REVIEW

Proteomic approaches to the analysis of multiprotein signaling complexes

Wei Yang^{1, 2}, Hanno Steen³ and Michael R. Freeman^{1, 2}

¹ The Urological Diseases Research Center, Department of Urology, Children's Hospital Boston, Boston, MA, USA

² Departments of Surgery, Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA

³ Department of Pathology, Harvard Medical School and Children's Hospital Boston, Boston, MA, USA

Signal transduction is one of the most active fields in modern biomedical research. Increasing evidence has shown that signaling proteins associate with each other in characteristic ways to form large signaling complexes. These diverse structures operate to boost signaling efficiency. ensure specificity and increase sensitivity of the biochemical circuitry. Traditional methods of protein analysis are inadequate to fully characterize and understand these structures, which are intricate, contain many components and are highly dynamic. Instead, proteomics technologies are currently being applied to investigate the nature and composition of multimeric signaling complexes. This review presents commonly used and potential proteomic methods of analyzing diverse protein complexes along with a discussion and a brief evaluation of alternative approaches. Challenges associated with proteomic analysis of signaling complexes are also discussed.

Keywords:

Cross-linking / Mass spectrometry / Post-translational modification / Quantitative proteomics / Signaling complexes

1 Introduction

The modern field of cell signaling can be traced back to the mid-1950s, when it was discovered that reversible phosphorylation could control enzyme activity [1]. Since the late

Correspondence: Dr. Michael R. Freeman, John F. Enders Research Laboratories, Suite 1161, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115, USA E-mail: michael.freeman@childrens.harvard.edu Fax: +1-617-730-0238

Abbreviations: BN-PAGE, blue native-PAGE; CBP, calmodulinbinding peptide; co-IP, co-immunoprecipitation; ECD, electron capture dissociation; EGFR, epidermal growth factor receptor; ETD, electron transfer dissociation; ICAT, isotope-coded affinity tags; NMDA, N-methyl-D-aspartate; PCS, peptide-concatenated standard; PKC, protein kinase C; QCAT, concatemer of Q peptides; SILAC, stable isotope labeling with amino acids in cell culture; TAP, tandem affinity purification.

© 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1980s, astonishingly rapid progress has been made in understanding the mechanisms of signal transduction. Many old concepts have been abandoned or revised and new ones have emerged. In the past two decades, one of the most important new concepts is that signaling molecules organize into multiprotein assemblies referred to as signaling complexes and, occasionally, signalosomes or transducisomes [2, 3].

Signaling complexes are composed of enzymes such as kinases and phosphatases, their substrates, and adaptor/scaffold proteins, which cohere using modular protein-protein interaction domains [4]. Adaptor/scaffold proteins are composed exclusively of these interaction domains and lack intrinsic enzymatic activity. The difference between adaptors and scaffolds is that adaptors only link together two partners whereas scaffolds link together three or more partners [4]. Recently, it has become clear that adaptor/scaffold proteins play essential roles in cell signaling by: (i) serving as backbones of signaling com-

InterScience[®]

Received: July 9, 2007 Revised: October 23, 2007 Accepted: October 23, 2007



plexes and achieve different outputs by using distinct combinations of binding partners [5]; (ii) controlling oligomerization [6, 7]; (iii) targeting signaling complexes to specific compartments [8, 9]; and (iv) acting as sorting adaptors and coincidence detectors to enhance specificity in cellular responses [10–12].

Rather than freely diffusing in the cytoplasm, signaling complexes are often stably associated with plasma membranes via protein-protein interactions (e.g. transmembrane receptor complexes), transiently attached to membranes via lipid-protein interactions (e.g. Akt complexes), or reversibly associated with membranes via synergistic combinations of protein-protein interactions and lipid-protein interactions (e.g. Ste5 complexes) (reviewed in [13]). In addition, signaling complexes are also found on various subcellular membranes such as Golgi, mitochondrial, and nuclear membranes and the cytoskeleton [9]. The assembly of various components into tight multiprotein clusters on or near cell membranes provides great advantages for the enhancement of signaling efficiency and specificity [14]. However, inappropriate organization of signaling complexes results in deranged signaling and has been linked to a variety of diseases like cancer [15, 16].

Due to the biological and therapeutic importance of these structures, understanding the nature of signaling complexes is a major focus of signal transduction research. To gain a comprehensive understanding of these structures, several basic questions need to be addressed. What is the composition and stoichiometry of these protein assemblies? When, where, and how are signaling complexes formed and regulated? How do these structures regulate signaling pathways downstream? A variety of techniques, including proteomics tools and fluorescence resonance energy transfer microscopy, have recently been employed to address these questions [17]. Among these techniques, proteomic approaches have been demonstrated to be very powerful for characterization of multiprotein complexes. These tools were adopted to investigate a number of signaling complexes such as those harboring the N-methyl-D-aspartate (NMDA) receptor [18-20], protein kinase C (PKC)-ɛ [21-23], NF-κB [24], and Akt/ protein kinase B [25]. Here we review commonly used and potential proteomic approaches to analyzing multiprotein signaling complexes (see Fig. 1). The strengths and shortcomings of these approaches, and the challenges ahead, are discussed.

2 Isolation of signaling complexes

Isolation of signaling complexes is a critical step for further MS analysis. The goal of this step is to purify the target structure with a minimum of contaminants, as well as to maintain the integrity of the native binding partners and their interactions. In other words, a balance has to be found between reducing the non-specifically bound proteins and maintaining the interactions with weak binding partners.



Figure 1. Proteomics workflow for the analysis of signaling complexes. First, complexes are separated and purified. Then the protein components are identified by MS. True interactors can be distinguished from nonspecific interactors by employing quantitative proteomic techniques. The stoichiometry of subunits can be accurately determined by using absolute quantitative proteomic techniques. Protein PTMs are characterized by MS after the PTM peptides are enriched. The spatial organization of the signaling complexes can be determined by combining chemical cross-linking with MS. Finally, the signaling complexes are reconstructed and their functions can be predicted.

2.1 Subcellular fractionation

Rather than being evenly distributed in the cytoplasm, signaling complexes are localized within specific compartments. Increasing evidence has shown that subcellular localization is a major determinant of their molecular composition and function. A case in point is the targeting of protein kinase A complexes to specific compartments, such as the plasma membrane, mitochondria, cytoskeleton, and centrosome, by distinct A-kinase anchoring proteins [8, 9]. Therefore, the first step is often to obtain a concentrated subcellular fraction of interest using well-established techniques [26-31]. After homogenization of cells or tissue (by douncing, sonicating, etc.), released organelles are usually separated by differential centrifugation, density gradient centrifugation or both procedures in sequence. Recently, many companies such as Sigma, Pierce and Calbiochem have launched subcellular fractionation kits for isolation of mitochondria, lysozyme, endoplasmic reticulum, peroxisome, plasma membrane, and synaptic vesicles. Though the kits are relatively expensive, they facilitate rapid and easy isolation of specific organelles with high purity, which can be

assessed using appropriate marker detection kits. Purified organelles are a robust starting material for subsequent isolation of subcellular location-specific protein complexes.

Notably, a variety of signaling proteins, such as seventransmembrane receptors, heterotrimeric G proteins, Gprotein coupled enzymes, and receptor and nonreceptor tyrosine kinases, are enriched in cholesterol-rich and liquidordered microdomains called lipid rafts or detergent-resistant membranes [32]. It has become clear that lipid rafts/ detergent-resistant membranes serve as a crucial platform for signal transduction and that many types of signaling complexes are assembled within cholesterol-rich microdomains upon ligand activation [33]. Thereby, isolation of lipid rafts may be a good starting point for studying a variety of signaling complexes, particularly those associated with plasma membrane receptor systems activated by extracellular soluble ligands. Numerous detergent- and nondetergent-based methods have been applied to isolate lipid rafts (reviewed in [34]). However, the classic isolation method utilizing Triton X-100 insolubility and subsequent floatation on sucrose density gradient is still the "gold standard" [35].

2.2 Affinity purification

Multiprotein complexes can be isolated *via* a variety of methods. Conventional methods such as LC, ultracentrifugation, and sucrose density gradient centrifugation have been used successfully. However, these methods are usually time-consuming. In contrast, affinity-based methods usually provide high-efficiency purification. Among these methods, co-immunoprecipitation (co-IP), epitope-tagging, tandem affinity purification (TAP), and GST-pulldown are commonly used.

2.2.1 Co-IP

Co-IP is probably the most frequently employed method for multiprotein complex purification. The principle of this approach is that when one protein in the complex is precipitated, additional members of the complex are captured as well, irrespective of whether they bind directly or indirectly. This method has been successfully applied to purify deathinducing signaling complexes [36], NMDA receptor complexes [18-20], PKC complexes [21-23], mGluR5 complexes [37], and Akt complexes [25]. An obvious advantage of this approach is that protein complexes are isolated in a state closest to the physiological condition. In addition, when a good quality antibody is available, co-IP is much faster to perform than the other commonly used methods, since there is no need to clone and express component(s) of the complex. The disadvantage of this method is that antibodies may cross-react with other nonspecific proteins. Consequently, interacting partners of these irrelevant proteins are also pulled down, resulting in a comparatively high background. Although using harsher washing conditions may reduce contamination, some weakly bound but physiologically relevant components will be lost. Another serious problem of co-IP is that antibodies have a tendency to leach from the support matrix during elution steps. Large amounts of antibodies present in the eluate may interfere with the detection of some low-abundance components, especially when samples are not separated by gel electrophoresis prior to MS analysis. In addition, good quality antibodies for co-IP are not always available, thus limiting the widespread use of the method. Signaling proteins are normally present at low levels, which can further complicate isolation by co-IP and subsequent MS analysis.

2.2.2 Epitope-tagging

To address the above-mentioned problems, epitope-tagging methods, first described in 1984 [38], have been developed and are widely used to isolate protein complexes in vivo. In this approach, a bait protein is genetically fused with an epitope tag, expressed in the host cell, and the complex captured by binding of the epitope tag (e.g. His, calmodulin-binding peptide (CBP), Strep II) or immunoaffinity tag (e.g. FLAG, HA, c-Myc) with the affinity column. Finally, specific elution is accomplished with the epitope-tag analog. Recently, the epitope-tagging method has been applied to purify target of rapamycin [39] and PAK1-BPIX-GIT1-Paxillin complexes [40]. Additionally, this method has been used for large-scale protein complex analysis [41, 42]. Compared with co-IP, the epitope-tagging method has remarkable advantages in many ways. First, it is generally quicker and easier to genetically tag a protein than to raise a high quality antibody. Second, protein complexes can be purified with standardized gentle conditions by using commercially available kits. In this way, high-throughput purification is achievable. Third, bait proteins can be overexpressed in a host cell, thereby facilitating detection and analysis of low abundance complexes. However, it must be kept in mind that, analogous to the inhibition of antibody-antigen precipitation in the precipitin test, the overexpression of certain proteins (e.g. adaptor/scaffold proteins), which may bind two or more other proteins, may inhibit the formation of signaling complexes [2, 40, 43]. In this case, the expression level of the bait must be tightly controlled, for example, by employing an inducible expression system. The major disadvantage of the epitope-tagging approach is that the artificially introduced tag may interfere with proper protein folding and function, as well as the ability of the bait to interact with other proteins. Therefore, it is highly recommended to create N-terminal and C-terminal fusions in parallel.

