these changes are often mediated far upstream by PTMs. MS is ideal for studying PTMs because it is a highly specific but generic detection method[46,47]. Addition of a PTM leads to a defined mass change of the peptide, which is directly measured by MS (TABLE 1). MS/MS then localizes the modification with single amino acid resolution. PTM analysis by MS is nevertheless very challenging because modified peptides may be present in low amounts. It is also more difficult to identify a modified than a non-modified peptide because the modification may lead to more complicated MS/MS spectra (containing fragments with and without the modification), and because of the increased database search space as software algorithms must match against many more possible peptide forms.

False discovery rate
(FDR). A statistical method used in multiple hypotheses testing to correct for multiple comparisons. In a list of positive calls, FDR controls the expected proportion of false positives. In proteomic data analysis, a 1% FDR is currently customary, which means that, at most, 1% of the identified proteins should be false positives.

Table 1 | **PTMs currently amenable to large-scale MS analysis**

PTM	Mass shift (Δm; Da)*	Enrichment methods	Largest MS study‡	Organism	Remarks
Phosphorylation	79.96633	IMAC, TiO_2 and antibodies	20,443 sites[54]	*Homo sapiens* (HeLa cell line)	Fe^{3+}-based IMAC and TiO_2 are most commonly used for enrichment of phosphoSer, phosphoThr and phosphoTyr-containing peptides
Acetylation	42.01056	Pan anti-acetyl-Lys antibodies	3,600 sites[50]	*Homo sapiens* (A459, Jurkat and MV4-11 cell lines)	Trypsin often cannot cleave acetylated peptides; thus, acetyl-Lys is located internally on modified peptides, which aids in the site-specific localization of PTMs
Ubiquitylation (diGly tag)	114.04292	Tagged ubiquitin	110 sites[59]	*Saccharomyces cerevisiae*	Cells express a tagged version of ubiquitin (the tag is used to isolate ubiquitin-conjugated proteins); tryptic peptides contain diGly-conjugated uncleaved Lys
Methylation	14.01565	Anti-methyl-Lys or anti-methyl-Arg antibodies	59 sites[55]	*Homo sapiens* (HeLa S3 cell lines)	Currently available pan anti-methyl-Lys and anti-methyl-Arg antibodies typically have low specificity
o-GlcNac	203.07937	Lectin	141 sites[124]	*Homo sapiens* (HeLa cell lines)	Very labile PTM; may benefit from fragmentation methods such as ETD and ECD

ECD, electron capture dissociation; ETD, electron transfer dissociation; IMAC, immobilized metal affinity chromatography; MS, mass spectrometry; o-GlcNac, o-linked β-N-acetylglucosamine; PTM, post-translational modification. *The Δ masses given here are monoisotopic masses of the PTM. ‡Reference to the largest proteomics study for each PTM to date.

As modified peptides only constitute a minority of all peptides, it is usually necessary to enrich them for proteome-wide modification analysis[48]. Ideally, this enrichment captures all modified peptides of interest and no others. In practice, modified peptides have a certain 'enrichment factor' with respect to the starting peptide mixture, which can range from over 100-fold for phosphorylation to only several fold for methylated peptides. In phosphoproteomics, metal affinity complexation of the phosphogroup is the most common principle, but there are many other strategies[49]. PTM enrichment methods can also be more or less 'specific'. This term refers to the proportion of modified peptides in the enriched population and can range from close to 100% for some fractions in phosphopeptide purification to about 5% for Lys-acetylated peptides captured by antibodies[50]. In our experience, enrichment of PTM-bearing peptides by affinity purification is preferable to chemical derivatization of PTMs. Chemical modifications are invariably accompanied by side reactions that need to be distinguished from *in vivo* modifications (see for example REF. 24). TABLE 1 contains a list of the most frequently studied PTMs and how they are analysed by MS.

Sequence-specific identification of PTMs. In terms of data quality, there are two distinct tasks in PTM characterization: confident identification of the peptide sequence bearing the PTM and unambiguous localization of the PTM to the correct amino acid. For example, it may be possible to identify a phosphorylated peptide with high confidence (> 99%) but the data may not be sufficient to distinguish which of two adjacent Ser residues is phosphorylated. Therefore, proteomic PTM data sets should contain both a peptide identification score and a PTM localization score[51–53], which is unfortunately not always the case. It is important that the FDRs of PTM determination are rigorously determined, especially when PTM studies are meant to serve as resources for the community. Unambiguous localization of PTMs benefits from high mass accuracy in the fragmentation spectra, which increases the confidence of both peptide identification and PTM localization (BOX 1; FIG. 3).

Comprehensiveness of PTM analysis by MS. One of the surprises of systems-wide PTM analysis has been the large number of sites that are routinely discovered. Phosphoproteome studies have identified more than 20,000 phosphorylation sites in a single project[54] and there may be more than 100,000 sites in the phosphoproteome of human cells. The numbers for other modifications, such as acetylation, ubiquitylation, methylation and gylcosylation, are also large, and no modification has been mapped to completion yet. Quantitative analysis is therefore essential to focus on the subset of regulated sites that are likely to be functionally important in the biological process of interest. Determining the occupancy of the site of modification is also becoming possible for many sites[54] and this may further help in pinpointing the sites that probably have functional roles.

In large-scale phosphoproteomics experiments, coverage of the well-known functional sites in the pathway under investigation can validate the quality of the data. Although some phosphorylation sites are in sequence contexts that are difficult to probe by MS (for example, in very short tryptic peptides), in general this goal seems to be within reach of current technology.

Non-phospho PTMs in signalling. Although most MS-based studies so far have focused on analysis of phosphorylation-dependent events, MS can in principle examine any other PTM involved in cell signalling. For example, an antibody-based enrichment approach and the 'heavy methyl' SILAC variant, in which heavy labelled Met serves as a donor for the methyl group, provide a tool to investigate *in vivo* methylation of Lys and Arg[55]. Enrichment of peptides containing acetylated Lys residues revealed a surprisingly large number of acetylation sites on mitochondrial proteins[56]. Recently, high

Metal affinity complexation
The coordinated binding between immobilized metal ions and charged peptides. Immobilized metals such as Fe^{3+} or Ga^{3+}, or metal oxides such as TiO_2 or ZrO_2, are commonly used to enrich phosphorylated peptides from non-phosphorylated peptides.