2.2.3 TAP

Although one-step affinity purification can be very effective in the isolation of multiprotein complexes, nonspecific binding is frequently observed. To achieve higher purity, a dual purification strategy termed TAP was introduced in 1999 [44]. The first proposed TAP-tag construct system includes protein A (ProtA) of Staphylococcus aureus and CBP as tandem tags with a tobacco etch virus protease cleavage site in between. The tag cassette can be fused to either the Nor C-terminal end of a target protein, and the fusion protein is expressed in a host cell or organism at, or close to, its natural level. The first purification step of the fusion protein and its associated components involves binding of ProtA to an IgG matrix. After gentle washing, purified complexes are released by tobacco etch virus enzyme cleavage while nonspecifically bound proteins are left behind. In the second purification step, the eluate of the first affinity step is incubated with calmodulin-coated beads in the presence of calcium. Following gentle washing, the fusion protein and its binding partners are specifically released via calcium chelation. Again, proteins that interact non-specifically with the support matrix are left behind.

Although this method was originally developed for use in yeast, it was quickly adapted to higher eukaryotes such as insect cells [45], human cells [46], and plants [47]. Recently, a variety of new TAP tags such as ProtA-FLAG [48], His6-FLAG [49], HPM (His-Myc) [50], and PTP (ProtA-ProtC) [51] as well as multiple affinity purification tags such as CHH (CBP-His-HA) [52], HCHH (His6-CBP-His8-HA) [53], and BT-ProtA-CBP [54] have been developed to increase recovery of protein complexes and/or to improve specificity. Very recently, two new types of TAP tags have been developed to preserve PTMs (e.g. ubiquitination) from enzymatic hydrolysis [55] and increase recovery by about ten-fold [56], respectively. In the first study [55], Kaiser and co-workers developed a histidinebiotin tandem tag to allow two-step purification under fully denaturing conditions such as 8 M urea. The authors demonstrated that the new TAP method can protect ubiquitination from hydrolysis by successfully purifying an in vivo cross-linked ubiquitin ligase complex, thereby allowing the identification of tens of ubiquitination sites with MS. In the second study [56], Superti-Furga and colleagues developed a much more efficient TAP procedure for purifying protein complexes in mammalian cells. Compared with the original yeast TAP tag, the new GS-TAP tag, which is based on protein G and the streptavidin-binding peptide, increases the yield of protein complexes by about an order of magnitude. With the GS-TAP tag, the authors successfully purified and identified a DNA-PK protein complex from only 5×10^7 HEK293 cells. Since much less starting material is required, the GS-TAP method is expected to be very useful for purification of protein complexes from a variety of cell types, including those not easily cultivated in large quantities (e.g. neuronal and immune cells) and even primary cells. With the high efficiency and robustness of the procedure, largescale analysis of human cellular machinery becomes more feasible.

The TAP strategy has not only been applied to purify static protein complexes on a normal lab scale [44, 57] and a proteomic scale [58, 59] but also been developed to study dynamic structures, such as signaling complexes in the human TNF- α /NF- κ B signal transduction pathway [24], human Par

protein complexes [46], worm nicotinic acetylcholine receptor complex [60], and human 14-3-3 ζ complex in transgenic mice [61]. The key advantage of TAP is that, compared with one-step purification, it significantly improves the purity of the isolated protein complex [44, 49, 55]. However, it also suffers from the aforementioned disadvantages of epitopetagging. In addition, time-consuming repetitive washes under non-physiological conditions often lead to the loss of weak or transient protein-protein interactions.

2.2.4 GST pulldown

GST pulldown is commonly used for purification of multiprotein complexes in vitro. In this approach, the protein/domain of interest is expressed in Escherichia coli as a GSTfusion and immobilized on a solid support (e.g. glutathionesepharose beads), while GST alone is often used as a control. Once a cellular lysate is applied to the beads or column, target protein/domain competes with endogenous protein for interacting proteins, forming protein complexes in vitro. Bound complexes can then be pulled-down with the beads or column and eluted after several washings to remove unbound proteins. In the control group, those proteins that bind to the GST coated beads or column are usually false positives and are subtracted from the protein complex data. Recently, this approach has been applied to purify activated human epidermal growth factor receptor (EGFR) complex [62] and mouse macrophage-colony stimulating factor receptor complex [63]. An advantage of this method is that it is robust and easy to use. Moreover, because high concentrations of the purified target can be immobilized on the affinity beads or column, proteins that have low affinity for the protein of interest can be retrieved. However, the method has several limitations. First, not all proteins can be easily overexpressed in a soluble form in E. coli. Second, fusion proteins expressed in E. coli may lack the PTMs required for proteinprotein interactions. Third, certain proteins that do not interact with the protein/domain of interest physiologically may also be purified, thereby increasing false positive rates.

2.3 Separation of affinity purified signaling complexes

Due to their dynamics and complexity, signaling complexes enriched by the above-mentioned approaches are sometimes a population of distinct multiprotein assemblies [9, 64–66]. To isolate an individual complex, further separation is required. This can be achieved by blue native (BN)-PAGE [67–69], sucrose density gradient ultracentrifugation [70–72], and SEC [23, 70].

BN-PAGE was firstly introduced by Schägger and Von Jagow for isolation of membrane protein complexes in 1991 [67]. In this method, the anionic dye Coomassie blue G-250, instead of ionic detergent, is introduced to make the multiprotein complexes negatively charged while retaining most protein-protein interactions. Consequently, the protein complexes can be electrophoretically separated by the sieving effect of the polyacrylamide gel with running buffer at a physiological pH of 7.5. Using this method, Hedman and coworkers separated multiple glucocorticoid receptor complexes and validated the existence of glucocorticoid receptor in at least four of the high molecular weight complexes with immunoblotting [66]. The advantage of BN-PAGE is that it is a robust method for separation of multiprotein complexes with high resolution [69]. Moreover, BN-PAGE can be coupled to colorless native PAGE to obtain even purer preparations [69, 73]. Nevertheless, some protein-protein interactions can be affected to some extent by the presence of the anionic CBB dye [69]. Another drawback of BN-PAGE is that giant protein complexes are restricted from entering the matrix, presumably due to the physical constraints of the polyacrylamide pore size. Although this problem can be alleviated by using an agarose-acrylamide composite native gel system [74], the high resolution of BN-PAGE is impaired.

Sucrose density gradient ultracentrifugation is a classic method for isolation of cell organelles and enveloped viruses, etc. This method has also been applied to separate multiprotein complexes. Different protein complexes differ in specific density, so they can float at different levels in a density gradient and thus be separated. Using this method, Kirschner and co-workers purified a WAVE1-containing complex from crude complex-containing chromatographic fractions [70]. SEC, also known as gel filtration chromatography, has also been used to separate multiprotein complexes. For example, Ping and colleagues used this method to separate signaling complexes containing PKCE-Akt-eNOS modules, some of which were much larger than thyroglobulin (669 kDa) [23]. Compared with BN-PAGE, sucrose density gradient ultracentrifugation and SEC only have moderate resolution but they work better for separation of giant protein complexes.

3 Protein identification, validation and stoichiometry determination

3.1 Protein identification by MS

In the post-genome era, MS is the preferred method for protein identification. Compared with the traditional protein sequencing method, Edman degradation, MS offers superior speed, sensitivity, versatility, and is applicable to protein mixtures. However, it is not suitable for analyzing native complex mixtures (*e.g.* a multiprotein complex) directly. Therefore, protein/peptide separation has to be performed prior to MS analysis. Among all the combinations of protein/ peptide separation and MS, 2-DE in combination with MALDI-MS, multi-dimensional LC coupled with ESI-MS/ MS, and 1-DE followed by LC in combination with ESI-MS/ MS are the most frequently used strategies.

2-DE is an established and widely accessible technology which has unparalleled resolving power (>1000 spots/gel)

and can separate different protein isoforms and PTM protein species [75-77]. Protein spots can be visualized by chemical stains such as CBB stain, silver stain, and fluorescent stains. In addition, glycoproteins and phosphoproteins can be selectively detected by fluorescent stains such as Pro-Q Emerald [78, 79] and Pro-O Diamond [80] or Phos-tag [81], respectively. After in-gel digestion, the peptide mixture is mixed with an appropriate matrix and analyzed by MALDI-TOF MS to obtain a PMF. The PMF is then in silico compared with "virtual" fingerprints obtained by theoretical cleavage of protein sequences stored in protein databases. The top-scoring proteins are retrieved as possible candidate proteins. A major advantage of MALDI-MS is that it is experimentally simple and fast to perform, relatively tolerant to salt and detergent, and very sensitive (fmol) and accurate (ppm). Another advantage of MALDI-MS is that the results are usually straightforward to interpret. However, proteins with low molecular weight (<15 kDa) or a high level of modifications present problems because effective proteolytic peptide number is not sufficient for unambiguous identification. In addition, PMF is dependent on the size of the available protein sequence databases. Nevertheless, with the advent of MALDI-TOF/TOF instrumentation [82], many of these problems can be effectively addressed.

Multi-dimensional LC-ESI-MS/MS is a gel-free approach for protein identification. In this method, isolated multiprotein complexes are digested in solution. The resulting peptides are separated by strong cation exchange chromatography followed by RP chromatography and analyzed in line by ESI-MS/MS. The mass spectrometer accurately and sensitively records the m/z of the precursor ions and their product ions generated by CID. The resulting mass spectra are then searched against protein databases to determine the amino acid sequences of the peptides and the identities of the proteins. In the past few years, a number of protein complex profiling studies have been successfully performed on this platform [50, 83-87]. Compared with the 2-DE gelbased method, multi-dimensional LC-ESI-MS/MS is more streamlined and overcomes several limitations of 2-DE such as the robust detection of hydrophobic and low-abundance proteins. However, proficiency and sophistication in multidimensional chromatography are required.

1-DE LC-ESI-MS/MS is a compromise between the advantages and disadvantages of the above-mentioned 2-DE and shotgun proteomic profiling methods [88]. In this method, proteins are normally separated by conventional SDS-PAGE. Gel slices or protein bands are excised and subjected to in-gel digestion. Generated peptides are separated and sequenced by LC-ESI-MS/MS and finally the proteins are identified. Compared with 2-DE-MALDI-MS and multi dimensional LC-ESI-MS/MS, 1-DE LC-ESI-MS/MS is technically simple and allows large-scale and high-throughput identification of proteins, including those with transmembrane domains.

For practical analysis of multiprotein complexes, the choice among the aforementioned methods should depend

on the primary interest of the investigators and the complexity of the interaction partners. If PTMs are of primary interest, a 2-DE gel-based method is a good choice. If the interest is in identifying as many proteins as possible, 1-D LC-ESI-MS/MS or multi-dimensional LC-ESI-MS/MS is more suitable. In addition, if the interaction partners are few (<50), LC-MS/MS alone may be sufficient for identification [89].

3.2 Validation of interaction partners by relative quantitative proteomics techniques

As mentioned before, isolated protein complexes are usually contaminated by irrelevant (nonspecific) proteins. Sometimes the problem is so severe that bona fide interaction partners have to be identified amidst a sea of false positive interactors. For example, Selbach and Mann showed that after co-IP, more than 95% of the identified proteins might be false positive interactors [90]. This critical issue highlights the necessity of validating the identified interaction partners. Traditionally, the identities of the proteins are commonly verified by immunoblotting experiments; the interactions between the interaction partners are usually validated by co-IP, GST-pulldown and yeast two hybrid experiments; and the colocalization of the proteins is determined via confocal microscopic analyses. However, these methods are usually labor-intensive and time-consuming. Fortunately, recently developed quantitative proteomics technologies provide a solution to distinguish specific from nonspecific interactions rapidly and sensitively. Consequently, efforts required to validate the interaction partners by the traditional methods are greatly reduced.

Although label-free quantification methods are under continuous development [91], they are still in their infancy. Currently, the most widely used quantitative proteomics technologies are label-based, including Stable Isotope Labeling with Amino acids in Cell culture (SILAC) [92], Isotope-Coded Affinity Tag (ICAT) [93], iTRAQ [94], and DIGE [95]. These methods have been proven to be able to yield quantitative results with reasonable accuracy [96–102]. A major difference among these methods is where the isotope or fluorescent labeling is performed (see Fig. 2). Given that this area has been well reviewed [103–107], the principles of these quantitative proteomic techniques and their strengths and shortcomings are only summarized below.

3.2.1 MS-based quantification methods

In the MS-based methods, the samples to be compared are labeled with 'light' (native) and 'heavy' stable isotopes (*e.g.* ²H, ¹³C, ¹⁵N, *etc.*), respectively, by either metabolic (*e.g.* SILAC) or chemical (*e.g.* ICAT, iTRAQ) methods. The principle of the quantification using isotope labeling is that isotopically labeled peptides have identical physico-chemical properties except for the mass, so they can be concurrently

eluted and ionized and subsequently distinguished according to their different masses. The ratio of their peak intensities corresponds to the relative abundance ratio of the peptides (and proteins).

3.2.1.1 SILAC

SILAC is performed at the cell culture stage and stable isotopes are metabolically introduced (Fig. 2A). In SILAC, cells are grown in the presence of isotopically distinguishable amino acids (e.g. L-arginine and ¹³C₆-labeled L-arginine). After at least five doublings, equal amounts of 'light' and 'heavy' labeled cells are combined. The proteins are extracted and digested with a sequence-specific protease (e.g. trypsin). The resulting peptides are then sequenced and quantified by MS. Because the 'light' and 'heavy' cell populations are combined immediately after harvesting, any protein/peptide losses in the remaining steps would occur equally for the two groups of samples. Therefore, compared with the other quantitative proteomics techniques, SILAC is expected to have the smallest technical variations. In addition, unlike chemical labeling methods, which can be tedious, this metabolic labeling method is relatively easy to perform. The disadvantage of SILAC is that the complexity of the sample is doubled, thereby resulting in more severe undersampling. Moreover, SILAC can not be directly applied to study samples that cannot be labeled metabolically (e.g. human biopsies), though the recently developed culture-derived isotope tag [108] strategy can overcome this problem to a large extent. However, this disadvantage is not greatly inhibiting with regard to signaling complex analysis, because much signal transduction research is generally carried out using cultured cells. In fact, exciting results from Matthias Mann's and other groups have demonstrated that SILAC is a powerful tool for characterization of cell signaling mechanisms (reviewed in [109]).

3.2.1.2 ICAT

ICAT is a cysteine-specific protein tagging method (Fig. 2B). The first-generation ICAT is composed of a cysteine-reactive iodoacetyl group, a differentially isotope-coded linker region, and a biotin affinity tag [93]. Later, several modified versions of the ICAT such as solid-phase ICAT [110], cleavable ICAT [111], and visible ICAT [112] were designed to alleviate certain problems encountered with the original ICAT method (reviewed in [107]). In the ICAT approach, two different protein mixtures are treated with isotopically 'light' (e.g. ¹H, ¹²C) and 'heavy' (e.g. ²H, ¹³C) ICAT reagents, respectively. The two protein mixtures are combined and proteolyzed (usually by trypsin). ICAT labeled peptides are purified by biotin-avidin affinity chromatography and then analyzed by LC-MS/MS. The major advantage of ICAT is that it allows selective capture and analysis of cysteine-containing peptides, thus significantly reducing the complexity of the peptide mixture.





However, proteins that lack cysteine residues (8% for yeast and less than 5% for human) cannot be detected by this method.

3.2.1.3 iTRAQ

iTRAQ is a global tagging method and is performed at the peptide level (Fig. 2C). An iTRAQ reagent consists of three components: (i) an amine-reactive group, which can covalently link to the N-terminal amine of peptides and the side chain of lysines; (ii) a charged reporter group with an MS/MS fragmentation site, which may give rise to diagnostic ions of m/z 114, 115, 116 or 117; and (iii) a neutral mass balance group to maintain an overall mass of 145 Da. Owing to the

labeled peptides appear as single peaks in an MS scan, thus reducing the possibility of peak overlap and increasing detection sensitivity. When iTRAQ-tagged peptides are subjected to MS/MS analysis, the mass balance moiety is released as a neutral fragment and isotope-encoded reporter ions are liberated to provide quantitative information on proteins. Another key advantage of iTRAQ is that it is a multiplexed technology, allowing the identification and quantification of up to four samples simultaneously. A new generation of iTRAQ (eight-channel), which has very recently been applied to compare the activity of six leukaemogenic tyrosine kinases [113], will be launched by Applied Biosystems soon. However, iTRAQ also has some limitations. First of all, since all peptides can be labeled by iTRAQ,

Proteomics 2008, 8, 832-851

Proteomics 2008, 8, 832-851



Figure 2. Comparison of the most commonly used quantitative proteomic techniques for discriminating *bona fide* interactors from nonspecific interactors. (A) SILAC. (B) ICAT. (C) iTRAQ. (D) DIGE. See Section 3.2 for more detail. In each panel, interacting A, B, C, D represent authentic binding partners in a multiprotein complex; N1 and N2 represent nonspecific binding proteins; red-violet spheres attached to peptides represent heavy isotopes.

sample complexity is maintained. Consequently, high-abundance proteins are frequently identified. Secondly, since the labeling is performed at the peptide level, samples have to be processed separately until after digestion. This increases the potential for errors introduced during cell lysis, protein extraction, and/or proteolysis. Finally, since the derivatized peptides are indistinguishable in MS, MS/MS spectra have to be acquired for quantification, thus requiring more analysis time than performing result-dependent analysis on differential peptide pairs in MS.

3.2.2 Gel-based quantification method – DIGE

DIGE is a gel-based quantitative proteomic technique (Fig. 2D). In this approach, proteins from different samples are labeled with spectrally-resolvable fluorescent dyes (Cy2, Cy3, Cy5), pooled, and separated by 2-DE. Protein spots are then imaged with different excitation and emission filters to detect the protein profiles for each sample. Differences in protein levels are then calculated according to spot color and intensity. Because proteins from different samples are sepa-

rated in a single gel, the advantage of DIGE is the elimination of gel-to-gel variation and, accordingly, overall technical variation is significantly reduced. Moreover, DIGE is a very sensitive technique, capable of detecting as little as 0.5 fmol of protein and protein differences down to \pm 15% [114]. However, running 2-D gels is labor-intensive and time-consuming. In addition, proteins have to be identified separately.

3.2.3 Quantitative proteomics for the differentiation of specific and non-specific binding partners

Recently, several groups have demonstrated that quantitative proteomic techniques are very effective in distinguishing specific from nonspecific interactions [62, 66, 115–119]. The principle of this strategy is straightforward. When an experimental sample/control pair is labeled with isotopically or fluorescently distinguishable reagents, nonspecific interactors should be present in a nearly 1:1 ratio, while specific intearctors can be identified by their significantly changed abundance (see Fig. 2). In 2003, Mann's and Aebersold's group independently reported the application of this strategy [62, 115]. Using SILAC, Mann and coworkers identified 228 Grb2-SH2-interacting proteins after GST pulldown, of which only 28 were selectively enriched (with ratio >1.3) upon EGF stimulation [62]. Using ICAT, Aebersold and colleagues identified 326 proteins in a RNA polymerase II preinitiation complex after a simple one-step DNA affinity purification, of which only 49 proteins (with ratio >1.9) were accepted as bona fide components of the complex. In fact, 45 of the 49 proteins were known components of the complex [115], demonstrating the effectiveness of this strategy in selecting for biologically important protein interactions. Subsequently, this strategy was successfully applied to study multiprotein complexes, such as (i) proline-rich Son of Sevenless peptide-binding partners using SILAC [116], (ii) MyD88-interacting partners using amino acid coded mass tagging [117], (iii) insulin-regulated glucose transporter-interacting proteins using SILAC [118], (iv) Grb2-interacting proteins using iTRAQ [119], and (v) glucocorticoid receptor-interacting proteins using DIGE [66]. Very recently, by combining SILAC, RNA interference, and co-IP, Mann and Selbach have developed an ingenious QUICK (QUantitative Immunoprecipitation Combined with Knockdown) technology to identify cellular interaction partners of endogenous proteins in mammalian cells with very high confidence [90]. This method offers substantial advantages over many other approaches. It can effectively eliminate false positives, such as proteins with cross-reactivity to the antibody and proteins non-specifically bound to the beads or the antibodies. Notably, the method only requires depletion of the target sufficient to yield significant protein ratios rather than complete depletion of the target. In summary, the quantitative proteomic technologybased strategy provides a powerful new way of reliably distinguishing specific complex components from co-purifying proteins. It obviates the need for extensive purification (*e.g.* TAP), thereby saving time, increasing protein complex recovery, and reducing the risk of losing weakly and transiently interacting proteins.

3.3 Stoichiometry determination by absolute quantitative proteomics techniques

When multiprotein complexes are resolved into individual subunits by SDS-PAGE or 2-DE, the stoichiometry of protein subunits can be assessed by comparing the intensity of the bands or the spots. However, this method is usually semiquantitative at best. Accurate stoichiometric information can be obtained by MS-based absolute quantification. In this strategy, known amounts of stable isotope-labeled peptides, either synthetic or proteolytic, are added to the digested complexes and used as internal standards (see Fig. 3). Consequently, the absolute amount of each subunit is determined and the stoichiometry can be calculated. The first reported absolute quantification using synthetic peptides with incorporated stable isotopes was in 1996 [120]. However, the idea was not significantly extended until 2003, when Gygi and co-workers developed a strategy termed AQUA (Absolute QUAntification) [121]. The major advantage of AQUA is that it can measure the levels, not only of proteins, but also of PTMs directly from cell lysates. Nevertheless, synthesis of isotopically labeled peptides is labor-intensive and time-consuming. Moreover, the accuracy of quantification may be impaired by variable purity of synthetic peptides and incomplete enzyme cleavage of target proteins. To overcome these weaknesses, two groups have independently developed novel absolute quantification strategies, concatemer of Q peptides (QCAT) [122] and peptide-concatenated standard (PCS) [123] using artificial peptide-concatenated proteins. The principles of the two strategies are similar. For each protein to be quantified, a unique Q peptide is selected according to several criteria [122, 123]. The sequences of the selected Q peptides are concatenated in silico and used to design a gene. Then the artificial gene is expressed in a heterologous expression system (e.g. E. coli), which permits metabolic labeling with stable isotopes. The synthetic protein is purified to homogeneity and used as a standard for subsequent quantification. Different from QCAT, PCS carries each peptide with its natural flanking sequences on both sides to precisely recapitulate the cleavage efficiency in its parental protein, thereby considerably improving the accuracy of quantification. The PCS strategy has been validated by accurately quantifying component stoichiometry of the eIF2BeIF2 complex [123]. The advantages of the artificial protein strategy for absolute quantification are significant. First, the daunting task of preparation and handling of many synthetic peptides is eliminated. Second, equimolar addition of multiple standard peptides to the sample is guaranteed. Third, technical variation caused by incomplete enzyme cleavage is reduced. Nevertheless, these methods do have their shortcomings. For example, the size of synthetic proteins is limited because gene construction and protein expression is



Figure 3. Schematic diagram of absolute quantitative proteomic techniques, *i.e.* absolute quantification (AQUA), QCAT and PCS. See Section 3.3 for more detail. In this figure, interacting A, B, C, D represent a multiprotein complex; red-violet spheres attached to peptides represent heavy isotopes.

much more difficult for very large proteins. Therefore, for stoichiometry determination of macromolecular protein complexes, multiple QCATs or PCSs have to be designed, expressed and purified.