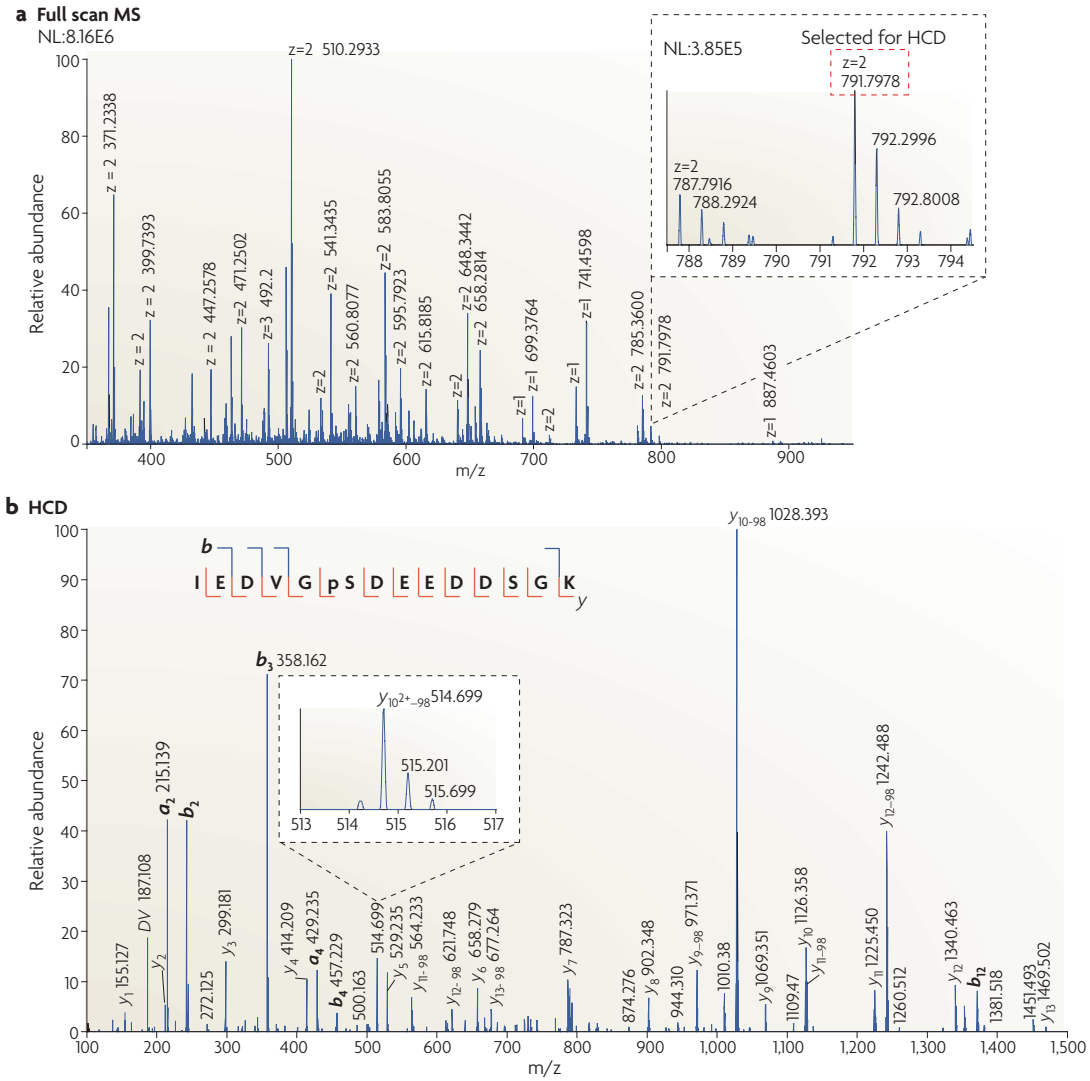

Figure 3 | **High resolution mass spectrum of a phosphopeptide. a** | One of several thousand mass spectrometry (MS) spectra from a fraction enriched for phosphopeptides. At least 20 peptides eluted simultaneously from the column. The charge state of the peptides are shown (z). A doubly charged phosphopeptide from heat shock protein 90β (HSP90β) at a mass to charge (m/z) ratio of 791.7978 is shown in the inset. This low intensity peptide ion was isolated by ejecting all other ions during accumulation in the linear ion trap. Despite the low intensity, it took only 91 milliseconds to accumulate a sufficient number of ions (50,000). This peptide was then selected for higher energy collisional dissociation (HCD). **b** | Following fragmentation by HCD, the high resolution tandem mass spectrum of the phosphopeptide was obtained in the orbitrap analyser of an LTQ Orbitrap Velos (Thermo Fisher Scientific). The presence of a complete series of fragments at the peptide bonds, each measured with mass deviations of a few parts per million, and the presence of fragments that include the phosphogroup, unambiguously identified the peptide sequence shown, as well as the phosphorylation site at Ser225. a_n and b_n ions are fragments at the n^{th} peptide bonds that contain the amino-terminal part of the peptide, whereas y_n ions contain the carboxy-terminal part. NL, normalized intensity level (counts per second).

resolution MS in combination with SILAC identified over 3,600 acetylation sites on 1,750 proteins[50]. Cells were treated with two different Lys deacetylase inhibitors (MS-275 and SAHA, which is used in the clinic), to investigate their *in vivo* cellular targets. This acetylome analysis revealed that Lys acetylation is a widespread modification that targets large macromolecular complexes involved in the regulation of diverse cellular processes. Lys acetylation also seems to have a previously unappreciated role in regulating metabolic pathways in human liver and *Salmonella enteric*[57,58].

MS is also increasingly used for measuring other complex PTMs such as ubiquitylation and sumoylation[59,60]. However, pinpointing the modified residue and differentiating similar PTMs is still a demanding task. For example, it is difficult to distinguish ubiquitin from other ubiquitin-like molecules such as neural precursor cell-expressed developmentally downregulated protein 8 (NEDD8) and interferon-induced 17 kDa protein (ISG15) as all leave the same diGly tag at the substrate site after tryptic digestion of the modified peptide. Likewise, sumoylated peptides contain a large

Table 2 | **MS-based global phosphoproteomics screens**

Number of phosphorylation sites identified*	Signalling processes studied‡	Quantification method	Organism
20,443 (Ser, Thr and Tyr)	Cell cycle[54]	SILAC	*Homo sapiens* (HeLa cell line)
14,000 (Ser, Thr and Tyr)	Mitosis[80]	SILAC	*Homo sapiens* (HeLa cell line)
12,186 (Ser, Thr and Tyr)	FLT3 signalling[94]	SILAC	*Mus musculus* (32D cell line)
10,665 (Ser, Thr and Tyr)	TCR signalling[72]	SILAC	*Homo sapiens* (Jurkat T cell lines)
10,000 (Ser, Thr and Tyr)	Identification of PTP1B substrates[101]	SILAC	*Drosophila melanogaster* (Schneider cells)
8,710 (Ser, Thr and Tyr)	Cdk1 substrates[103]	SILAC	*Saccharomyces cerevisiae*
6,600 (Ser, Thr and Tyr)	EGF signalling[51]	SILAC	*Homo sapiens* (HeLa cell line)
6,451 (Ser, Thr and Tyr)	EGF signalling, with or without MEK1, MEK2 or p38 kinase inhibitors[98]	SILAC	*Homo sapiens* (HeLa cell line)
5,534 (Ser, Thr and Tyr)	Osmotic stress[125]	SILAC	*Saccharomyces cerevisiae*
5,433 (Ser, Thr and Tyr)	Phosphatase inhibitors[126]	SILAC	*Mus musculus* (Hepa1-6 cell line)
4,551 (Tyr)	Lung cancer[90]	No quantification	*Homo sapiens* (41 different NSCLC cell lines, and 150 NSCLC tumours)
3,067 (Ser, Thr and Tyr)	Stem cell differentiation[77]	SILAC	*Homo sapiens* (embryonic stem cell line HUES-7)
2,689 (Ser, Thr and Tyr)	Mec1, Tel1 and Rad53 substrates[102]	N-isotag	*Saccharomyces cerevisiae*
2,546 (Ser, Thr and Tyr)	Stem cell differentiation[78]	Label free	*Homo sapiens* (embryonic stem cell line H1)
1,940 (Ser, Thr and Tyr)	Mitotic spindles[85]	SILAC	*Homo sapiens* (HeLa S3 cell line)
1,000 (Ser, Thr and Tyr)	Cell cycle[84]	SILAC	*Homo sapiens* (HeLa S3 cell line)
900 (Ser/Thr and Glu)	ATM and ATR substrates[82]	SILAC	*Homo sapiens* (293T cell line)