4 Characterization of PTMs

PTMs play a crucial role in protein activity, localization, turnover, and interactions with other proteins. In higher eukaryotic cells, signaling complexes are mainly assembled on demand [2]. The assembly of an intracellular signaling complex on the tail of a membrane anchored receptor is usually triggered by reversible PTMs, phosphorylation in particular [2, 13]. Subsequently, some signaling proteins that are recruited to form the receptor complexes are post-translationally modified, especially by phosphorylation, to achieve the active forms that trigger downstream signaling. Due to the great importance of PTMs for biological function, mapping PTMs is indispensable for full characterization of protein complexes. In view of the central importance of phosphorylation, next we focus on the commonly used methodologies for detection and enrichment of phosphorylated proteins and peptides as well as the progress in MS analysis of phosphorylation, following a brief introduction to bioinformatic analysis of identified proteins. Readers interested in learning about the recent advances in proteomic analysis of other PTMs are directed to recent reviews [124–131].

4.1 Bioinformatic prediction of protein phosphorylation

With the accelerating progress in applications of proteomics, the number of experimentally identified PTM sites grows exponentially. Accordingly, numerous PTM databases have been constructed to compile the colossal amount of data. Based on these data, a variety of programs have been developed to predict PTM sites from a protein sequence with fairly high accuracy. For instance, the PhosphoSite (http:// www.phosphosite.org) [132] database contains more than 58000 phosphorylation sites, among which more than 17000 are curated from the literature. The Phospho.ELM (http://phospho.elm.eu.org) [133] database contains more than 16000 experimentally verified phosphorylation sites in eukaryotic proteins. The dbPTM (http://dbptm.mbc.nctu. edu.tw) [134] database contains more than 22 000 experimental and 1 800 000 putative phosphorylation sites. The human protein reference database (HPRD, http:// www.hprd.org) [135, 136] lists more than 8800 expert-curated phosphorylation sites and compiles more than 13 000 additional phosphorylation sites in Human Proteinpedia. In addition to the protein phosphorylation databases, there are

W. Yang et al.

several bioinformatic tools available for the prediction of phosphorylation sites such as NetPhos (http:// www.cbs.dtu.dk/services/NetPhos) [137] and ScanSite (http://scansite.mit.edu) [138]. Although not a replacement for experimental validation, *in silico* prediction of putative PTM sites is a promising strategy to conduct preliminary analysis and can effectively direct subsequent MS experiments, as demonstrated in the multiple reaction monitoring strategy [139].

4.2 Detection of phosphoproteins

A classic method for detecting protein phosphorylation is radioactive ³²P labeling combined with 2-DE. Although this method is highly sensitive, it is not compatible with subsequent MS analysis because of safety issues associated with handling the material and contamination of instrumentation. For proteomic analysis of protein phosphorylation. western blotting and phospho-specific staining are the most frequently used methods to detect phosphorylated proteins. High-quality anti-pTyr antibodies are commercially available. with which tyrosine phosphorylated proteins can be detected with high sensitivity. However, anti-pSer/pThr antibodies that can universally detect Ser/Thr phosphorylation are still not available because existing antibodies are sensitive to amino acid sequence context. Recently, two universal phosphoprotein dve technologies termed Pro-Q Diamond [80] and Phos-tag [81] have been developed and commercialized by Molecular Probes (Eugene, OR, USA) and PerkinElmer (Waltham, MA, USA), respectively. These fluorescent dves can detect pSer-, pThr-, and pTyr-containing proteins directly in SDS-PAGE gels and 2-DE gels with high sensitivity, and are fully compatible with subsequent MS analysis.

4.3 Enrichment of phosphoproteins and phosphopeptides

Protein phosphorylation is usually present at low stoichiometry and difficult to detect in complex mixtures. Specific enrichment of phosphorylated proteins and/or peptides is generally required to facilitate subsequent MS analysis. Most commonly used enrichment methods are based on immunoreactivity, charge and chemical reactivity of phosphate groups. Anti-pTyr antibodies are widely used to purify tyrosine-phosphorylated proteins [140-143]. In contrast, antipSer/pThr antibodies are seldom used because these antibodies typically do not work in IP experiments [144]. IMAC, in which a resin is chelated with trivalent metal ions such as Fe(III) [145] or Ga(III) [146], is frequently used to enrich for phosphorylated peptides. The selectivity of this method has been further improved by methylesterification of acidic residues prior to IMAC enrichment [147, 148]. An alternative to IMAC involves enrichment methods based on metal oxides/ hydroxides such as titanium dioxide [149, 150], aluminium hydroxide [151], or zirconium dioxide [152], which can be used for phosphopeptide purification with high specificity. Based on the finding that most phosphopeptides carry a much lower net solution charge state than most tryptic peptides do, strong cation exchange chromatography has also been successfully applied to enrich phosphopeptides [153]. In addition to the above-mentioned methods, chemical tagging strategies have attracted a lot of interest. In most of these approaches, phosphate groups in pSer/pThr residues are removed by β -elimination under alkaline conditions. The resulting double bonds can react with various nucleophilic compounds by a Michael-type addition, which allows subsequent isolation of phosphopeptides [154] or phospho-specific proteolysis of phosphoproteins [155]. These β-elimination-based strategies have several limitations including (i) the occurrence of numerous side reactions and (ii) the inability to label tyrosine phosphorylated proteins or peptides. Another issue is that unmodified serine or threonine residues may also be affected to some degree [156]. In addition to these β -elimination-based approaches, another type of chemical tagging strategy termed phosphoramidate chemistry has been developed to allow equal purification of phosphopeptides containing pSer, pThr or pTyr residues [157, 158]. Very recently, Aebersold and colleagues have used phosphoramidate chemistry, IMAC, and TiO₂ to enrich phosphopeptides on a large scale [159]. They concluded that these methods detect different, partially overlapping segments of the phosphoproteome, which implies that none of the methods can enrich phosphopeptides in an unbiased way and highlights the necessity of combining several strategies for a comprehensive analysis of protein phosphorylation.

4.4 MS analysis of phosphorylation

The most frequently used method for fragmentation of phosphopeptides is CID. Because the phosphoester linkage is more susceptible to cleavage by fragmentation than the peptide backbone, a characteristic product ion is usually generated when phosphopeptides are fragmented in the mass spectrometer (reviewed in [124]). The product ion gives rise to a specific peak in the product spectrum, which can be monitored to trigger MS/MS and MS³ sequencing of the candidate phosphopeptide (reviewed in [160]). Determination of phosphorylation sites in phosphopeptides containing only one candidate site is usually straightforward. However, precise annotation of phosphorylation sites in phosphopeptides containing two or more candidate sites can be problematic and requires manual validation. To relieve this bottleneck, Gygi and colleagues have recently developed computational techniques for automatic assignment of protein phosphorylation sites [161].

Although CID is a very powerful method for phosphopeptide sequencing, it usually produces limited or weak fragment ions spectra for certain phosphopeptides, especially large, multiply charged and/or multiply phosphorylated ones. Consequently, many false-negative and false-positive identifications are produced. To address this problem, alternative fragmentation techniques such as electron capture dissociation (ECD) [162] and electron transfer dissociation (ETD) [163] have been developed. In contrast to CID, ECD and ETD are gentler fragmentation techniques, which preferentially cleave along the peptide backbone while retaining the labile phosphate moiety. The obvious difference between ECD and ETD is that ECD uses free electrons while ETD employs radical anions (e.g. anthracene [163] or azobenzene [164]) to fragment peptides or proteins. Since its introduction in 1998, ECD has been demonstrated to be powerful for analysis of triply or higher charged phosphopeptides and is complementary to CID [165, 166]. However, to date, this type of fragmentation is almost exclusively available in combination with FT-ICR [163, 167], the most expensive type of MS instrumentation. To develop an ECDlike fragmentation method for use with a low-cost, widely accessible mass spectrometer, Hunt and co-workers introduced ETD in 2004 [163] and demonstrated that this methodology is well suited for characterization of large multiply charged phosphopeptides on a large scale [168]. Pandey and colleagues compared ETD versus CID for global phosphoproteomics analysis and concluded that ETD identifies more phosphopeptides than CID and that these two methods are highly complementary [169]. Obviously, ETD is a very promising approach for the analysis of phosphopeptides. The combination of ETD with CID will become a common platform for comprehensive analysis of protein phosphorylation in the near future.

5 Analysis of the spatial organization of signaling complexes

The identification of interaction partners and their PTMs is a milestone but not the endpoint of the characterization of a multiprotein complex. Rather than being randomly distributed, components in a functional signaling complex are highly organized spatially. In fact, inappropriate spatial distribution of signaling complexes has been linked to a variety of pathological conditions like cancer [15, 16].

Several well-established techniques are available to yield information on spatial organization of protein complexes, but each has its shortcomings. For example, X-ray crystallography and NMR spectroscopy can produce high resolution atomic data but these approaches require relatively large amounts (milligrams) of pure analyte in a particular crystalline or solution state. Moreover, it usually takes months or even years to generate a molecular structure by applying these methodologies. In contrast, electron microscopy only requires limited amounts (micrograms) of material and image acquisition and analysis take a relatively short time. However, in most cases the resolution available from electron microscopy is too low (>10 Å) to determine the precise location of specific subunits [170].

As an alternative, chemical cross-linking technology provides an attractive way to obtain low resolution interatomic distance information about a macromolecular structure. Recently, this technology has been combined with rapidly developing MS technology (see Fig. 4). In this new strategy, two functional groups (e.g. amines, sulfhydryls, carboxylic acids, etc.), which are located on the interacting surfaces of two interaction partners within a protein complex, are covalently linked by cross-linkers. The protein complex is digested by sequence specific enzymes (commonly trypsin) and the cross-linked peptides are analyzed by MS. Finally, the specific sites of interaction are determined and the architecture of the complex is deduced [171, 172]. The advantage of this strategy is that a broad range of cross-linking reagents with different specificity and lengths are commercially available and only minute amounts of material are required. However, MS analysis of the cross-linked peptides is hampered by the enormous complexity of the reaction mixtures. After cross-linking and proteolytic digestion, a variety of peptides exists in the reaction mixtures, including predominantly unmodified peptides, three types of singly modified peptides (i.e. dead-end, intra-, and inter-cross-linked peptides), multiply modified peptides, and nonspecifically



Figure 4. Schematic representation of the use of chemical crosslinking in combination with MS to determine the interacting sites of binding partners in signaling complexes. Four types of chemical cross-linkers are used in different steps to reduce the complexity of the reaction mixtures and facilitate analysis by MS.

modified peptides. Consequently, the identification of intercross-linked peptides is akin to looking for a needle in a haystack. Moreover, the most commonly used cross-linking reagents react with ε -amine groups of lysine residues, resulting in a high frequency of missed cleavages and a loss of positive charge. As a result, cross-linked peptides are commonly large and sometimes may not be detected by MS.

To reduce the complexity of the reaction mixtures and facilitate MS analysis, a variety of cross-linkers have recently been developed, which contain an affinity tag (e.g. biotin) [173], a fluorogenic group [174], a stable isotope label [175], a chemical- [176], photo- [177], or MS/MS-cleavable [178] site, or, more recently, a combination of these groups [179-181] (see Fig. 4). A comprehensive introduction to these crosslinkers is beyond the scope of this review. Interested readers are referred to a recently published review [172]. Moreover, powerful MS technologies such as FT-ICR can generate high resolution and high accuracy data and can fragment large peptides by using ECD or ETD, thus allowing for an unambiguous assignment of cross-linked products. In addition, more and more bioinformatic tools, such as CLPM [182], VIRTUALMSLAB [183] and Pro-CrossLink [184], have recently been developed to assign the cross-linked peptides in the mass spectra. Taking the rapid progress into consideration, MS analysis of cross-linked peptides of multiprotein complexes with high throughput is feasible. Continuous efforts in synthesizing novel cross-linking reagents, improving the resolution and accuracy of MS, and improvements in software will facilitate the application of this technology.

6 Reconstruction of signaling complexes

Signaling complexes are usually composed of tens or even hundreds of different proteins. For example, 186 proteins have been identified through a proteomic analysis of the NMDA receptor complexes [20]. This represents a formidable challenge: how does one predict the functions of signaling complexes based on the knowledge of the identities of the components and their PTMs? In other words, how to reconstruct a descriptive model that reflects the underlying biological processes as accurately as possible?

Functional annotation of the identified proteins is usually the first step for modeling multiprotein complexes. Gene Ontology (http://www.geneontology.org/) is a powerful tool for molecular function classification as well as biological process classification (reviewed in [185]). In addition, the InterPro (http://www.ebi.ac.uk/interpro/) database provides a wealth of information about domains, families, and other structural characteristics. Protein sequences can be submitted for automatic InterProScan analysis to retrieve such information. Moreover, the relationship between human genes and genetic disorders is sometimes available from Online Mendelian Inheritance in Man (OMIM; http:// www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM). Public and commercial databases such as HPRD (http://www.hprd.org) [135, 136] and HumanPSD (http://www.biobase-inter national.com/pages/index.php?ide=humanpsd) [186] collect expert-curated information for human proteins or proteins of human, mouse and rat origin, respectively, including their molecular functions and relationships to diseases. Nevertheless, the information retrieved from these databases is often not sufficient for functional annotation. Therefore, it is usually necessary to query the scientific literature, either by manual searching or text mining (reviewed in [187, 188]). To better interpret the relevance of biological annotations of a set of proteins, statistical analysis should be used. Likely functional players, such as kinases, phosphatases and adaptor/scaffold proteins in a signaling complex can be identified by MS sequence analysis and database query and/or structural comparisons with known proteins.

Protein-protein interaction information is also required for reconstruction of signaling complexes. Cross-linking in combination with MS can provide first-hand information of protein-protein interactions in a signaling complex. In addition, information about protein-protein interactions can be retrieved from a variety of protein-protein interaction databases such as BIND, DIP, HPRD, IntAct, MINT, MIPS, PDZBase, and Reactome (reviewed in [189]). Information about domain interactions can be retrieved from a recently constructed public database called InterPare [190]. Moreover, the scientific literature is also an important source for known protein-protein interactions.

To model the pathway(s) intersecting with a signaling complex, graphic visualization tools should be employed. There are a number of such programs available for visualizing molecular interaction networks, among which Cytoscape [191] is one of the most advanced. The highly connected proteins (mostly scaffold proteins) in a signaling complex are particularly important for network integrity. Knockout or mutation of such proteins usually has dramatic effects on cell and organ phenotype [4].

By integrating information from functional annotations, known protein-protein interactions, and known signaling pathways, a network model may be constructed. The model may be useful in making functional predictions about the complex, for example, about which pathways are activated or blocked following complex formation. Compared with static modeling of signaling complexes, dynamic and quantitative modeling of signaling complexes is much more challenging. A wealth of quantitative information must be integrated into a mathematical model. As to a large-scale analysis of signaling complexes, more powerful modeling strategies have to be developed for a highly integrated description of the network system.

7 Challenges ahead

Although a number of signaling complexes have been successfully identified and studied by the application of employing proteomic techniques, a comprehensive proteomic analysis of signaling complexes is currently unattainable due to several unfavorable properties of these multiprotein assemblies. First, signaling complexes are mostly low abundant. Second, many protein-protein interactions in signaling complexes are weak or transient. Third, their assembly and disassembly is highly dynamic and usually happens over a short time.

7.1 Low abundance of signaling complexes

In higher eukaryotes, signaling complexes are mainly present at very low levels. Typically, adequate amounts of material for MS identification are isolated from $>10^8$ cells. For example, Blagoev et al. used 2×10^8 HeLa cells to isolate EGFR complexes [62]. Luo et al. started with $\sim 1.5 \times 10^8$ Mv1Lu cells to purify Smad2 complexes [192]. Very recently, Vandermoere et al. isolated Akt complexes from MCF-7 cells, starting with 15 mg of total protein [25]. Although not a major problem for analyzing one or several signaling complexes under static conditions, low abundance can be a serious issue for characterizing dynamic signaling events or for large-scale analysis. In those cases, starting from $10^9 \sim 10^{10}$ cells is usually a prerequisite, which means hundreds of 75 cm² flasks of monolayer mammalian cells have to be cultured. To address this problem, ultra-sensitive proteomic approaches have to be developed and made routine.

7.2 Weak and transient protein-protein interactions

Many components of signal transduction networks interact with each other through weak or transient interactions. When using affinity purification methods, TAP in particular, to isolate signaling intermediates, weak and transient interactions are often lost. In vivo chemical cross-linking is an attractive approach to "cement" the weak and transient protein-protein interactions. However, this could lead to some false positive and false negative results. Cross-linkers combine any protein in the vicinity and thus proteins that are not in direct contact may be cross-linked. On the contrary, when functional groups (e.g. ɛ-amino group of lysine) are not positioned in the interacting surface, the two interacting proteins may not be cross-linked. Therefore, the type and concentration of cross-linker and reaction time must be optimized to reduce false positives and negatives. Additionally, as we discuss in Section 5, MS analysis of cross-linked peptides or proteins is still a challenge.

7.3 Dynamics of signaling complexes

The formation and breakdown of signaling complexes is highly dynamic. This has been directly observed by highresolution multicolor imaging [193]. However, this technology is only applicable to known proteins. Fortunately, multiplexed quantitative proteomic approaches have recently been developed to characterize the dynamics of cell signaling. For example, three isotopic forms of SILAC (with mass differences of +6 and +10 Da) have been used to investigate the early events of EGFR signaling in HeLa cells [194]. iTRAQ has been used to study the dynamic regulation of specific tyrosine phosphorylation sites in the EGFR signaling network [143]. All these multiplexed quantitative techniques can be used directly to study the dynamics of signaling complexes.

A critical issue in studying the dynamics of signaling complexes is that the assembly of signaling complexes is usually very fast. For example, stimulation of the T cell receptor results in the recruitment of many different proteins within 15 s of receptor activation [195]. Moreover, by tagging signaling proteins with variants of enhance green fluorescent protein, Bunnell et al. found that T cell receptor complexes can form within seconds of T cell receptor complexes engagement [193]. It is difficult for conventional techniques to deal with this short time scale with high resolution and robustness. Very recently, an automated, continuous quench-flow system has been developed to allow quantitative proteomic assessment of very early cellular signaling events, qPACE, with a resolution of 1 s [196]. Although the qPACE approach was only applied to quantitatively analyze the changes of the phosphorylation of EGFR, Shc (Src homologous and collagen-like protein) and PLC_Y (phospholipase C gamma 1), it can be readily applied to quantitatively investigate the formation of signaling complexes. Nevertheless, qPACE has not yet been commercialized.

8 Conclusions

The important concept of the "signaling complex" has emerged in the past two decades. The dynamic formation of signaling complexes plays a critical role in enhancing signaling efficiency, ensuring specificity and increasing sensitivity [14]. Powerful proteomic techniques, which can characterize proteins with high sensitivity, high confidence, and high throughput, have recently been employed in this area. In this review, we described the most commonly used approaches and potential and emerging approaches involving MS techniques. After separation and purification of signaling complexes, MS can identify specific protein components as well as characterize their PTMs. Especially in combination with chemical cross-linking, MS can determine many aspects of the architecture of large multiprotein assemblies. This information can then be used to make testable predictions about their biological function. We also discussed several challenges ahead of us before we can achieve comprehensive proteomic analysis of dynamic signaling complexes.

Every technology has its own strengths and shortcomings. To fully understand the formation, function and composition of signaling complexes, other technologies such as confocal microscopy, fluorescence resonance energy transfer microscopy, bioluminescence resonance energy transfer microscopy, isothermal titration calorimetry, site directed mutagenesis, mouse models with directed mutations (reviewed in [17]) have to be employed in combination with MS methods. Never-

W. Yang et al.

theless, the unique ability of proteomic techniques to identify proteins, characterize PTMs, and discover protein-protein interactions with high throughput ensures that proteomics will be a fundamental driver in the coming years in the signal transduction field. It is a fair bet that, with the continuous development of proteomic techniques, proteomic approaches will play a major role in the integration of biochemical information with higher-order biological systems.

The authors have declared no conflict of interest.

9 References

- [1] Hunter, T., Signaling–2000 and beyond. *Cell* 2000, *100*, 113–127.
- [2] Bray, D., Signaling complexes: Biophysical constraints on intracellular communication. Annu. Rev. Biophys. Biomol. Struct. 1998, 27, 59–75.
- [3] Tsunoda, S., Sierralta, J., Zuker, C. S., Specificity in signaling pathways: Assembly into multimolecular signaling complexes. *Curr. Opin. Genet. Dev.* 1998, *8*, 419–422.
- [4] Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., Lim, W. A., Domains, motifs, and scaffolds: The role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* 2006, *75*, 655–680.
- [5] Hoshi, N., Langeberg, L. K., Scott, J. D., Distinct enzyme combinations in AKAP signalling complexes permit functional diversity. *Nat. Cell Biol.* 2005, *7*, 1066–1073.
- [6] Kim, E., Cho, K. O., Rothschild, A., Sheng, M., Heteromultimerization and NMDA receptor-clustering activity of Chapsyn-110, a member of the PSD-95 family of proteins. *Neuron* 1996, *17*, 103–113.
- [7] Hsueh, Y. P., Kim, E., Sheng, M., Disulfide-linked head-tohead multimerization in the mechanism of ion channel clustering by PSD-95. *Neuron* 1997, *18*, 803–814.
- [8] Carnegie, G. K., Scott, J. D., A-kinase anchoring proteins and neuronal signaling mechanisms. *Genes Dev.* 2003, 17, 1557– 1568.
- [9] Wong, W., Scott, J. D., AKAP signalling complexes: Focal points in space and time. *Nat. Rev. Mol. Cell Biol.* 2004, *5*, 959–970.
- [10] Kagan, J. C., Medzhitov, R., Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling. *Cell* 2006, *125*, 943–955.
- [11] Fitzgerald, K. A., Chen, Z. J., Sorting out Toll signals. *Cell* 2006, *125*, 834–836.
- [12] Pawson, T., Dynamic control of signaling by modular adaptor proteins. *Curr. Opin. Cell Biol.* 2007, *19*, 112–116.
- [13] Cho, W., Building signaling complexes at the membrane. *Sci. STKE* 2006, *321*, pe7.
- [14] Zhang, M., Wang, W., Organization of signaling complexes by PDZ-domain scaffold proteins. Acc. Chem. Res. 2003, 36, 530–538.
- [15] Pawson, T., Warner, N., Oncogenic re-wiring of cellular signaling pathways. *Oncogene* 2007, *26*, 1268–1275.
- [16] Kuhn, S., Koch, M., Nubel, T., Ladwein, M. et al., A complex of EpCAM, claudin-7, CD44 variant isoforms, and tetra-

spanins promotes colorectal cancer progression. *Mol. Cancer Res.* 2007, *5*, 553–567.