ATM, ataxia telangiectasia mutated; ATR, ataxia telangiectasia and RAD3-related; EGF, epidermal growth factor; FLT3, FMS-like Tyr kinase 3; MEK, MAPK ERK kinase; NSCLC, non-small cell lung cancer cell lines; PTP1B, protein Tyr phosphatase 1B; SILAC, stable isotope labelling by amino acids in cell culture; TCR, T cell receptor.
*Phosphorylated amino acids are shown in brackets. ‡The list includes recent large-scale, mostly quantitative, phosphoproteomics studies and is not comprehensive.

modification after digestion (32 amino acids for human small ubiquitin-related modifier 2 (SUMO2) and SUMO3), leading to large cross-linked peptides. These peptides result in complex fragmentation spectra and require specialized MS techniques for efficient detection[61,62]. Complex glycosylation structures are difficult to characterize in detail but sites of N-glycosylation can now be mapped on a large scale and at high precision.

Phospho-analysis of signalling networks

MS has been used as a tool in biological research for decades[63] and has helped to identify important signalling proteins, for example, caspase 8, a key mediator of apoptosis[64]. Likewise, there is a long history of methods development in proteomics in general, and in phosphoproteomics in particular[49,65–67]. However, the technology for large-scale, quantitative and high accuracy proteomics has only been developed during the past few years. Here, we review papers that apply global and quantitative MS-based proteomics to signalling related questions (TABLE 2).

Investigating cellular processes by proteomics. PhosphoTyr signalling initiates signalling cascades, and Tyr phosphorylated proteins or peptides can be affinity enriched by specific antibodies before MS. Several studies monitored the dynamics of Tyr phosphorylation in response to insulin stimulation, providing an overview of receptor proximal signalling events and yielding many new insulin-regulated substrates[68–70]. An unbiased phosphoproteomics study on the pheromone pathway in yeast has quantified more than 700 phosphorylated peptides[71]. A quantitative and time-resolved study in mammalian cells analysed 6,600 unique Ser, Thr or Tyr phosphorylation sites for their dynamic regulation in response to epidermal growth factor (EGF) stimulation[51]. This study provided the first dynamic view of a global signalling network in mammalian cells and revealed the unparalleled sensitivity and capacity of MS to map cell signalling networks at a systems-wide level. A similar approach for T cell receptor (TCR) signalling uncovered a role for Ser and Thr phosphorylation in the formation of the F-actin cup[72]. In multicellular organisms, cells not only receive unidirectional signals from extracellular soluble cytokines and growth factors, but also exchange signals with neighbouring cells in tissues to control their behaviour. For example, in bidirectional signalling in ephrin type B receptor 2 (EPHB2)- and ephrin B1-expressing cells, membrane-bound ephrin ligands interact with membrane-bound EPHB2 RTKs, resulting in signalling in both the ligand- and receptor-expressing cells. To study this complex interplay of signals, researchers differentially labelled the two cell lines by SILAC, making signalling events in each of them distinguishable by MS. Ten minutes after mixing, EPHB2- and ephrin B1-expressing cells were lysed and phosphoTyr-containing peptides enriched and quantified[73]. Tyr phosphorylation sites on more

F-actin cup
(Filamentous actin cup). A polymer of globular actin (G-actin) subunits, which accumulates underneath phagocytic cups — cup-shaped extensions of the plasma membrane that encircle foreign particles during early processes of phagocytosis.

than 300 proteins were modulated significantly as a result of bringing the two cell populations into contact (although many other sites had 'inconclusive quantification', probably owing to the use of SILAC labels with overlapping isotope patterns), revealing asymmetric and distinct signalling networks engaged in cell positioning.

Exposure of cells to specific growth factors provides crucial signals during cell fate decisions. For example, EGF and platelet-derived growth factor (PDGF) induce mesenchymal stem cells to differentiate into bone forming cells or to proliferate, respectively. The reason for the different response of this system to closely related signals was uncovered by quantifying Tyr phosphorylated proteins. This revealed that PDGF, but not EGF, activated the phosphoinositide 3-kinase (PI3K) pathway in addition to the canonical mitogen-activated protein kinase (MAPK) pathway. Chemically blocking the PI3K pathway converted the PDGF pathway into a differentiation signal[74]. To date, most MS-based quantitative signalling studies have been carried out in immortalized cell lines, but other cell types, such as embryonic stem cells, are also amenable to SILAC labelling and MS analysis[75,76]. Two recent studies analysed phosphorylation changes associated with the differentiation of human embryonic stem cells. A SILAC-based study identified phosphorylation at 3,067 sites in response to differentiation. Surprisingly, nearly half of all quantified sites seemed to be modulated within one hour of differentiation induction[77]. A label-free MS approach uncovered 2,546 phosphorylation sites and suggested that activation of multiple RTKs cooperatively supports the growth of undifferentiated embryonic stem cells[78]. Using two different peptide fragmentation approaches (CID and ETD), another recent embryonic stem cell study reported more than 10,000 phosphorylation events[79].

Cellular processes such as the cell cycle and perturbations such as DNA damage have overlapping yet distinct phosphorylation signatures compared to signalling networks activated by extracellular growth factors. Phosphorylation changes between mitotically arrested and interphase cells[80,81], or across the entire cell cycle[54], have recently been reported. To distinguish changes in phosphorylation from changes in protein abundance, both the proteome and phosphoproteome should be measured. If sufficient depth of analysis is achieved, such that the abundance change of the non-phosphorylated and phosphorylated peptide, as well as that of the protein, are measured, then the occupancy of phosphorylation sites can be calculated. This has revealed that thousands of phosphorylation sites can have close to complete occupancy, at least during mitosis[54].

The depth of phosphorylation site coverage on a specific protein subclass or in a specific subcellular compartment can be increased by using upstream enrichment or fractionation methods, albeit at the expense of systems-wide analysis. For example, to affinity enrich ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and RAD3-related (ATR) substrates in the DNA damage response pathway, phosphospecific-antibodies have been used that are directed against the Ser/Thr-Gln motif, which ATM and ATR phosphorylates[82,83]. Phosphorylation of protein kinases on key residues is often used as proxy to analyse their activation state in the cell. To obtain an in-depth coverage of kinase phosphorylation, a chemical proteomics approach using immobilized ATP mimetics was used to enrich protein kinases, which allowed SILAC quantification of their phosphorylation sites in mitosis[84]. Similarly, phosphorylation sites on mitotic spindle-associated proteins were analysed after purification of mitotic spindles at different stages of mitosis[85].

Large-scale mapping of oncogenic signalling. Many human malignancies are caused by inappropriate activation of signalling networks[86]. The mutational analysis of kinases and other drivers of transformation provides important clues about their involvement in human diseases. However, a key attraction of MS is that it can quantify transformation-associated changes in the entire signalling network. MS-based approaches have identified Tyr phosphorylation signalling pathways that are activated by oncogenes such as SRC, EGF receptor (EGFR) and KRAS[87–89]. Large-scale phosphoTyr profiling of non-small cell lung cancer (NSCLC) cell lines and tumours uncovered over 4,000 Tyr residues that seemed to be phosphorylated in this system, many of them presumably owing to inappropriate activation of kinases in NSCLC[90]. However, this pioneering study used low-resolution MS, making it difficult to estimate FDRs in this extensive and heterogeneous data set. MS-based phosphoproteomic analysis can readily be extended to study Tyr phosphorylation[90] or Ser, Thr and Tyr phosphorylation[91] in primary tumour samples.

Because of the negatively charged phosphate group, phosphopeptides may ionize more readily if electrospray is carried out in a negative rather than the commonly used positive ion mode. Looking back to precursor-ion studies using triple-quadrupole instruments, a recent study of BRAF signalling in cancer cell lines detected the phosphogroup in negative ion mode followed by peptide sequencing in the positive ion mode[92]. This work did not use affinity enrichment and therefore the number of identified sites is relatively low. However, using a MAPK extracellular signal-regulated kinase (ERK) kinase 1 (MEK1; also known as MAPKK1) and MEK2 (also known as MAPKK2) inhibitor, the authors identified FAM129B (also known as MINERVA), a previously uncharacterized protein, as an effector of BRAF–MEK–ERK MAPK signalling in melanoma, and established its functional role in the migration of melanoma cells.

The cellular compartmentalization influences the activation of signalling pathways and it is well known that endocytosed RTKs can signal along the endocytic route[93]. A recent phosphoproteomics study combining quantitative MS with chemical and genetic tools discovered that oncogenic RTKs can activate non-conventional signalling along the biosynthetic route. The RTK FMS-like Tyr kinase 3 with a juxtamembrane internal tandem duplication (FLT3-ITD), was allowed to proceed to the cell surface

Endocytic route
The trafficking of cell surface receptors from the plasma membrane to intracellular compartments by receptor endocytosis, which generally involves early and late endosomes, multivesicular bodies and lysosomes. It is a major pathway that regulates the amplitude and duration of receptor signalling at the plasma membrane.

Biosynthetic route
The trafficking of receptors from their intracellular biosynthesis compartments to the surface, involving protein synthesis at the ER, relocation to the Golgi and, finally, transport to the cell surface.

FLT3-ITD
The FLT3 receptor containing an in-frame insertion of amino acid sequence (internal tandem duplication (ITD); of ~3–100 amino acids in length) in the intracellular juxtamembrane region, which results in ligand-independent receptor activation. These oncogenic mutations are found in about 20% of human acute myeloid leukaemia and are associated with oncogenic transformation.