- [17] Houtman, J. C., Barda-Saad, M., Samelson, L. E., Examining multiprotein signaling complexes from all angles. *FEBS J.* 2005, 272, 5426–5435.
- [18] Husi, H., Ward, M. A., Choudhary, J. S., Blackstock, W. P., Grant, S. G., Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. *Nat. Neurosci.* 2000, *3*, 661– 669.
- [19] Husi, H., Grant, S. G., Isolation of 2000-kDa complexes of Nmethyl-D-aspartate receptor and postsynaptic density 95 from mouse brain. J. Neurochem. 2001, 77, 281–291.
- [20] Collins, M. O., Husi, H., Yu, L., Brandon, J. M. et al., Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. J. Neurochem. 2006, 97, 16–23.
- [21] Ping, P., Zhang, J., Pierce, W. M., Jr., Bolli, R., Functional proteomic analysis of protein kinase C epsilon signaling complexes in the normal heart and during cardioprotection. *Circ. Res.* 2001, *88*, 59–62.
- [22] Edmondson, R. D., Vondriska, T. M., Biederman, K. J., Zhang, J. *et al.*, Protein kinase C epsilon signaling complexes include metabolism- and transcription/translationrelated proteins: complimentary separation techniques with LC/MS/MS. *Mol. Cell. Proteomics* 2002, *1*, 421–433.
- [23] Zhang, J., Baines, C. P., Zong, C., Cardwell, E. M. et al., Functional proteomic analysis of a three-tier PKCepsilon-Akt-eNOS signaling module in cardiac protection. Am. J. Physiol. Heart Circ. Physiol. 2005, 288, H954–H961.
- [24] Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P. O. et al., A physical and functional map of the human TNF-alpha/NFkappa B signal transduction pathway. *Nat. Cell Biol.* 2004, 6, 97–105.
- [25] Vandermoere, F., El Yazidi-Belkoura, I., Demont, Y., Slomianny, C. *et al.*, Proteomics exploration reveals that actin is a signaling target of the kinase Akt. *Mol. Cell. Proteomics* 2007, *6*, 114–124.
- [26] Huber, L. A., Pfaller, K., Vietor, I., Organelle proteomics: implications for subcellular fractionation in proteomics. *Circ. Res.* 2003, *92*, 962–968.
- [27] Taylor, S. W., Fahy, E., Ghosh, S. S., Global organellar proteomics. *Trends Biotechnol.* 2003, *21*, 82–88.
- [28] Brunet, S., Thibault, P., Gagnon, E., Kearney, P. *et al.*, Organelle proteomics: Looking at less to see more. *Trends Cell Biol.* 2003, *13*, 629–638.
- [29] Dreger, M., Subcellular proteomics. Mass Spectrom. Rev. 2003, 22, 27–56.
- [30] Warnock, D. E., Fahy, E., Taylor, S. W., Identification of protein associations in organelles, using mass spectrometrybased proteomics. *Mass Spectrom. Rev.* 2004, *23*, 259–280.
- [31] Andersen, J. S., Mann, M., Organellar proteomics: turning inventories into insights. *EMBO Rep.* 2006, 7, 874–879.
- [32] Rebois, R. V., Hebert, T. E., Protein complexes involved in heptahelical receptor-mediated signal transduction. *Receptors Channels* 2003, *9*, 169–194.
- [33] Simons, K., Toomre, D., Lipid rafts and signal transduction. Nat. Rev. Mol. Cell Biol. 2000, 1, 31–39.
- [34] Pike, L. J., Lipid rafts: heterogeneity on the high seas. Biochem. J. 2004, 378, 281–292.

- [35] Raimondo, F., Ceppi, P., Guidi, K., Masserini, M. *et al.*, Proteomics of plasma membrane microdomains. *Expert Rev. Proteomics* 2005, *2*, 793–807.
- [36] Scaffidi, C., Kischkel, F. C., Krammer, P. H., Peter, M. E., Analysis of the CD95 (APO-1/Fas) death-inducing signaling complex by high-resolution two-dimensional gel electrophoresis. *Methods Enzymol.* 2000, *322*, 363–373.
- [37] Farr, C. D., Gafken, P. R., Norbeck, A. D., Doneanu, C. E. et al., Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. J. Neurochem. 2004, 91, 438–450.
- [38] Munro, S., Pelham, H. R., Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp 70. *EMBO J.* 1984, *3*, 3087–3093.
- [39] Loewith, R., Jacinto, E., Wullschleger, S., Lorberg, A. et al., Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Mol. Cell* 2002, 10, 457–468.
- [40] Mayhew, M. W., Webb, D. J., Kovalenko, M., Whitmore, L. et al., Identification of protein networks associated with the PAK1-betaPIX-GIT1-paxillin signaling complex by mass spectrometry. J. Proteome Res. 2006, 5, 2417–2423.
- [41] Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D. et al., Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature 2002, 415, 180–183.
- [42] Ewing, R. M., Chu, P., Elisma, F., Li, H. et al., Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol. Syst. Biol.* 2007, *3*, 89.
- [43] Burack, W. R., Shaw, A. S., Signal transduction: hanging on a scaffold. *Curr. Opin. Cell Biol.* 2000, *12*, 211–216.
- [44] Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M. *et al.*, A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 1999, *17*, 1030–1032.
- [45] Forler, D., Kocher, T., Rode, M., Gentzel, M. et al., An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat. Biotechnol.* 2003, 21, 89–92.
- [46] Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T., Drewes, G., Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. J. Biol. Chem. 2004, 279, 12804–12811.
- [47] Rivas, S., Romeis, T., Jones, J. D., The Cf-9 disease resistance protein is present in an approximately 420-kilodalton heteromultimeric membrane-associated complex at one molecule per complex. *Plant Cell* 2002, *14*, 689–702.
- [48] Knuesel, M., Wan, Y., Xiao, Z., Holinger, E. *et al.*, Identification of novel protein-protein interactions using a versatile mammalian tandem affinity purification expression system. *Mol. Cell. Proteomics* 2003, *2*, 1225–1233.
- [49] Kaneko, A., Umeyama, T., Hanaoka, N., Monk, B. C. et al., Tandem affinity purification of the Candida albicans septin protein complex. Yeast 2004, 21, 1025–1033.
- [50] Graumann, J., Dunipace, L. A., Seol, J. H., McDonald, W. H. et al., Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. *Mol. Cell. Proteomics* 2004, *3*, 226–237.
- [51] Schimanski, B., Nguyen, T. N., Gunzl, A., Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot. Cell* 2005, 4, 1942–1950.

- [52] Honey, S., Schneider, B. L., Schieltz, D. M., Yates, J. R., Futcher, B., A novel multiple affinity purification tag and its use in identification of proteins associated with a cyclin-CDK complex. *Nucleic Acids Res.* 2001, *29*, E24.
- [53] Auty, R., Steen, H., Myers, L. C., Persinger, J. et al., Purification of active TFIID from Saccharomyces cerevisiae. Extensive promoter contacts and co-activator function. J. Biol. Chem. 2004, 279, 49973–49981.
- [54] Drakas, R., Prisco, M., Baserga, R., A modified tandem affinity purification tag technique for the purification of protein complexes in mammalian cells. *Proteomics* 2005, *5*, 132– 137.
- [55] Tagwerker, C., Flick, K., Cui, M., Guerrero, C. *et al.*, A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with *in vivo* cross-linking. *Mol. Cell. Proteomics* 2006, *5*, 737–748.
- [56] Burckstummer, T., Bennett, K. L., Preradovic, A., Schutze, G. et al., An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat. Methods* 2006, *3*, 1013–1019.
- [57] Chang, I. F., Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes. *Proteomics* 2006, *6*, 6158–6166.
- [58] Gavin, A. C., Bosche, M., Krause, R., Grandi, P. et al., Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 2002, 415, 141–147.
- [59] Gavin, A. C., Aloy, P., Grandi, P., Krause, R. *et al.*, Proteome survey reveals modularity of the yeast cell machinery. *Nature* 2006, 440, 631–636.
- [60] Gottschalk, A., Almedom, R. B., Schedletzky, T., Anderson, S. D. et al., Identification and characterization of novel nicotinic receptor-associated proteins in *Caenorhabditis elegans. EMBO J.* 2005, 24, 2566–2578.
- [61] Angrand, P. O., Segura, I., Volkel, P., Ghidelli, S. *et al.*, Transgenic mouse proteomics identifies new 14-3-3-associated proteins involved in cytoskeletal rearrangements and cell signaling. *Mol. Cell. Proteomics* 2006, *5*, 2211–2227.
- [62] Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M. *et al.*, A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* 2003, *21*, 315–318.
- [63] Cross, M., Nguyen, T., Bogdanoska, V., Reynolds, E., Hamilton, J. A., A proteomics strategy for the enrichment of receptor-associated complexes. *Proteomics* 2005, *5*, 4754– 4763.
- [64] Sheng, M., Lee, S. H., Growth of the NMDA receptor industrial complex. *Nat. Neurosci.* 2000, *3*, 633–635.
- [65] Vondriska, T. M., Ping, P., Multiprotein signaling complexes and regulation of cardiac phenotype. *J. Mol. Cell Cardiol.* 2003, *35*, 1027–1033.
- [66] Hedman, E., Widen, C., Asadi, A., Dinnetz, I. *et al.*, Proteomic identification of glucocorticoid receptor interacting proteins. *Proteomics* 2006, *6*, 3114–3126.
- [67] Schagger, H., von Jagow, G., Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* 1991, *199*, 223–231.
- [68] Wittig, I., Braun, H. P., Schagger, H., Blue native PAGE. Nat. Protoc. 2006, 1, 418–428.
- [69] Krause, F., Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel

electrophoresis: (Membrane) protein complexes and supercomplexes. *Electrophoresis* 2006, *27*, 2759–2781.

- [70] Eden, S., Rohatgi, R., Podtelejnikov, A. V., Mann, M., Kirschner, M. W., Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck. *Nature* 2002, *418*, 790– 793.
- [71] Dudkina, N. V., Eubel, H., Keegstra, W., Boekema, E. J., Braun, H. P., Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 3225–3229.
- [72] Hartman, N. T., Sicilia, F., Lilley, K. S., Dupree, P., Proteomic complex detection using sedimentation. *Anal. Chem.* 2007, 79, 2078–2083.
- [73] Schagger, H., Cramer, W. A., von Jagow, G., Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* 1994, *217*, 220–230.
- [74] Suh, M. H., Ye, P., Datta, A. B., Zhang, M., Fu, J., An agaroseacrylamide composite native gel system suitable for separating ultra-large protein complexes. *Anal. Biochem.* 2005, 343, 166–175.
- [75] Rabilloud, T., Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains. *Proteomics* 2002, *2*, 3–10.
- [76] Gorg, A., Weiss, W., Dunn, M. J., Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 2004, 4, 3665–3685.
- [77] Wittmann-Liebold, B., Graack, H. R., Pohl, T., Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry. *Proteomics* 2006, *6*, 4688–4703.
- [78] Steinberg, T. H., Pretty On Top, K., Berggren, K. N., Kemper, C. et al., Rapid and simple single nanogram detection of glycoproteins in polyacrylamide gels and on electroblots. *Proteomics* 2001, 1, 841–855.
- [79] Hart, C., Schulenberg, B., Steinberg, T. H., Leung, W. Y., Patton, W. F., Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate Schiff-base stain. *Electrophoresis* 2003, 24, 588–598.
- [80] Steinberg, T. H., Agnew, B. J., Gee, K. R., Leung, W. Y. et al., Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology. *Proteomics* 2003, *3*, 1128– 1144.
- [81] Kinoshita, E., Takahashi, M., Takeda, H., Shiro, M., Koike, T., Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc(II) complex. *Dalton Trans.* 2004, 1189–1193.
- [82] Medzihradszky, K. F., Campbell, J. M., Baldwin, M. A., Falick, A. M. et al., The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. Anal. Chem. 2000, 72, 552–558.
- [83] Link, A. J., Eng, J., Schieltz, D. M., Carmack, E. *et al.*, Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* 1999, *17*, 676–682.
- [84] Sato, S., Tomomori-Sato, C., Parmely, T. J., Florens, L. et al., A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol. Cell* 2004, 14, 685–691.

- [85] Carrozza, M. J., Florens, L., Swanson, S. K., Shia, W. J. et al., Stable incorporation of sequence specific repressors Ash1 and Ume6 into the Rpd3L complex. *Biochim. Biophys. Acta* 2005, 1731, 77–87.
- [86] Regonesi, M. E., Del Favero, M., Basilico, F., Briani, F. et al., Analysis of the Escherichia coli RNA degradosome composition by a proteomic approach. Biochimie 2006, 88, 151– 161.
- [87] Koch, H. B., Zhang, R., Verdoodt, B., Bailey, A. *et al.*, Largescale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. *Cell Cycle* 2007, *6*, 205– 217.
- [88] Schirle, M., Heurtier, M. A., Kuster, B., Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol. Cell. Proteomics* 2003, *2*, 1297–1305.
- [89] Chen, E. I., Cociorva, D., Norris, J. L., Yates, J. R., 3rd, Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. J. Proteome Res. 2007, 6, 2529–2538.
- [90] Selbach, M., Mann, M., Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat. Methods* 2006, *3*, 981–983.
- [91] Kuhner, S., Gavin, A. C., Towards quantitative analysis of proteome dynamics. *Nat. Biotechnol.* 2007, 25, 298–300.
- [92] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B. et al., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 2002, 1, 376–386.
- [93] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F. *et al.*, Quantitative analysis of complex protein mixtures using isotopecoded affinity tags. *Nat. Biotechnol.* 1999, *17*, 994–999.
- [94] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B. et al., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol. Cell. Proteomics 2004, 3, 1154–1169.
- [95] Unlu, M., Morgan, M. E., Minden, J. S., Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997, *18*, 2071–2077.
- [96] Molina, H., Parmigiani, G., Pandey, A., Assessing reproducibility of a protein dynamics study using in vivo labeling and liquid chromatography tandem mass spectrometry. *Anal. Chem.* 2005, *77*, 2739–2744.
- [97] Molloy, M. P., Donohoe, S., Brzezinski, E. E., Kilby, G. W. et al., Large-scale evaluation of quantitative reproducibility and proteome coverage using acid cleavable isotope coded affinity tag mass spectrometry for proteomic profiling. *Proteomics* 2005, 5, 1204–1208.
- [98] Choe, L. H., Aggarwal, K., Franck, Z., Lee, K. H., A comparison of the consistency of proteome quantitation using twodimensional electrophoresis and shotgun isobaric tagging in *Escherichia coli* cells. *Electrophoresis* 2005, *26*, 2437– 2449.
- [99] Qian, W. J., Monroe, M. E., Liu, T., Jacobs, J. M. et al., Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using 160/180 labeling and the accurate mass and time tag approach. *Mol. Cell. Proteomics* 2005, 4, 700–709.

- [100] Kolkman, A., Dirksen, E. H., Slijper, M., Heck, A. J., Double standards in quantitative proteomics: direct comparative assessment of difference in gel electrophoresis and metabolic stable isotope labeling. *Mol. Cell. Proteomics* 2005, *4*, 255–266.
- [101] Wu, W. W., Wang, G., Baek, S. J., Shen, R. F., Comparative study of three proteomic quantitative methods, DIGE, clCAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. J. Proteome Res. 2006, 5, 651–658.
- [102] Chong, P. K., Gan, C. S., Pham, T. K., Wright, P. C., Isobaric tags for relative and absolute quantitation (iTRAQ) reproducibility: Implication of multiple injections. *J. Proteome Res.* 2006, *5*, 1232–1240.
- [103] Julka, S., Regnier, F. E., Recent advancements in differential proteomics based on stable isotope coding. *Brief. Funct. Genomic. Proteomic.* 2005, 4, 158–177.
- [104] Ong, S. E., Mann, M., Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 2005, 1, 252–262.
- [105] Unwin, R. D., Evans, C. A., Whetton, A. D., Relative quantification in proteomics: new approaches for biochemistry. *Trends Biochem. Sci.* 2006, *31*, 473–484.
- [106] Roe, M. R., Griffin, T. J., Gel-free mass spectrometry-based high throughput proteomics: tools for studying biological response of proteins and proteomes. *Proteomics* 2006, *6*, 4678–4687.
- [107] Leitner, A., Lindner, W., Chemistry meets proteomics: the use of chemical tagging reactions for MS-based proteomics. *Proteomics* 2006, *6*, 5418–5434.
- [108] Ishihama, Y., Sato, T., Tabata, T., Miyamoto, N. et al., Quantitative mouse brain proteomics using culturederived isotope tags as internal standards. *Nat. Biotechnol.* 2005, 23, 617–621.
- [109] Mann, M., Functional and quantitative proteomics using SILAC. Nat. Rev. Mol. Cell Biol. 2006, 7, 952–958.
- [110] Zhou, H., Ranish, J. A., Watts, J. D., Aebersold, R., Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat. Biotechnol.* 2002, *20*, 512–515.
- [111] Hansen, K. C., Schmitt-Ulms, G., Chalkley, R. J., Hirsch, J. et al., Mass spectrometric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and multidimensional chromatography. *Mol. Cell. Proteomics* 2003, 2, 299–314.
- [112] Lu, Y., Bottari, P., Turecek, F., Aebersold, R., Gelb, M. H., Absolute quantification of specific proteins in complex mixtures using visible isotope-coded affinity tags. *Anal. Chem.* 2004, *76*, 4104–4111.
- [113] Pierce, A., Unwin, R. D., Evans, C. A., Griffiths, S. et al., Eight-channel iTRAQ enables comparison of the activity of 6 leukaemogenic tyrosine kinases. *Mol. Cell. Proteomics* 2007, Epub.
- [114] Viswanathan, S., Unlu, M., Minden, J. S., Two-dimensional difference gel electrophoresis. *Nat. Protoc.* 2006, *1*, 1351– 1358.
- [115] Ranish, J. A., Yi, E. C., Leslie, D. M., Purvine, S. O. *et al.*, The study of macromolecular complexes by quantitative proteomics. *Nat. Genet*, 2003, *33*, 349–355.
- [116] Schulze, W. X., Mann, M., A novel proteomic screen for peptide-protein interactions. J. Biol. Chem. 2004, 279, 10756–10764.
- [117] Wang, T., Gu, S., Ronni, T., Du, Y. C., Chen, X., In vivo dualtagging proteomic approach in studying signaling path-

© 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ways in immune response. J. Proteome Res. 2005, 4, 941–949.

- [118] Foster, L. J., Rudich, A., Talior, I., Patel, N. et al., Insulin-dependent interactions of proteins with GLUT4 revealed through stable isotope labeling by amino acids in cell culture (SILAC). J. Proteome Res. 2006, 5, 64–75.
- [119] Zieske, L. R., A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J. Exp. Bot. 2006, 57, 1501–1508.
- [120] Barr, J. R., Maggio, V. L., Patterson, D. G., Jr., Cooper, G. R. et al., Isotope dilution-mass spectrometric quantification of specific proteins: Model application with apolipoprotein A-I. Clin. Chem. 1996, 42, 1676–1682.
- [121] Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., Gygi, S. P., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 6940–6945.
- [122] Beynon, R. J., Doherty, M. K., Pratt, J. M., Gaskell, S. J., Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat. Methods* 2005, *2*, 587–589.
- [123] Kito, K., Ota, K., Fujita, T., Ito, T., A synthetic protein approach toward accurate mass spectrometric quantification of component stoichiometry of multiprotein complexes. J. Proteome Res. 2007, 6, 792–800.
- [124] Spickett, C. M., Pitt, A. R., Morrice, N., Kolch, W., Proteomic analysis of phosphorylation, oxidation and nitrosylation in signal transduction. *Biochim. Biophys. Acta* 2006, *1764*, 1823–1841.
- [125] Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J. *et al.*, The utility of ETD mass spectrometry in proteomic analysis. *Biochim. Biophys. Acta* 2006, *1764*, 1811–1822.
- [126] Gesellchen, F., Bertinetti, O., Herberg, F. W., Analysis of posttranslational modifications exemplified using protein kinase A. *Biochim. Biophys. Acta* 2006, *1764*, 1788–1800.
- [127] Morelle, W., Canis, K., Chirat, F., Faid, V., Michalski, J. C., The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* 2006, *6*, 3993–4015.
- [128] Larsen, M. R., Trelle, M. B., Thingholm, T. E., Jensen, O. N., Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *Biotechniques* 2006, 40, 790– 798.
- [129] Kettenhofen, N. J., Broniowska, K. A., Keszler, A., Zhang, Y., Hogg, N., Proteomic methods for analysis of S-nitrosation. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2007, 851, 152–159.
- [130] Gianazza, E., Crawford, J., Miller, I., Detecting oxidative post-translational modifications in proteins. *Amino Acids* 2007, 33, 51–56.
- [131] Reinders, J., Sickmann, A., Modificomics: Posttranslational modifications beyond protein phosphorylation and glycosylation. *Biomol. Eng.* 2007, 24, 169–177.
- [132] Hornbeck, P. V., Chabra, I., Kornhauser, J. M., Skrzypek, E., Zhang, B., PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* 2004, *4*, 1551–1561.
- [133] Diella, F., Cameron, S., Gemund, C., Linding, R. et al., Phospho.ELM: A database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* 2004, *5*, 79.

- [134] Lee, T. Y., Huang, H. D., Hung, J. H., Huang, H. Y. et al., dbPTM: an information repository of protein post-translational modification. *Nucleic Acids Res.* 2006, 34, D622– D627.
- [135] Peri, S., Navarro, J. D., Amanchy, R., Kristiansen, T. Z. et al., Development of human protein reference database as an initial platform for approaching systems biology in humans. Genome Res. 2003, 13, 2363–2371.
- [136] Mishra, G. R., Suresh, M., Kumaran, K., Kannabiran, N. et al., Human protein reference database–2006 update. *Nucleic Acids Res.* 2006, 34, D411–D414.
- [137] Blom, N., Gammeltoft, S., Brunak, S., Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 1999, 294, 1351–1362.
- [138] Obenauer, J. C., Cantley, L. C., Yaffe, M. B., Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* 2003, *31*, 3635–3641.
- [139] Unwin, R. D., Griffiths, J. R., Leverentz, M. K., Grallert, A. et al., Multiple reaction monitoring to identify sites of protein phosphorylation with high sensitivity. *Mol. Cell. Proteomics* 2005, 4, 1134–1144.
- [140] Pandey, A., Podtelejnikov, A. V., Blagoev, B., Bustelo, X. R. et al., 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. USA* 2000, *97*, 179–184.
- [141] Steen, H., Kuster, B., Fernandez, M., Pandey, A., Mann, M., Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway. *J. Biol. Chem.* 2002, 277, 1031–1039.
- [142] Rush, J., Moritz, A., Lee, K. A., Guo, A. *et al.*, Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* 2005, *23*, 94–101.
- [143] Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J. et al., Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Prote*omics 2005, 4, 1240–1250.
- [144] Gronborg, M., Kristiansen, T. Z., Stensballe, A., Andersen, J. S. et al., A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol. Cell. Proteomics* 2002, *1*, 517–527.
- [145] Porath, J., Carlsson, J., Olsson, I., Belfrage, G., Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 1975, *258*, 598–599.
- [146] Posewitz, M. C., Tempst, P., Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* 1999, 71, 2883–2892.
- [147] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J. et al., Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. Nat. Biotechnol. 2002, 20, 301–305.
- [148] Haydon, C. E., Eyers, P. A., Aveline-Wolf, L. D., Resing, K. A. et al., Identification of novel phosphorylation sites on Xenopus laevis Aurora A and analysis of phosphopeptide enrichment by immobilized metal-affinity chromatography. Mol. Cell. Proteomics 2003, 2, 1055–1067.

- [149] Pinkse, M. W., Uitto, P. M., Hilhorst, M. J., Ooms, B., Heck, A. J., Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal. Chem.* 2004, *76*, 3935–3943.
- [150] Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., Jorgensen, T. J., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* 2005, *4*, 873– 886.
- [151] Wolschin, F., Wienkoop, S., Weckwerth, W., Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC). *Proteomics* 2005, *5*, 4389–4397.
- [152] Kweon, H. K., Hakansson, K., Selective zirconium dioxidebased enrichment of phosphorylated peptides for mass spectrometric analysis. *Anal. Chem.* 2006, 78, 1743–1749.
- [153] Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E. *et al.*, Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 12130–12135.
- [154] Oda, Y., Nagasu, T., Chait, B. T., Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* 2001, *19*, 379–382.
- [155] Knight, Z. A., Schilling, B., Row, R. H., Kenski, D. M. et al., Phosphospecific proteolysis for mapping sites of protein phosphorylation. Nat. Biotechnol. 2003, 21, 1047–1054.
- [156] McLachlin, D. T., Chait, B. T., Improved beta-eliminationbased affinity purification strategy for enrichment of phosphopeptides. *Anal. Chem.* 2003, *75*, 6826–6836.
- [157] Zhou, H., Watts, J. D., Aebersold, R., A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* 2001, *19*, 375–378.
- [158] Tao, W. A., Wollscheid, B., O'Brien, R., Eng, J. K. et al., Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. Nat. Methods 2005, 2, 591–598.
- [159] Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., Aebersold, R., Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* 2007, *4*, 231–237.
- [160] Collins, M. O., Yu, L., Choudhary, J. S., Analysis of protein phosphorylation on a proteome-scale. *Proteomics* 2007, 7, 2751–2768.
- [161] 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. *Nat. Biotechnol.* 2006, *24*, 1285–1292.
- [162] Zubarev, R. A., Kelleher, N. L., McLafferty, F. W., Electron capture dissociation of multiply charged protein cations. A nonergodic process. J. Am. Chem. Soc. 1998, 120, 3265– 3266.
- [163] Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 9528–9533.
- [164] Han, H., Xia, Y., McLuckey, S. A., Ion trap collisional activation of c and z* ions formed via gas-phase ion/ion electrontransfer dissociation. J. Proteome Res. 2007, 6, 3062–3069.
- [165] Stingl, C., Ihling, C., Ammerer, G., Sinz, A., Mechtler, K., Application of different fragmentation techniques for the

analysis of phosphopeptides using a hybrid linear ion trap-FTICR mass spectrometer. *Biochim. Biophys. Acta* 2006, *1764*, 1842–1852.

- [166] Sweet, S. M., Cooper, H. J., Electron capture dissociation in the analysis of protein phosphorylation. *Expert Rev. Prote*omics 2007, 4, 149–159.
- [167] Ding, L., Brancia, F. L., Electron capture dissociation in a digital ion trap mass spectrometer. *Anal. Chem.* 2006, 78, 1995–2000.
- [168] Chi, A., Huttenhower, C., Geer, L. Y., Coon, J. J. et al., Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. Proc. Natl. Acad. Sci. USA 2007, 104, 2193–2198.
- [169] Molina, H., Horn, D. M., Tang, N., Mathivanan, S., Pandey, A., Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 2199–2204.
- [170] Rubinstein, J. L., Structural analysis of membrane protein complexes by single particle electron microscopy. *Meth*ods 2007, 41, 409–416.
- [171] Rappsilber, J., Siniossoglou, S., Hurt, E. C., Mann, M., A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. *Anal. Chem.* 2000, *72*, 267–275.
- [172] Sinz, A., Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and proteinprotein interactions. *Mass Spectrom. Rev.* 2006, *25*, 663– 682.
- [173] Alley, S. C., Ishmael, F. T., Jones, A. D., Benkovic, S. J., Mapping protein-protein interactions in the bacteriophage T4 DNA polymerase holoenzyme using a novel trifunctional photo-cross-linking and affinity reagent. J. Am. Chem. Soc. 2000, 122, 6126–6127.
- [174] Sinz, A., Wang, K., Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: application to nebulin-calmodulin complexes. *Biochemistry* (*Mosc.*) 2001, 40, 7903–7913.
- [175] Muller, D. R., Schindler, P., Towbin, H., Wirth, U. *et al.*, Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* 2001, *73*, 1927–1934.
- [176] Bennett, K. L., Kussmann, M., Bjork, P., Godzwon, M. et al., Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping–a novel approach to assess intermolecular protein contacts. Protein Sci. 2000, 9, 1503–1518.
- [177] Marriott, G., Ottl, J., Synthesis and applications of heterobifunctional photocleavable cross-linking reagents. *Methods Enzymol.* 1998, 291, 155–175.
- [178] Back, J. W., Hartog, A. F., Dekker, H. L., Muijsers, A. O. et al., A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes. J. Am. Soc. Mass Spectrom. 2001, 12, 222–227.
- [179] Tang, X., Munske, G. R., Siems, W. F., Bruce, J. E., Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. *Anal. Chem.* 2005, 77, 311– 318.
- [180] Chowdhury, S. M., Munske, G. R., Tang, X., Bruce, J. E., Collisionally activated dissociation and electron capture

Cell Biology 851

dissociation of several mass spectrometry-identifiable chemical cross-linkers. *Anal. Chem.* 2006, *78*, 8183–8193.

- [181] Chu, F., Mahrus, S., Craik, C. S., Burlingame, A. L., Isotopecoded and affinity-tagged cross-linking (ICATXL): an efficient strategy to probe protein interaction surfaces. J. Am. Chem. Soc. 2006, 128, 10362–10363.
- [182] Tang, Y., Chen, Y., Lichti, C. F., Hall, R. A. *et al.*, CLPM: a cross-linked peptide mapping algorithm for mass spectrometric analysis. *BMC Bioinformatics* 2005, *6*, S9.
- [183] de Koning, L. J., Kasper, P. T., Back, J. W., Nessen, M. A. *et al.*, Computer-assisted mass spectrometric analysis of naturally occurring and artificially introduced cross-links in proteins and protein complexes. *FEBS J.* 2006, *273*, 281–291.
- [184] Gao, Q., Xue, S., Doneanu, C. E., Shaffer, S. A. *et al.*, Pro-CrossLink. Software tool for protein cross-linking and mass spectrometry. *Anal. Chem.* 2006, *78*, 2145–2149.
- [185] Beissbarth, T., Interpreting experimental results using gene ontologies. *Methods Enzymol.* 2006, 411, 340–352.
- [186] Hodges, P. E., Carrico, P. M., Hogan, J. D., O'Neill, K. E. et al., Annotating the human proteome: the Human Proteome Survey Database (HumanPSD) and an in-depth target database for G protein-coupled receptors (GPCR-PD) from Incyte Genomics. *Nucleic Acids Res.* 2002, 30, 137–141.
- [187] Erhardt, R. A., Schneider, R., Blaschke, C., Status of textmining techniques applied to biomedical text. *Drug Discov. Today* 2006, *11*, 315–325.
- [188] Ananiadou, S., Kell, D. B., Tsujii, J., Text mining and its potential applications in systems biology. *Trends Biotechnol.* 2006, 24, 571–579.
- [189] Mathivanan, S., Periaswamy, B., Gandhi, T. K., Kandasamy, K. *et al.*, An evaluation of human protein-protein interaction data in the public domain. *BMC Bioinformatics* 2006, 7, S19.
- [190] Gong, S., Park, C., Choi, H., Ko, J. et al., A protein domain interaction interface database: InterPare. BMC Bioinformatics 2005, 6, 207.
- [191] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S. et al., Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003, 13, 2498–2504.
- [192] Luo, O., Nieves, E., Kzhyshkowska, J., Angeletti, R. H., Endogenous transforming growth factor-beta receptor-mediated Smad signaling complexes analyzed by mass spectrometry. *Mol. Cell. Proteomics* 2006, *5*, 1245–1260.
- [193] Bunnell, S. C., Hong, D. I., Kardon, J. R., Yamazaki, T. et al., T cell receptor ligation induces the formation of dynamically regulated signaling assemblies. J. Cell Biol. 2002, 158, 1263–1275.
- [194] Blagoev, B., Ong, S. E., Kratchmarova, I., Mann, M., Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* 2004, *22*, 1139–1145.
- [195] Pacini, S., Valensin, S., Telford, J. L., Ladbury, J., Baldari, C. T., Temporally regulated assembly of a dynamic signaling complex associated with the activated TCR. *Eur. J. Immunol.* 2000, *30*, 2620–2631.
- [196] Dengjel, J., Akimov, V., Olsen, J. V., Bunkenborg, J. et al., Quantitative proteomic assessment of very early cellular signaling events. *Nat. Biotechnol.* 2007, 25, 566–568.