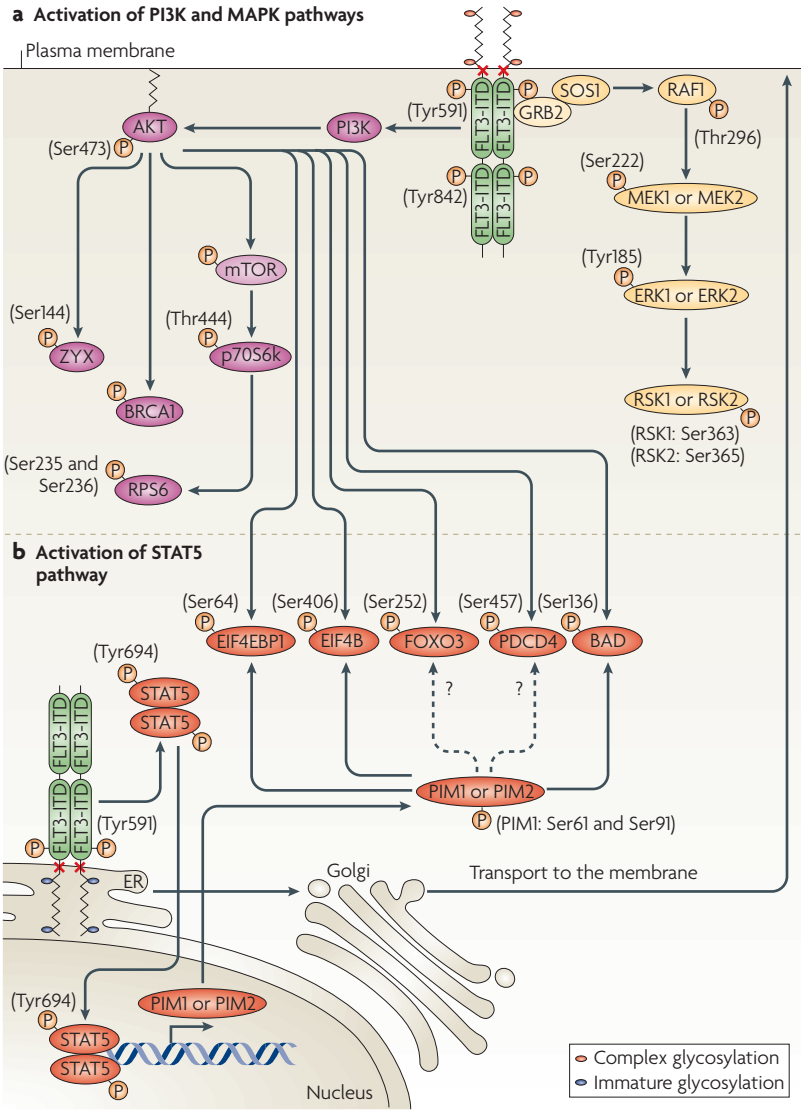

Figure 4 | **Quantitative proteomic analysis of oncogenic RTK signalling compartmentalization. a** | In normal cells, receptor tyrosine kinases (RTKs) are activated at the plasma membrane by ligand stimulation. However, the oncogenic FMS-like tyrosine kinase 3 with internal tandem duplication (FLT3-ITD) receptor is activated independently of ligands at the cell surface and at compartments of the biosynthetic route, such as the endoplasmic reticulum (ER) and the Golgi. Choudhary *et al.*[94] showed by classical techniques and by mass spectrometry (MS)-based proteomics that, at the plasma membrane, FLT3-ITD activates the classical extracellular signal-regulated kinase (ERK) mitogen-activated protein kinase (MAPK) and phosphoinositide-3 kinase (PI3K) signalling pathways and their downstream effectors. **b** | On the ER and Golgi, FLT3-ITD aberrantly activates signal transducer and activator of transcription 5 (STAT5) but not PI3K or MAPK pathways. Phosphorylated STAT5 forms dimers, migrates to the nucleus and upregulates its downstream genes, such as *PIM1* and *PIM2*. Pim kinases then phosphorylate several potential downstream substrates such as BCL2 antagonist of cell death (BAD), eukaryotic translation initiation factor 4B (EIF4B), EIF4E-binding protein 1 (EIF4EBP1), programmed cell death protein 4 (PDCD4) and forkhead box protein O3 (FOXO3). Phosphorylation of these substrates by Pim (and Akt) kinases prevents apoptosis, promotes cell survival and induces cellular transformation. The indicated phosphorylation sites on the signalling components were quantified by stable isotope labelling by amino acids in cell culture (SILAC)-based quantitative proteomics. BCRA1, breast cancer type 1 susceptibility protein; GRB2, growth factor receptor-bound protein 2; MEK, MAPK ERK kinase; mTOR, mammalian target of rapamycin; p70S6K, p70 ribosomal S6 kinase-α; RPS6, ribosomal protein S6; RSK, ribosomal protein S6 kinase-α; SOS1, Son of Sevenless homologue 1; ZYX, zyxin. Figure is modified, with permission, from REF. 94 © (2009) Elsevier.

or kept in the endoplasmic reticulum (ER). Comparison of the signalling patterns by classical tools and by SILAC revealed that specific pathways are initiated at the ER as opposed to the plasma membrane[94]. In particular, intracellular activation of FLT3-ITD aberrantly triggers transformation-associated signal transducer and activator of transcription 5 (STAT5)–Pim kinase signalling pathways, whereas on reaching the cell surface, the same receptor activates a different spectrum of substrates corresponding to its wild-type function (FIG. 4).

Mapping the effects of network perturbations. Global and unbiased measurements of the effects of network perturbations have so far typically been at the level of the transcriptome. MS can directly quantify these effects at the level of early information processing (phosphorylation levels) and at the end of the signal cascade (protein abundance). This is particularly important for understanding the effects and degree of specificity of kinase inhibitors intended as drugs. Quantitative MS combined with global phosphoproteomics is an attractive method for identifying targets of such inhibitors and for assessing their specificity[95,96], and goes far beyond current assays on recombinant proteins. Specialized proteomics techniques can measure the binding affinity of chemical inhibitors to all expressed kinases in a cell type, which is a good indicator of their inhibitory potency, and can also profile their specificity *in vivo*[95–97]. Furthermore, unbiased analysis of phosphorylation changes of cells treated with kinase inhibitors provides a systems view of signalling networks affected by inhibitor treatment[98,99].

Alternatively, signalling networks of protein kinases and phosphatases can be unravelled either by using specific genetic perturbation (such as gene knockout or knockdown) or by modified enzymes that can be selectively blocked with chemical inhibitors. Substrates of protein Tyr phosphatase 1B (PTP1B; also known as PTPN1) were analysed in mouse and *Drosophila melanogaster* cell lines in which PTP1B was either deleted genetically or depleted with a small interfering RNA approach[100,101]. PTP1B-knockout mouse fibroblasts showed increased phosphorylation of 18 Tyr sites, 8 of which directly bound to the substrate-trapping mutant of PTP1B. In *D. melanogaster* S2 cells, 28 Tyr sites showed such changes, but ~250 Ser or Thr sites were also upregulated, suggesting that there is crosstalk between PTP1B substrates and Ser/Thr signalling. Similarly, phosphoproteomics of *Saccharomyces cerevisiae* mutant strains deficient for Mec1, Tel1 and Rad53 (orthologues of human ATR, ATM and CHK2, respectively) has provided important clues to the substrates of these kinases, which are components of DNA damage checkpoint signalling[102]. As an alternative approach to gene deletion, many new substrates of yeast Cdk1 (also known as Cdc28), an essential cell cycle kinase, were revealed by phosphoproteomic analysis of a yeast strain expressing an analogue-sensitive mutant of Cdk1 (REF. 103); that is, a Cdk1 protein that has been engineered to accept analogues of ATP that are not efficiently used by wild-type kinases.

Network analysis of phosphoproteomics data. High quality phosphorylation data sets containing thousands of quantified phosphorylation sites are an attractive resource for systems-wide *in silico* analysis of signalling networks. Tyr phosphorylation analysis downstream of human EGFR2 (HER2; also known as ERBB2) provided data for network modelling to identify crucial phosphorylation sites that are key control points for cell migration and proliferation[104]. Integration of information about tissue- or cell type-specific kinase expression, cellular localization, substrate phosphorylation motifs and other functional interaction data with global phosphorylation data is useful in predicting kinase–substrate relationships in signalling networks[105]. Connecting linear sequence kinase motifs and phosphorylation-dependent binding domains with *in vivo* phosphorylation data further helps in matching phosphorylation sites to their putative upstream kinases[106], and computational analysis of rich phosphoproteomics data can examine the molecular evolution of signalling networks[107]. The above studies are interesting examples of how phosphoproteomics data can be integrated with other biochemical data. However, we believe that this field is still in its infancy because of the current scarcity of high quality and large-scale information about kinase sequence motifs, redundancy in these motifs and methods to experimentally validate thousands of predicted kinase–substrate relationships.

Functions of PTMs in inducing protein–protein interactions. Almost all biological functions are mediated by protein–protein interactions and their large-scale study is an important area in functional proteomics[108,109]. Recent examples in the signalling field include the determination of several novel components of the PP2A phosphatase complex[110,111] and of the dynamic interactome of ERK MAPKs during the neuronal differentiation of PC12 cells[112]. However, in addition to finding binding partners for full length proteins, MS-based proteomics is uniquely suited to explore PTM-mediated interactions in signalling networks. The power of quantitative proteomics to discover PTM-induced protein interactions was first shown in the following example, inspired by the classical experiments in growth factor-induced assembly of signalling complexes. The immobilized SH2 domain of the adaptor protein growth factor receptor-bound protein 2 (GRB2) specifically precipitated the protein complex containing the activated (Tyr phosphorylated) EGFR. SILAC-based quantification of EGF-stimulated versus non-stimulated cells then distinguished specific protein interactors — those binding owing to EGF-mediated EGFR activation — from hundreds of background proteins that bound to the bait or beads[113]. In a reverse approach, synthetic peptides containing phosphorylated or non-phosphorylated Tyr were incubated with SILAC-labelled cell lysates to characterize the phosphoTyr-dependent interactome of the ERBB receptor kinase family and of the insulin receptor family and its downstream effectors insulin receptor substrate 1 (IRS1) and IRS2 (REFS 114,115). The same approach also characterized mammalian interactors of bacterial proteins that are phosphorylated when injected into the host cell to subvert its functions[116].

Integration and perspectives

The exciting studies discussed above clearly show that MS-based quantitative proteomics is already delivering on its promise to map signalling pathways in depth and in a systems-wide manner. At a minimum, large-scale and high accuracy studies are an extremely useful resource for the community as publicly stored data from such studies enables research groups without access to MS technology to functionally study PTMs on the proteins of their interest. We envision that many projects will start with a high resolution and quantitative proteomics screen, which provides the basis for functional hypotheses. These hypotheses are then followed up either with standard functional assays or in combination with quantitative proteomics methods. In the long term, the unmatched specificity, quantitative accuracy, unbiased nature and increasing sensitivity of MS-based methods will make them a routine part of all phases of signalling research.

What are the next steps for the field? In our view it is first important that the technology becomes more widely accessible and that the depth of coverage and data quality demonstrated in some pioneering publications become much more wide-spread. In the next few years, instrumental developments, improved protocols and computational tools will all work together to make this vision a reality. Integration of these proteomics tools with high throughput genetic and imaging technologies[117] and phenotypic screens will provide powerful tools to the cell biology community. Ongoing and future improvement in technology will also be important in addressing the challenges of analysing a large number of samples, including *in vivo* animal and human tumour materials, and the analysis of very small numbers of cells[118].

Currently, phosphorylation is by far the most extensively investigated PTM, hence our focus on phosphoproteomics in this Review. However, ubiquitylation, acetylation and some other PTMs that can be efficiently enriched are also becoming amenable to large-scale analysis. These developments should make it possible to investigate multiple PTMs in the same system, yielding direct data on their crosstalk at the global level.

Given high precision and in-depth data, the challenge now is to assign function to as many of the signalling nodes as possible. Quantification is indispensible for this, and clever experimental design helps to reduce the number of sites to be experimentally investigated to a manageable number. Furthermore, these large and quantitative data sets are already providing input for sophisticated bioinformatic algorithms that attempt to extract information on kinase–substrate relationships and the evolution of signalling networks, and ultimately to determine the crucial control points in cellular decision making.

Today, many large-scale studies infer novel biological mechanisms but do not prove them. We hope that increased availability and quality of proteomics hardware, protocols, strategies and analysis tools will allow researchers to increasingly focus on these aspects, which are ultimately the important goals of biological research.

1. Seet, B. T., Dikic, I., Zhou, M. M. & Pawson, T. Reading protein modifications with interaction domains. *Nature Rev. Mol. Cell Biol.* **7**, 473–483 (2006).
2. Hunter, T. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* **28**, 730–738 (2007).
3. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protoc.* **1**, 2856–2860 (2006).
4. Link, A. J. *et al.* Direct analysis of protein complexes using mass spectrometry. *Nature Biotech.* **17**, 676–682 (1999).
5. Washburn, M. P., Wolters, D. & Yates, J. R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotech.* **19**, 242–247 (2001).
6. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212–217 (2006).
7. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b-range mass accuracies and proteome-wide protein quantification. *Nature Biotech.* **26**, 1367–1372 (2008).
8. Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nature Rev. Mol. Cell Biol.* **5**, 699–711 (2004).
9. Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F. & Coon, J. J. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal. Chem.* **76**, 3590–3598 (2004).
10. Olsen, J. V. *et al.* Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* **4**, 709–712 (2007).
11. Olsen, J. V. *et al.* A dual pressure linear ion trap orbitrap instrument with very high sequencing speed. *Mol. Cell. Proteomics* **8**, 2759–2769 (2009).
12. Zubarev, R. A. *et al.* Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* **72**, 563–573 (2000).
13. Syka, J. E. *et al.* Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.* **3**, 621–626 (2004).
14. Ong, S. E. & Mann, M. Mass spectrometry-based proteomics turns quantitative. *Nature Chem. Biol.* **1**, 252–262 (2005).
15. Ong, S. E. *et al.* 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 (2002).
16. Mann, M. Functional and quantitative proteomics using SILAC. *Nature Rev. Mol. Cell Biol.* **7**, 952–958 (2006).
17. Kruger, M. *et al.* SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* **134**, 353–364 (2008).
18. Geiger, T., Cox, J. Pawel, O., Wisniewski, J. R. & Mann, M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature Methods* **7**, 383–385 (2010).
19. Ross, P. L. *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154–1169 (2004).
20. Ow, S. Y. *et al.* iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *J. Proteome Res.* **8**, 5347–5355 (2009).
21. Zhang, Y. *et al.* A robust error model for iTRAQ quantification reveals divergent signaling between oncogenic FLT3 mutants in acute myeloid leukemia. *Mol. Cell. Proteomics* 17 Dec 2009 (doi: 10.1074/mcp.M900452-MCP200).
22. Hsu, J. L., Huang, S. Y., Chow, N. H. & Chen, S. H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **75**, 6843–6852 (2003).
23. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protoc.* **4**, 484–494 (2009).
24. Nielsen, M. L. *et al.* Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nature Methods* **5**, 459–460 (2008).
25. Mueller, L. N., Brusniak, M. Y., Mani, D. R. & Aebersold, R. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J. Proteome Res.* **7**, 51–61 (2008).
26. Trost, M. *et al.* The phagosomal proteome in interferon-γ-activated macrophages. *Immunity* **30**, 143–154 (2009).
27. Wong, J. W. & Cagney, G. An overview of label-free quantitation methods in proteomics by mass spectrometry. *Methods Mol. Biol.* **604**, 273–283.
28. Luber, C. *et al.* Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279–289 (2010).
29. de Godoy, L. M. *et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251–1254 (2008).
Haploid yeast were quantified against diploid yeast after SILAC encoding, and the differential roles of members of the pheromone pathway were revealed by this expression analysis. This is the first complete proteome analysis, as judged by comparison to previous genome-wide tagging experiments.
30. Malmstrom, J. *et al.* Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* **460**, 762–765 (2009).
31. Rivers, J., Simpson, D. M., Robertson, D. H., Gaskell, S. J. & Beynon, R. J. Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol. Cell. Proteomics* **6**, 1416–1427 (2007).
32. Brunner, E. *et al.* A high-quality catalog of the *Drosophila melanogaster* proteome. *Nature Biotech.* **25**, 576–583 (2007).
33. Hanke, S., Besir, H., Oesterhelt, D. & Mann, M. Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *J. Proteome Res.* **7**, 1118–1130 (2008).
34. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
35. Huh, W. K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
36. Wisniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359–362 (2009).
37. Schmidt, A., Claassen, M. & Aebersold, R. Directed mass spectrometry: towards hypothesis-driven proteomics. *Curr. Opin. Chem. Biol.* **13**, 510–517 (2009).
38. Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D. A. & White, F. M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc. Natl Acad. Sci. USA* **104**, 5860–5865 (2007).
39. Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L. & Park, B. K. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **877**, 1229–1239 (2009).
40. Unwin, R. D., Griffiths, J. R. & Whetton, A. D. A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). *Nature Protoc.* **4**, 870–877 (2009).
41. Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B. & Aebersold, R. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* **138**, 795–806 (2009).
The targeted analysis of ~100 proteins from different yeast abundance classes, using multiple reaction monitoring (MRM), detected proteins with as few as an estimated 41 copies per cell. Expression changes in the glycolytic pathway on nutrient switch were quantified.
42. Kumar, C. & Mann, M. Bioinformatics analysis of mass spectrometry-based proteomics data sets. *FEBS Lett.* **583**, 1703–1712 (2009).
43. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol. Biol.* **604**, 55–71.
44. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genet.* **25**, 25–29 (2000).
45. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, D277–D280 (2004).
46. Jensen, O. N. Interpreting the protein language using proteomics. *Nature Rev. Mol. Cell Biol.* **7**, 391–403 (2006).
47. Witze, E. S., Old, W. M., Resing, K. A. & Ahn, N. G. Mapping protein post-translational modifications with mass spectrometry. *Nature Methods* **4**, 798–806 (2007).
48. Zhao, Y. & Jensen, O. N. Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* **9**, 4632–4641 (2009).
49. Macek, B., Mann, M. & Olsen, J. V. Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu. Rev. Pharmacol. Toxicol.* **49**, 199–221 (2009).
50. Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834–840 (2009).
51. Olsen, J. V. *et al.* Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635–648 (2006).
Time resolved, quantitative analysis of 6,600 phosphorylation sites provided the first dynamic and global view of the effects of growth factor stimulation on a cell.
52. Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J. & Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature Biotech.* **24**, 1285–1292 (2006).
53. Gnad, F. *et al.* PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol.* **8**, R250 (2007).
54. Olsen, J. V. *et al.* Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* **3**, ra3 (2010).
55. Ong, S. E., Mittler, G. & Mann, M. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. *Nature Methods* **1**, 119–126 (2004).
56. Kim, S. C. *et al.* Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* **23**, 607–618 (2006).
57. Wang, Q. *et al.* Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**, 1004–1007.
58. Zhao, S. *et al.* Regulation of cellular metabolism by protein lysine acetylation. *Science* **327**, 1000–1004.
Together with reference 50, references 56–58 identify thousands of new Lys acetylation sites and functionally implicate them in the cell cycle, 14-3-3-based signalling and central metabolic pathways.
59. Peng, J. *et al.* A proteomics approach to understanding protein ubiquitination. *Nature Biotech.* **21**, 921–926 (2003).
60. Golebiowski, F. *et al.* System-wide changes to SUMO modifications in response to heat shock. *Sci. Signal.* **2**, ra24 (2009).
61. Pedrioli, P. G. *et al.* Automated identification of SUMOylation sites using mass spectrometry and SUMmOn pattern recognition software. *Nature Methods* **3**, 533–539 (2006).
62. Matic, I. *et al. In vivo* identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an *in vitro* to *in vivo* strategy. *Mol. Cell. Proteomics* **7**, 132–144 (2008).
63. Burlingame, A. L., Baillie, T. A. & Russell, D. H. Mass spectrometry. *Anal. Chem.* **64**, 467R–502R (1992).
64. Muzio, M. *et al.* FLICE, a novel FADD-homologous ICE/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death-inducing signaling complex. *Cell* **85**, 817–827 (1996).
65. Ficarro, S. B. *et al.* Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nature Biotech.* **20**, 301–305 (2002).
The first large-scale MS-based phosphoproteome analysis, reporting the identification of several hundred yeast phosphorylation sites.
66. Nita-Lazar, A., Saito-Benz, H. & White, F. M. Quantitative phosphoproteomics by mass spectrometry: past, present, and future. *Proteomics* **8**, 4433–4443 (2008).
67. Dengjel, J., Kratchmarova, I. & Blagoev, B. Receptor tyrosine kinase signaling: a view from quantitative proteomics. *Mol. Biosyst.* **5**, 1112–1121 (2009).
68. Blagoev, B., Ong, S. E., Kratchmarova, I. & Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nature Biotech.* **22**, 1139–1145 (2004).
69. Schmelzle, K., Kane, S., Gridley, S., Lienhard, G. E. & White, F. M. Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes* **55**, 2171–2179 (2006).
70. Kruger, M. *et al.* Dissection of the insulin signaling pathway via quantitative phosphoproteomics. *Proc. Natl Acad. Sci. USA* **105**, 2451–2456 (2008).

71. Gruhler, A. *et al.* Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **4**, 310–327 (2005).
72. Mayya, V. *et al.* Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. *Sci. Signal.* **2**, ra46 (2009).
73. Jorgensen, C. *et al.* Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science* **326**, 1502–1509 (2009).
74. Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M. & Mann, M. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* **308**, 1472–1477 (2005).
Differential quantification of Tyr phosphorylation in response to EGF or PDGF, which elucidated the mechanism responsible for the proliferation or differentiation of mesenchymal stem cells.
75. Graumann, J. *et al.* Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell. Proteomics* **7**, 672–683 (2008).
76. Prokhorova, T. A. *et al.* Stable isotope labeling by amino acids in cell culture (SILAC) and quantitative comparison of the membrane proteomes of self-renewing and differentiating human embryonic stem cells. *Mol. Cell. Proteomics* **8**, 959–970 (2009).
77. Van Hoof, D. *et al.* Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell* **5**, 214–226 (2009).
78. Brill, L. M. *et al.* Phosphoproteomic analysis of human embryonic stem cells. *Cell Stem Cell* **5**, 204–213 (2009).
79. Swaney, D. L., Wenger, C. D., Thomson, J. A. & Coon, J. J. Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. *Proc. Natl Acad. Sci. USA* **106**, 995–1000 (2009).
80. Dephoure, N. *et al.* A quantitative atlas of mitotic phosphorylation. *Proc. Natl Acad. Sci. USA* **105**, 10762–10767 (2008).
81. Steen, J. A. *et al.* Different phosphorylation states of the anaphase promoting complex in response to antimitotic drugs: a quantitative proteomic analysis. *Proc. Natl Acad. Sci. USA* **105**, 6069–6074 (2008).
82. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).
Used a panel of antibodies recognizing phosphorylation sites of the ATM and ATR consensus motif (Ser/Thr-Gln) to enrich substrates of these kinases and their regulation on DNA damage.
83. Stokes, M. P. *et al.* Profiling of UV-induced ATM/ATR signaling pathways. *Proc. Natl Acad. Sci. USA* **104**, 19855–19860 (2007).
84. Daub, H. *et al.* Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol. Cell* **31**, 438–448 (2008).
85. Malik, R. *et al.* Quantitative analysis of the human spindle phosphoproteome at distinct mitotic stages. *J. Proteome Res.* **8**, 4553–4563 (2009).
86. Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411**, 355–365 (2001).
87. Qiao, Y., Molina, H., Pandey, A., Zhang, J. & Cole, P. A. Chemical rescue of a mutant enzyme in living cells. *Science* **311**, 1293–1297 (2006).
88. Guo, A. *et al.* Signaling networks assembled by oncogenic EGFR and c-Met. *Proc. Natl Acad. Sci. USA* **105**, 692–697 (2008).
89. Amanchy, R. *et al.* Identification of c-Src tyrosine kinase substrates using mass spectrometry and peptide microarrays. *J. Proteome Res.* **7**, 3900–3910 (2008).
90. Rikova, K. *et al.* Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* **131**, 1190–1203 (2007).
91. Zanivan, S. *et al.* Solid tumor proteome and phosphoproteome analysis by high resolution mass spectrometry. *J. Proteome Res.* **7**, 5314–5326 (2008).
92. Old, W. M. *et al.* Functional proteomics identifies targets of phosphorylation by B-Raf signaling in melanoma. *Mol. Cell* **34**, 115–131 (2009).
93. Sorkin, A. & von Zastrow, M. Endocytosis and signalling: intertwining molecular networks. *Nature Rev. Mol. Cell Biol.* **10**, 609–622 (2009).
94. Choudhary, C. *et al.* Mislocalized activation of oncogenic RTKs switches downstream signaling outcomes. *Mol. Cell* **36**, 326–339 (2009).
Used chemical inhibitors, genetic manipulation and SILAC-based MS to analyse the role of cellular compartmentalization in oncogenic signalling on a global scale. The oncogenic FLT3 receptor is found to signal inappropriately during the biosynthetic route before reaching the plasma membrane.
95. Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nature Biotech.* **25**, 1035–1044 (2007).
96. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nature Biotech.* **26**, 127–132 (2008).
97. Sharma, K. *et al.* Proteomics strategy for quantitative protein interaction profiling in cell extracts. *Nature Methods* **6**, 741–744 (2009).
References 95–97 developed and applied chemical proteomics tools to identify *in vivo* targets of kinase inhibitors, to profile their kinase specificity and binding affinity.
98. Pan, C., Olsen, J. V., Daub, H. & Mann, M. Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics. *Mol. Cell. Proteomics* **8**, 2796–2808 (2009).
99. Huber, A. *et al.* Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. *Genes Dev.* **23**, 1929–1943 (2009).
100. Mertins, P. *et al.* Investigation of protein-tyrosine phosphatase 1B function by quantitative proteomics. *Mol. Cell. Proteomics* **7**, 1763–77 (2008).
101. Hilger, M., Bonaldi, T., Gnad, F. & Mann, M. Systems-wide analysis of a phosphatase knock-down by quantitative proteomics and phosphoproteomics. *Mol. Cell. Proteomics* **8**, 1908–1920 (2009).
102. Smolka, M. B., Albuquerque, C. P., Chen, S. H. & Zhou, H. Proteome-wide identification of *in vivo* targets of DNA damage checkpoint kinases. *Proc. Natl Acad. Sci. USA* **104**, 10364–10369 (2007).
103. Holt, L. J. *et al.* Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325**, 1682–1686 (2009).
Used an analogue sensitive mutant of Cdk1, a mutant-specific inhibitor and quantitative MS to reveal new substrates of this kinase.
104. Kumar, N., Wolf-Yadlin, A., White, F. M. & Lauffenburger, D. A. Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput. Biol.* **3**, e4 (2007).
105. Linding, R. *et al.* Systematic discovery of *in vivo* phosphorylation networks. *Cell* **129**, 1415–1426 (2007).
Developed a novel software algorithm, NetworKIN, to predict likely kinase–substrate relationships from large-scale phosphoproteomic data sets.
106. Miller, M. L. *et al.* Linear motif atlas for phosphorylation-dependent signaling. *Sci. Signal.* **1**, ra2 (2008).
107. Tan, C. S. *et al.* Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci. Signal.* **2**, ra39 (2009).
108. Gingras, A. C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nature Rev. Mol. Cell Biol.* **8**, 645–654 (2007).
109. Vermeulen, M., Hubner, N. C. & Mann, M. High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr. Opin. Biotechnol.* **19**, 331–337 (2008).
110. Glatter, T., Wepf, A., Aebersold, R. & Gstaiger, M. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol. Syst. Biol.* **5**, 237 (2009).
111. Goudreault, M. *et al.* A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol. Cell. Proteomics* **8**, 157–171 (2009).
112. von Kriegsheim, A. *et al.* Cell fate decisions are specified by the dynamic ERK interactome. *Nature Cell Biol.* **11**, 1458–1464 (2009).
113. Blagoev, B. *et al.* A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nature Biotech.* **21**, 315–318 (2003).
114. Schulze, W. X., Deng, L. & Mann, M. Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol. Syst. Biol.* **1**, 2005.0008 (2005).
115. Hanke, S. & Mann, M. The phosphotyrosine interactome of the insulin receptor family and its substrates IRS-1 and IRS-2. *Mol. Cell. Proteomics* **8**, 519–534 (2009).
116. Selbach, M. *et al.* Host cell interactome of tyrosine-phosphorylated bacterial proteins. *Cell Host Microbe* **5**, 397–403 (2009).
117. Trinkle-Mulcahy, L. & Lamond, A. I. Toward a high-resolution view of nuclear dynamics. *Science* **318**, 1402–1407 (2007).
118. Waanders, L. F. *et al.* Quantitative proteomic analysis of single pancreatic islets. *Proc. Natl Acad. Sci. USA* **106**, 18902–18907 (2009).
119. Mallick, P. *et al.* Computational prediction of proteotypic peptides for quantitative proteomics. *Nature Biotech.* **25**, 125–131 (2007).
120. Makarov, A. *et al.* Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* **78**, 2113–2120 (2006).
121. Olsen, J. V. *et al.* Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021 (2005).
122. Zubarev, R. & Mann, M. On the proper use of mass accuracy in proteomics. *Mol. Cell. Proteomics* **6**, 377–381 (2007).
123. Mann, M. & Kelleher, N. L. Precision proteomics: the case for high resolution and high mass accuracy. *Proc. Natl Acad. Sci. USA* **105**, 18132–18138 (2008).
124. Wang, Z. *et al.* Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Sci. Signal.* **3**, ra2 (2010).
125. Soufi, B. *et al.* Global analysis of the yeast osmotic stress response by quantitative proteomics. *Mol. Biosyst.* **5**, 1337–1346 (2009).
126. Pan, C., Gnad, F., Olsen, J. V. & Mann, M. Quantitative phosphoproteome analysis of a mouse liver cell line reveals specificity of phosphatase inhibitors. *Proteomics* **8**, 4534–4546 (2008).

Acknowledgements
We thank our co-workers for fruitful discussions, especially J. Olsen for providing figure 3. The Protein Research Center (CPR) is supported by a generous donation from the Novo Nordisk Foundation. This project was supported by the European Commission's 7th Framework Program PROteomics SPECification in Time and Space (PROSPECTS, HEALTH-F4-2008-021,648).

Competing interests statement
The authors declare no competing financial interests.

DATABASES
UniProtKB: http://www.uniprot.org
ATM | ATR | Cdk1 | EGF | EPHB2 | GRB2 | MEK1 | MEK2 | PTP1B | STAT5 | TNFα

FURTHER INFORMATION
Matthias Mann's homepage:
http://www.biochem.mpg.de/mann/
NNF Center for Protein Research:
http://www.cpr.ku.dk/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF