Published in final edited form as:

Adv Enzyme Regul. 2011 ; 51(1): 293–305. doi:10.1016/j.advenzreg.2010.10.005.

Receptor Tyrosine Kinase Signaling – A Proteomic Perspective

Jordane Biarc, Robert J. Chalkley, A. L. Burlingame, and Ralph A. Bradshaw

Dept of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco CA 94158 USA

Abstract

The stimulation of various cellular processes through extracellular signals is of paramount importance in biological systems and is a central focus in the diagnosis, treatment and prevention of disease. The information transfer is accomplished in a variety of ways by the interaction of soluble, matrix-associated and cell bound ligands that either bind specifically to plasma membrane-associated proteins that act as receptors, or penetrate to the cytoplasmic/nuclear compartments to bind and activate receptors located there. The former class of entities generates intracellular signals that are transmitted and amplified by chemical modifications that are manifested as protein post-translational modifications (PTMs). These are both reversible and irreversible and range from phosphorylation of tyrosine, threonine and serine residues to endoproteolytic cleavages. Although the PTMs alter the activity and functions of many of the proteins in these cascades, the major outcomes of most of the signaling pathways are the activation/deactivation of transcriptional regulators with the concomitant changes in gene expression that generally underlie biological responses.

Among the several types of cell surface receptors involved in signal transduction, the receptor tyrosine kinases (RTKs) are a large and important group (Schlessinger 2000). They are organized with an extracellular domain devoted to ligand recognition and binding, a single transmembrane segment, usually of 20–24 amino acids in length and quite hydrophobic in character that exists as an alpha helix, and an intracellular (or endo-) domain that contains a tyrosine kinase that is flanked by variously sized juxtamembrane regions and C-terminal extensions. With a couple of exceptions, these receptors phosphorylate themselves autocatalytically upon ligand binding that in turn provide phosphotyrosine docking sites for effectors, adaptors and scaffold proteins (Pawson 2002). These then perpetrate the signal in various ways. The initial events involve only the phosphorylation of tyrosine residues, but the amplification processes mostly involve serine/threonine phosphorylations. These downstream events also are marked by other types of modification, such as acetylation, methylation and O-GlcNAcylation. The overall flow of information is vectorially directed to the nuclear compartment, where the transcriptional alterations and resulting changes in gene expression occur.

Proteomic analyses have been applied to several aspects of cellular signaling, including the characterization of protein-protein interactions (both constitutive and induced), the identification of PTMs introduced in the various adaptor/effector participants, and elucidation of the overall networks that control and distribute the flow of information imparted by the initial signal. Coupled with expression analyses, these studies have provided

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

substantially increased knowledge of these crucial biological processes at the molecular level and are illustrated in the examples discussed in the following sections.

An Introduction to Proteomic Analyses

Proteomics, derived from the term coined by Marc Wilkins in 1993 (Cohen 2001; Huber 2003), is generally described as the study of the protein component of an organism, cell or tissue. In the global sense, this means the identification and characterization of all the proteins present in a germane sample, including all of their variants, modifications and interactions. It would also require the complete description of the three dimensional structure of these entities and all of their interactions with small and macromolecules. As living systems are dynamic, this would necessarily need to include determining the changes in all of these parameters as a function of time (Bradshaw and Burlingame 2005). Such an analysis is well beyond the scope of the presently available technology and, before it will be achieved, it will require substantive advances, not only in proteomic methodology, but also in the techniques and approaches of the related sciences of genomics, transcriptomics and metabolomics, and also in bioinformatics, the computational component that ties all of the 'omic sciences together.

The concept of considering the entire complement of proteins from a single source as an entity in its own right, as opposed to studying individual proteins in isolation (the so-called reductionist approach), can arguably be dated from the introduction of 2D gel analyses (Klose 1975; O'Farrell 1975). This methodology allowed investigators to look for changes in an array that allowed the visualization of several hundred or more proteins and showed relative differences in expression. Its value was limited by the resolution and the difficulty in accurately identifying the proteins in the interesting spots. The 2D gel electrophoresis method eventually gave rise to other arrays, both protein and nucleic acid-based, that materially advanced the usefulness of this type of analysis (MacBeath 2002; Sreekumar and Chinnaiyan 2002; Schweitzer *et al.* 2003). Independently, highly significant advances in mass spectrometry (MS) were developed that made the application of this very powerful analytical method to biological materials practical. The invention of FAB (LSIMS) (Barber *et al.* 1981; Aberth *et al.* 1982), MALDI (Karas and Hillencomp 1988; Tanaka *et al.* 1988) and ESI (Fenn *et al.* 1989) as three radical new ways to inject protonated or deprotonated polar molecules into the gas phase for analysis by mass spectrometry rapidly made this the dominant approach for proteomic experimentation (Aebersold and Mann 2003). However, the value of MS was not fully manifested until the human genome sequence was deduced and, through the technological developments in nucleic acid sequencing that were accomplished as a result, the avalanche of genomic data that followed. The wedding of these two technological breakthroughs was a true paradigm shift and can be taken as the beginning of the modern proteomic era (Bradshaw and Burlingame 2010).

Table 1 illustrates the active areas of proteomic research and all are applicable to some degree to signaling systems. These have lead to very significant changes in our understanding of the function and organization of cellular activity at the molecular level. Overall, there have been significant gains in identifying proteins in complex samples with only a minimum amount of prefractionation, and in elucidating PTMs at unprecedented scales (Table 2). One of the most impressive achievements has been the determination of the extent to which protein-protein interactions occur in the highly complex networks that underlie the regulation of virtually all cellular activity (Collins *et al.* 2007). In addition, a subset of these interactions describes various protein machines, such as the nuclear pore, the proteosome or various transcriptional complexes. The combination of mass spectrometry with several of the tools of structural biology, including X-ray crystallography, 2D NMR and electron microscopy, have deduced the detailed architecture of these structures resulting

in significant understanding of their function and regulation (Alber *et al.* 2007). Many of these advances have come about through the development of improved methodology such as faster, more sensitive mass spectrometers and new peptide fragmentation methods inside the mass spectrometer, new chromatographic materials, and an extended array of useful reagents (in particular antibodies). It is noteworthy that all of these gains have had translational applications that have provided new advances in the clinical arena.

There are however, many areas for which proteomics has yet to provide significant gains (Table 3). In the first place, approximately a third of the annotated human genome (~6-7000 genes) has never been observed in a proteomic experiment and the function of most of these putative proteins is unknown. Whether they are expressed in only very specialized situations or are not expressed at all remains to be determined. The Human Proteome Project (www.hupo.org/research/hpp/) is directed toward developing assays (based on MS analyses, using multiple or selective reaction monitoring; M- or SRM) of all of the potential human proteins along with suitable antibodies for their detection in germane tissues (www.proteinatlas.org/). The achievement of these goals will be an important step in the development of proteomics. A second problem area is the determination and the correct localization of PTMs, which is of major importance to transmembrane signaling experiments (Bradshaw *et al.* 2010). Although large datasets, often containing several thousand identifications, are becoming commonplace, particularly in phosphorylation studies, the quality of data is quite variable, and misidentification and/or mislocalization is prevalent. This materially damages the value of these analyses and is a cause for some concern as the information often finds its way into databases where it then develops a certain permanency. Greater awareness of this problem and the deposition of raw data are two countermeasures, but improved PTM search engines would also help. A third problem has been the failure, so far, to produce a significant number of biomarkers of clinical utility based on proteomic experimentation. As shown in Fig 1, the rate and number of diagnostic tests approved by the FDA over the past several years has been minimal and has not been particularly affected by the introduction of proteomic technologies (Anderson 2010). There have been a large number of candidate biomarkers but only a relative few have been validated. The lack of progress in finding biomarkers for hard to diagnose diseases (at least at an early stage), such as ovarian and pancreatic cancer, is particularly disconcerting. Finally, the integration of proteomic data with the data obtained from the other 'omics (principally genomics, transcriptomics and metabolomics) to form a seamless continuum has not yet been effectively achieved although there is growing interest in doing so. Indeed, this can be taken to be a major objective of the rapidly developing field of systems biology.

Receptor Tyrosine Kinase Dimerization/Activation

The determination that the receptors for some growth factors/hormones contained tyrosine kinases was a major development in cell signaling (Schlessinger and Ullrich 1992). This feature was presaged by the findings of Hunter and colleagues (Eckert *et al.* 1979; Hunter and Sefton 1980) that the Src protein contained this activity followed by the results of Ushiro and Cohen (Ushiro and Cohen 1980) that demonstrated that the epidermal growth factor (EGF) receptor also catalyzed this reaction when stimulated by ligand binding. The subsequent cloning of both the EGF and insulin receptors (Ullrich *et al.* 1984; Ebina *et al.* 1985; Ullrich *et al.* 1985) confirmed the presence of this entity in their endodomain portions. This led to the postulation that the activation of receptor activity was synonymous with ligand-induced dimerization of receptor protomers, that brought the two kinase domains into juxtaposition to allow intermolecular phosphorylations to occur, a concept that has survived for nearly two decades (Hubbard 2000) (Fig 2). Of course, this model did not apply to the insulin receptor because it was already established that the protomer of this receptor underwent limited proteolysis to produce alpha and beta chains that were held together by

both inter- and intraprotomer disulfide bonds to form a covalently stabilized heterotetramer. Hence, ligand binding could not induce oligomerization, as it was already preformed. As cloning and genome sequencing revealed the full portrait of the human RTK super family (Blume-Jensen and Hunter 2001) (Fig 3), which contains 58 members grouped in 19 families, it became clear that the insulin receptor family, including the two other members, is the only one with covalently associated subunits. All of the rest are associated non-covalently.

The ligand-induced model for RTK activation was thus based primarily on the EGF receptor and was in part due to the then widely held view that EGF itself was also a dimeric species (Taylor *et al.* 1974). It was only established later on that this is really only true of the high molecular weight complex found in mouse submandibular glands and that EGF alone in solution is basically monomeric at concentrations appropriate for receptor binding studies. Ultimately structural analyses of EGF receptor ectodomain-ligand complexes showed that EGF was bound in a monomeric fashion to each protomer of the active dimer and that there was no contact between ligand molecules in the bound state (Garrett *et al.* 2002; Ferguson *et al.* 2003). Nonetheless, the concept that ligands induce receptor oligomerization for essentially all receptors remains a commonly held view (For example (Choudhary and Mann 2010)).

The first suggestions that activation might proceed through a ligand-induced conformational change and/or protomer rearrangement in a preformed oligomer was made by Gadella and Jovin (Gadella and Jovin 1995) (Fig 4). This model didn't really address whether monomeric species might exist in equilibrium with dimers (or higher order oligomers), but was rather predicated on the idea that ligand binding occurred only on the aggregated form and that activation and autophosphorylation resulted from the release of inhibition of the kinase domain (in some manner). This presumably is related to the movement of the activation loop from the closed (inactive) to open (active) conformation. Indeed, if monomers do exist, as required by the ligand-induced oligomerization model, they would necessarily have to exist in some kinase-inhibited form to allow for ligand-based activation. Refinements on this constitutive oligomerization model have subsequently appeared (Moriki *et al.* 2001; Martin-Fernandez *et al.* 2002; Zhu *et al.* 2003).

The constitutive oligomerization model has been buttressed by a considerable amount of chemical modification/mutation, physical chemical and structural data. Definitive measurements for a number of the RTK families support this model with similar observations for other receptor classes, such as the cytokine and GPCR types (Livnah *et al.* 1999; Remy *et al.* 1999; Krause *et al.* 2002; Kniazeff *et al.* 2004; Tateyama *et al.* 2004). It now appears that the kinase activation loop prefers the open conformation and that there are various means for maintaining the loop in the closed conformation in the preformed unliganded dimers (or oligomers). When this inhibition is released by the ligand-induced protomer reorganization, the now active kinases are free to autophosphorylate tyrosine residues in the activation loop (thus anchoring it in the open conformation) as well as other tyrosine residues located throughout the endodomain (See Fig 5). These then provide the sites for effector/adaptor/scaffold binding that initiates downstream signaling.

Phosphoproteomics

The induction of protein phosphorylation in eukaryotic cells by RTKs and other receptor classes is the major currency of cell signaling. There are over 500 protein kinases (of which 90 are tyrosine kinases) and more than 100 protein phosphatases in the human genome, which underscores the investment that has been made in this modification (Manning *et al.* 2002). Although phosphorylation is less than 1% of the total PTMs that occur, it is very

widespread (even in non-stimulated cells), can occur on multiple sites in the same protein and often does so in substoichiometric amounts. Early analyses (identifications) were largely based on Western blots that depended on specific antibody reagents to identify the various downstream participants and were therefore limited by reagent availability (Schlessinger 2000). Nonetheless, these studies were effective in establishing the main signaling pathways, which in the case of RTKs is the Ras/mitogen activated protein kinase (MAPK or ERK) pathway, the phospholipase C γ (PLC γ) pathway and the phosphatidylinositol-3-kinase (PI-3-K)/AKT pathway (See Fig 6). Depending on the cell and receptor type there are variations in both the constituents involved and in the phenotypic responses that are manifestations of the induction of additional PTMs and various transcriptional regulators. Importantly, there are other additional pathways, many of which are primarily used by different receptor systems, that can also be activated including those effected by receptor cross talk.

The unbiased nature of large scale mass spectrometric analyses, both to identify proteins and their modifications and to determine protein-protein interactions through the elucidation of complex composition, greatly amplified the details of these pathways and revealed the extent to which PTMs are used by cells to proliferate information transfer and response. For the most part these large scale identification experiments have depended on so-called bottom up strategies where complex samples with little or no pre-fractionation, such as the lysate of a growth factor-stimulated cell preparation, are proteolyzed, most commonly with trypsin, and the digested sample subjected to mass spectrometric analysis following one or more chromatographic purification steps. The identification of the peptides is deduced by generating peptide fragments of preselected precursor peaks that reveal all or part of the amino acid sequence of the parent peptide, which is then matched to an entry in an appropriate database by a search engine. This general proteomic strategy is depicted schematically in Fig 7. If the object of the experiment is to identify peptides (or the proteins from which they are derived) that bear a certain PTM, then an enrichment step designed to capture peptides so modified is usually included before the MS step. In the case of phosphorylation this usually means either using a phospho-specific antibody or chromatography on an immobilized-metal (IMAC) or TiO_2 column. Phosphoramidate chemistry can also be employed (Bodenmiller *et al.* 2007). The resulting list of proteins can be subjected to multiple additional analyses to group, classify or connect them in ways intended to reveal aspects of the biological process(es) under investigation. Importantly, to accurately determine the levels of the changes observed (in either the proteins themselves or their PTMs), some quantitative strategy for the mass spectrometric measurements is required (Ong and Mann 2005).

Although all peptide/protein identifications by mass spectrometry are subject to a certain amount of error, PTMs are a particular problem for two reasons: the added mass found on a residue may not be readily identified and, even it is, its location in the sequence may be difficult to ascertain with certainty (Bradshaw *et al.* 2010). Indeed, this is presently among the most challenging problems facing proteomics. There are no immediate solutions but better software that does not make arbitrary assignments when the data do not warrant it and the deposition of raw data for each experiment that can be interrogated by other investigators (Bradshaw and Burlingame 2010) are likely to improve this situation as is a general appreciation of the extent to which this is a problem by the scientific community.

The application of global mass spectrometric-based screens to elucidate protein phosphorylation profiles in various biological systems has produced a substantial atlas for this modification (<http://www.phosphosite.org/homeAction.do>; <http://141.61.102.18/phosida/index.aspx>; <http://phospho.elm.eu.org/>). More than two dozen such experiments have been reported in a variety of organisms (mostly human, mouse, and

yeast) and most relate to some form of signaling; for example (Choudhary and Mann 2010). These range from growth factor stimulation to stem cell differentiation and include studies of cell cycle changes, the effect of various inhibitors and immune responses, as examples. Several studies have directly measured RTK-based responses, the largest being a detailed kinetic study (using SILAC technology) of the stimulation of HeLa cells by EGF (Olsen *et al.* 2006). Some 6600 sites of phosphorylation on serine, threonine and tyrosine were reported. A second study of EGF signaling using several downstream kinase inhibitors detected a similar number of sites and suggested that a high percentage of the growth factor induced modifications were affected by one or another of the MAPK inhibitors tested (Pan *et al.* 2009). Several studies on RTKs have been performed in order to identify the phosphotyrosine proteome (Blagoev *et al.* 2004; Schmelzle *et al.* 2006; Kruger *et al.* 2008). In general, the tyrosine-phosphorylated proteins are immunoprecipitated and analyzed by high resolution mass spectrometry in combination with the SILAC method. In one study (Kruger *et al.* 2008) they identified 40 insulin-induced effectors, seven of which had not previously been described in insulin signaling, including SDR, PKCdelta binding protein, LRP-6, and PISP/PDZK11, a potential calcium ATPase binding protein. The bidirectional signaling of the ephrin B receptor 2 and ephrin B1 (both receptors and ligands are membrane bound in their respective cells) have been explored by using differentially labeled (SILAC) receptor and ligand cells (Jorgensen *et al.* 2009). Finally, a comparative study of EGF and platelet-derived growth factor (PDGF) stimulation of mesenchymal cells (which differentiate into bone progenitor cells) showed that PDGF was able to activate the PI3K pathway but EGF was not (Kratchmarova *et al.* 2005).

In experiments utilizing stably transfected PC12 cells, we have used chimeric receptors, formed from the extracellular domain of the PDGFR and the transmembrane and endodomain of TrkA (the receptor for NGF) to examine the downstream phosphorylation profile at 20 minutes following stimulation, that primarily features serine and threonine modifications. As PC12 cells do not express the PDGFR this approach allowed us to examine TrkA signals without the stimulation of either endogenous TrkA or the pan neurotrophin receptor P75. It also allowed us to examine the effect of ablating the two principal autocatalytically-generated phosphotyrosine docking sites at residues 490 and 785 in the TrkA endodomain (Obermeier *et al.* 1994). This is the first global phosphorylation analysis of TrkA signaling and the first to use mutant receptors* to ascertain the effects of ablating select signaling pathways, wholly or in part. Interestingly the overlap in phosphorylated proteins between this study (J.Biarc, R.J. Chalkley, A. L. Burlingame, and R.A. Bradshaw, manuscript in preparation) and EGF stimulated HeLa cells at a comparable time point was less than 20 percent, despite the fact that both RTKs stimulate basically the same major signaling pathways.¹

The Role of Other PTMs in Cell Signaling

Although there are a large number of biologically-introduced PTMs that have been reported (> 200) and certainly several more to be identified as judged by the detection of undetermined side chain masses using mass spectrometry, not all are likely to participate in cell signaling. Furthermore, only a relatively small number are amenable at this juncture to large-scale global mass spectrometric analysis. Leaving aside limited proteolysis, which is the basis of a separate area of study (often referred to as degradomics) and phosphorylation, already described above, acetylation, methylation and O-GlcNAcylation are the main transient, regulatory alternatives. The first two are of particular importance in histone modification and are key to understanding the epigenetic role of these nucleosome proteins

¹*Choudhary et al (Choudhary *et al.* 2009) reported a detailed phosphoproteomic analysis of the oncogenic RTK, FLT3 which is constitutively active.

and their effect on gene transcription. Ubiquitination (including all of the closely related derivatives such as NEDD8, SUMO, etc) is becoming of increasing interest because of the wide variety of different activities that is apparently associated with it and it should probably be included in any list of key PTMs (Kerscher *et al.* 2006). Ubiquitin family proteins produce distinctive proteolytic fingerprints upon enzymatic digestion that can be used for identification of sites using mass spectrometry. Some of these modifications are large, even after proteolytic cleavage, but appropriate sequence mutations can be introduced to facilitate their site identification (Knuesel *et al.* 2005).

O-GlcNAcylation, which has been linked to phosphorylation in a number of cases, presents many interesting contrasts to the phosphorylation regulatory approach (Zeidan and Hart 2010). In the first place, this intracellular modification, which places a single N-acetylglucosamine on a serine or threonine residue, is catalyzed by a single enzyme as opposed to the several hundred protein kinases involved in protein phosphorylation (Hart *et al.* 2007). Instead, it is activated as part of multi-protein complexes, where specificity is controlled through protein-protein interactions. Secondly, this modification is reversible, but again only a single O-GlcNAcase is known. Finally, several hundred sites for this modification have been reported and there is no clear recognition motif, although there are apparently preferred amino acid residues flanking the attachment residue (Chalkley *et al.* 2009). This lack of sequence specificity is curious, but explains why all the various modifications of this class can be accomplished with a single catalyst.

The detection of O-GlcNAc modifications in wide-scale mass spectrometric experiments requires enrichment of the modified peptides prior to the analysis. A number of strategies have been employed including immobilized lectins in long affinity chromatography columns (Vosseller *et al.* 2006) or chemoenzymatic strategies that specifically attach an azido sugar to the modification, to which a biotin affinity tag can be ligated and used for enrichment of the O-GlcNAc peptides (Wang *et al.* 2010). O-GlcNAc modification performs a wide range of roles from regulating transcriptional activation to preventing proteins from being targeted for proteosomal degradation. There have been a number of cases where this modification competes with phosphorylation, suggesting that the two PTMs can act in a “Ying-Yang” fashion; one regulating the other, presumably with opposite effects on the function of the protein. For example, the insoluble neurofibrillary tangles that are characteristic of Alzheimer's disease contain hyperphosphorylated tau. O-GlcNAcylation has been shown to compete with phosphorylation on tau and elevating O-GlcNAcylation using an O-GlcNAcase inhibitor reduces tau phosphorylation levels and reverting it back to a soluble state and removing the protein aggregation (Yuzwa *et al.* 2008). There is also evidence that these two modifications may affect one another without directly competing for the same site (Fig 8). In a large scale study of a mouse brain post-synaptic density (PSD) preparation that had previously been analyzed extensively for phosphorylations (Trinidad *et al.* 2006), some 86 modified peptides were detected after lectin purification (Chalkley *et al.* 2009). One of the proteins known to be involved in synaptic functions, bassoon, had 28 O-GlcNAc sites identified as compared to 37 phosphorylation sites, with three sites in common. The reasons for this extensive derivatization by both modifications is unknown but is likely to involve the regulation of protein-protein interactions.

Concluding Remarks

The RTKs are one of the most important families mediating transmembrane signaling and they participate and are instrumental in regulating a broad range of physiological activities. Indeed, tyrosine kinases in general, and the processes that they control and/or stimulate, provide a rich source of drug targets, particularly in growth related disorders such as cancer (Zwick *et al.* 2002; Krause and Van Etten 2005). However, there remain many questions

regarding their activation and downstream signaling and the application of proteomic analyses promises to answer many of them. There have been relatively few detailed studies of this type to date and it will require considerably more of them to better define the pathways with respect to both the major and minor PTMs that, along with the protein-protein interactions, are the means to direct the flow of the signals generated. It will take such approaches to define the specificity that characterize the individual families, even appreciating that to some degree all are capable of activating many, if not all, of the principal pathways. It will also be necessary to understand, in the highly complex networks of intracellular phosphorylation (that contain thousands of sites of modification and clearly have not yet been fully determined in any paradigm), exactly which kinases modify which substrates, and to work out the inter-relationships with other modifications such as O-GlcNAcylation and acetylation. Only then will it be possible to determine which modifications are physiologically significant and which are simply background. Along the way, these studies should continue to provide potential drug targets and perhaps improve the current lackluster biomarker discovery track record.

Acknowledgments

Portions of the work described were supported by NIH NCRR grant P41 RR001614. The authors thank Erin Schalles for assistance in preparing this article.

References

- Aberth W, Straub KM, Burlingame AL. Secondary ion mass spectrometry with cesium ion primary beam and liquid target matrix for analysis of bioorganic compounds. *Anal Chem*. 1982; 54:2029–2034.
- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422:198–207. [PubMed: 12634793]
- Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprpto A, Karni-Schmidt O, Williams R, Chait BT, Rout MP, Sali A. Determining the architectures of macromolecular assemblies. *Nature*. 2007; 450(7170):683–94. [PubMed: 18046405]
- Anderson NL. The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin Chem*. 2010; 56(2):177–85. [PubMed: 19884488]
- Barber, M.; Bordoli, RS.; Sedgwick, RD. Fast atom bombardment mass spectrometry. In: Morris, HR., editor. *Soft ionization Biological Mass spectrometry*. London: Heyden and Son, Ltd; 1981. p. 137-152.
- Blagoev B, Ong SE, Kratchmarova I, Mann M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nature biotech*. 2004; 22:1139–1145.
- Blume-Jensen P, Hunter T. Oncogenic kinase signalling. *Nature*. 2001; 411(6835):355–65. [PubMed: 11357143]
- Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. Reproducible isolation of distinct, overlapping segments of the proteome. *Nature Methods*. 2007; 4:231–237. [PubMed: 17293869]
- Bradshaw RA, Burlingame AL. From Proteins to Proteomics. *IUBMB Life*. 2005; 57:267–272. [PubMed: 16036609]
- Bradshaw RA, Burlingame AL. Technological Innovation Revisited. *Mol Cell Proteomics*. 2010;9.
- Bradshaw RA, Medzihradszky KF, Chalkley RJ. Protein PTMs: Post-translational modifications or pesky trouble makers. *JMass Spectrom*. 2010; 45:1095–1097. [PubMed: 20635432]
- Chalkley RJ, Thalhammer A, Schoepfer R, Burlingame AL. Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. *Proc Natl Acad Sci U S A*. 2009; 106(22):8894–9. [PubMed: 19458039]
- Choudhary C, Mann M. Decoding signalling networks by mass spectrometry-based proteomics. *Nature RevMol Cell biol*. 2010; 11:427–439.

- Choudhary C, Olsen JV, Brandts C, Coc J, Reddy PNG, Bohmer FD, Gerke V, Schmidt-Arras D-E, Berdel WE, Mueller-Tidow C, Mann M, Serve H. Mislocalized Activation of Oncogenic RTKs Switches Downstream Signaling Outcomes. *Mol Cell*. 2009; 36:326–339. [PubMed: 19854140]
- Cohen J. The Proteomics Payoff. *Tech Rev*. 2001; 104:55–63.
- Collins SR, Kemmeren P, Zhao X-C, Greenblatt JF, Spencer F, Holstege FCP, Weissman JS, Krogan NJ. Toward a Comprehensive Atlas of the Physical Interactome of *Saccharomyces cerevisiae*. *Mol Cell Proteomics*. 2007; 6:439–450. [PubMed: 17200106]
- Ebina Y, Ellis L, Jarnagin K, Edery M, Graf L, Clauser E, Ou JH, Masaiarz F, Kan YW, Goldfine ID, Roth RA, Rutter W. The human insulin receptor cDNA: the structural basis for hormone transmembrane signalling. *Cell*. 1985; 40:747–758. [PubMed: 2859121]
- Eckert W, Hutchinson MA, Hunter T. An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell*. 1979; 18:925–933. [PubMed: 229973]
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 1989; 246:64–71. [PubMed: 2675315]
- Ferguson KM, Berger MB, Mendrola JM, Cho HS, Leahy DJ, Lemmon MA. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol Cell*. 2003; 11(2): 507–17. [PubMed: 12620237]
- Gadella TW, Jovin TM. Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. *J Cell Biol*. 1995; 129(6):1543–58. [PubMed: 7790353]
- Garrett TPJ, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Zhu H-J, Walker F, Frenkel MJ, Hoyne PA, Jorissen RNECN, Burgess AW, Ward CW. Crystal structure of a Truncated Epidermal Growth Factor Receptor Extracellular Domain Bound to Transforming Growth Factor Alpha. *Cell*. 2002; 110:763–773. [PubMed: 12297049]
- Hart GW, Housley MP, Slawson C. Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature*. 2007; 446(7139):1017–22. [PubMed: 17460662]
- Hubbard SR. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J*. 1997; 16(18):5572–81. [PubMed: 9312016]
- Hubbard SR, Wei L, Ellis L, Hendrickson WA. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*. 1994; 372(6508):746–54. [PubMed: 7997262]
- Hubbard, SRaTJH. Protein Tyrosine Kinase Structure and Function. *Annu Rev Biochem*. 2000; 69:373–398. [PubMed: 10966463]
- Huber LJ. Is proteomics heading in the wrong direction? *Nat Rev Mol Cell Biol*. 2003; 4:74–80. [PubMed: 12511871]
- Hunter T, Sefton BM. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proc Natl Acad Sci USA*. 1980; 77:1311–1315. [PubMed: 6246487]
- Huse M, Kuriyan J. The conformational plasticity of protein kinases. *Cell*. 2002; 109(3):275–82. [PubMed: 12015977]
- Jorgensen C, Sherman A, Chen GI, Pasculescu A, Poliakov A, Hsiung M, Larsen B, Wilkinson DG, Linding R, Pawson T. Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science*. 2009; 326(5959):1502–9. [PubMed: 20007894]
- Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons. *Anal Chem*. 1988; 60:2299–2301. [PubMed: 3239801]
- Kerscher O, Felberbaum R, Hochstrasser M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol*. 2006; 22:159–80. [PubMed: 16753028]
- Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik*. 1975; 26:231–243. [PubMed: 1093965]
- Kniazeff J, Bessis AS, Maurel D, Ansanay H, Prezeau L, Pin JP. Closed state of both binding domains of the homodimeric mGlu receptors is required for full activity. *Nat Struct Mol Biol*. 2004; 11:706–713. [PubMed: 15235591]
- Knuesel M, Cheung HT, Hamady M, Barthel KK, Liu X. A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. *Mol Cell Proteomics*. 2005; 4(10):1626–36. [PubMed: 16020427]

- Kratchmarova I, Blagoev B, Haack-Sorenson M, Kassem M, Mann M. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science*. 2005; 308:1472–1477. [PubMed: 15933201]
- Krause CD, Mei E, Xie J, Jia Y, Bopp MA, Hochstrasser RM, Pestka S. Seeing the Light: Preassembly and Ligand-Induced Changes of the Interferon gamma Receptor Complex in Cells. *Mol Cell Proteomics*. 2002; 1:805–815. [PubMed: 12438563]
- Krause DS, Van Etten RA. Tyrosine kinases as targets for cancer therapy. *N Engl J Med*. 2005; 353(2):172–87. [PubMed: 16014887]
- Kruger M, Kratchmarova I, Blagoev B, Tseng Y-H, Kahn CR, Mann M. Dissection of the insulin signaling pathway via quantitative phosphoproteomics. *Proc Natl Acad Sci USA*. 2008; 105:245–2456.
- Livnah O, Stura EA, Middleton SA, Johnson DL, Jolliffe LK, Wilson IA. Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science*. 1999; 283(5404):987–90. [PubMed: 9974392]
- MacBeath G. Protein microarrays and proteomics. *Nature Genetics Supplement*. 2002; 32:526–531.
- Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science*. 2002; 298:1912–1934. [PubMed: 12471243]
- Martin-Fernandez M, Clarke DT, Tobin MJ, Jones SV, Jones GR. Preformed oligomeric epidermal growth factor receptors undergo an ectodomain structure change during signaling. *Biochem J*. 2002; 362:2415–2427.
- Moriki T, Maruyama H, Maruyama IN. Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. *J Mol Biol*. 2001; 311(5):1011–26. [PubMed: 11531336]
- O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*. 1975; 250:4007–4021. [PubMed: 236308]
- Obermeier A, Bradshaw RA, Seedorf K, Choidas A, Schlessinger J, Ullrich A. Neuronal differentiation Signals are Controlled by Nerve Growth Factor receptor/Trk Binding Sites for SHC and PLC gamma. *EMBO J*. 1994; 13:1585–1590. [PubMed: 8156997]
- Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell*. 2006; 127:635–648. [PubMed: 17081983]
- Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol*. 2005; 1(5):252–62. [PubMed: 16408053]
- Pan C, Olsen JV, Daub H, Mann M. Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics. *Mol Cell Proteomics*. 2009; 8:2796–2808. [PubMed: 19651622]
- Pawson T. Regulation and targets of receptor tyrosine kinases. *Eur J Cancer*. 2002; 38(Suppl 5):S3–S10. [PubMed: 12528767]
- Remy I, Wilson IA, Michnick SW. Erythropoietin receptor activation by a ligand-induced conformation change. *Science*. 1999; 283(5404):990–3. [PubMed: 9974393]
- Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell*. 2000; 103:211–225. [PubMed: 11057895]
- Schlessinger J, Ullrich A. Growth factor Signaling by Receptor Tyrosine Kinases. *Neuron*. 1992; 9:383–391. [PubMed: 1326293]
- Schmelzle K, Kane S, Gridley S, Lienhard GE, White MF. Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes*. 2006; 55:2171–2179. [PubMed: 16873679]
- Schweitzer B, Predki P, Snyder M. Microarrays to characterize protein interactions on a whole-proteome scale. *Proteomics*. 2003; 3:2190–2199. [PubMed: 14595818]
- Sreekumar A, Chinnaiyan AM. Using protein microarrays to study cancer. *BioTechniques*. 2002; 33:S46–S53.
- Tanaka K, Waki H, Ido Y, Akita S, Yoshida T, Matsuo T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Comm Mass Spectrometry*. 1988; 2:151–153.

- Tateyama M, Abe H, Nakata H, Saito O, Kubo Y. Ligand-induced rearrangement of the dimeric metabotropic glutamate receptor 1alpha. *Nat Struct Mol Biol.* 2004; 11:637–42. [PubMed: 15184890]
- Taylor JM, Mitchell WM, Cohen S. Characterization of the High Molecular Weight Form of Epidermal Growth factor. *J Biol Chem.* 1974; 249:3198–3203. [PubMed: 4830241]
- Trinidad JC, Specht CG, Thalhammer A, Schoepfer R, Burlingame AL. Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Mol Cell Proteomics.* 2006; 5(5): 914–22. [PubMed: 16452087]
- Ullrich A, Bell JR, Chen EY, Herrera R, Petruzzelli LM, Dull TJ, Gray A, Coussens L, Liao YC, Tsubokawa M, Mason A, Seeburg PH, Grunfeld O, Rosen M, Ramachandran J. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature.* 1985; 313:756–761. [PubMed: 2983222]
- Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Libermann TA, Schlessinger J, Downward J, Mayes ELV, Whittle N, Waterfield MD, Seeburg PH. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature.* 1984; 309:418–25. [PubMed: 6328312]
- Ushiro H, Cohen S. Identification of phosphotyrosine as a product of epidermal growth factor-activated protein kinase in A-431 cell membranes. *J Biol Chem.* 1980; 255(18):8363–5. [PubMed: 6157683]
- Vosseller K, Trinidad JC, Chalkley RJ, Specht CG, Thalhammer A, Lynn AJ, Snedecor JO, Guan S, Medzihradszky KF, Maltby DA, Schoepfer R, Burlingame AL. O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol Cell Proteomics.* 2006; 5(5):923–34. [PubMed: 16452088]
- Wang Z, Udeshi ND, O'Malley M, Shabanowitz J, Hunt DF, Hart GW. Enrichment and site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry. *Mol Cell Proteomics.* 2010; 9(1): 153–60. [PubMed: 19692427]
- Yuzwa SA, Macauley MS, Heinonen JE, Shan X, Dennis RJ, He Y, Whitworth GE, Stubbs KA, McEachern EJ, Davies GJ, Vocadlo DJ. A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nat Chem Biol.* 2008; 4(8):483–90. [PubMed: 18587388]
- Zeidan Q, Hart GW. The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways. *J Cell Sci.* 2010; 123(Pt 1):13–22. [PubMed: 20016062]
- Zhu HJ, Iaria J, Orchard S, Walker F, Burgess AW. Epidermal growth factor receptor: association of extracellular domain negatively regulates intracellular kinase activation in the absence of ligand. *Growth Factors.* 2003; 21(1):15–30. [PubMed: 12795333]
- Zwick E, Bange J, Ullrich A. Receptor tyrosine kinases as targets for anticancer drugs. *Trends Mol Med.* 2002; 8(1):17–23. [PubMed: 11796262]

Only a handful of new protein diagnostics Introduced in past 15 yrs and current rate is only ~1.5/yr

Figure 1.

FDA-cleared or -approved protein tests approved under CLIA regulations since 1993. IGFBP, insulin-like growth factor-binding protein; IL-2R, interleukine-2 receptor; TnT, troponin T; IGF-II, insulin-like growth factor II; CA27.29, cancer antigen 27.29 (same as CA15-3); LP-PLA2, lipoprotein-associated phospholipase A2; PCT, procalcitonin; MPO, myeloperoxidase; sMesothelin, soluble-related peptide; HE4, cancer antigen, human epididymis protein4. Reproduced from (Anderson 2010) with permission.

Figure 2.

The «ligand-induced oligomerization» model of growth factor receptor activation. (A) a growth factor binds a single receptor protomer that causes its dimerization; (B) the dimer formed allows the autophosphorylation in a trans fashion, creating docking sites that (C) link effectors and adaptors to the activation of several signaling pathways.

Figure 3.

Human receptor protein-tyrosine kinase family. The prototypic receptor for each family is indicated above the receptor, and the known members are listed below. Abbreviations of the prototypic receptors: EGFR, epidermal growth factor receptor; InsR, insulin receptor; PDGFR, platelet-derived growth factor receptor; VEGFR; vascular endothelial growth factor receptor; FGFR, fibroblast growth factor receptor; KLG/CCK, colon carcinoma kinase; NGFR, nerve growth factor receptor; HGFR, hepatocyte growth factor receptor, EphR, ephrin receptor; Axl, a Tyro3 PTK; TIE, tyrosine kinase receptor in endothelial cells; RYK, receptor related to tyrosine kinases; DDR, discoidin domain receptor; Ret, rearranged during transfection; ROS, RPTK expressed in some epithelial cell types; LTK, leukocyte tyrosine kinase; ROR, receptor orphan; MuSK, muscle-specific kinase; LMR, Lemur. Other abbreviations: AB, acidic box; CadhD, cadherin-like domain; CRD, cysteine-rich domain; DiscD, discoidin-like domain; EGFD, epidermal growth factor-like domain; FNIII, fibronectin type III-like domain; IgD, immunoglobulin-like domain; KrinD, kringle-like domain; LRD, leucine-rich domain. The symbols α and β denote distinct RPTK subunits. An asterisk indicates that the member is devoid of intrinsic kinase activity. Reproduced from (Blume-Jensen and Hunter 2001) with permission.

Figure 4.

The «rotation-activation» model, where the ligand induces a conformational change and/or a protomer rearrangement in a pre-formed oligomer. In this model the ligand binding occurs only on the associated form and the activation and autophosphorylation results from the release of inhibition of the kinase domain. Modeled after Gadella and Jovin (Gadella and Jovin 1995).

Figure 5.

The catalytically active conformation of the protein kinase domain. Representation of the kinase domain of the insulin receptor in the active (left) and inactive (right) forms (Hubbard *et al.* 1994; Hubbard 1997). The two structures have been schematized to clearly depict the conformational transitions involved in the regulation of kinase activity, with particular emphasis on the α C helix and activation loop. The catalytic lysine and glutamate are shown in each schematic as are the tyrosine sites of phosphorylation. Reproduced from (Huse and Kuriyan 2002) with permission.

Figure 6.

Main signaling pathways activated by RTKs. After stimulation, the receptor undergoes autophosphorylation, creating docking sites that link effectors and adaptors to the activation of several signaling pathways, including those depicted. The first adaptor protein leading to the activation of the MAPK pathways can be Grb2, Shc, IRS/FRS, etc.

Figure 7.

Typical work flow for proteome and PTM analysis using shotgun proteomics. (a) Proteins extracted from organs, tissues or cells are separated by one-dimensional polyacrylamide gel electrophoresis (1D PAGE) and in-gel digested into peptides using proteases such as trypsin. The peptides containing specific post-translational modifications (PTMs) can be enriched using different approaches. Non-modified peptides are used to identify and quantify total cellular proteins. (b) Purified peptides are separated on a miniaturized reverse phase chromatography column with an organic solvent gradient. Peptides eluting from the column are ionized by electrospray at the tip of the column, directly in front of the mass spectrometer (known as on-line coupling). (c) The electrosprayed ions are transferred into the vacuum of the mass spectrometer. In the mass spectrometry (MS) mode, all ions are moved to the mass analyser, where they are measured (top mass spectrum). The first mass analyser then selects a particular peptide ion and fragments it in a collision cell. The inset in the MS panel indicates the stable isotope labelling by amino acids in cell culture (SILAC) ratio of one of the peptides. For modified peptides, the peptide mass will be shifted by the mass of the modification, as will all fragments containing the modification, allowing the unambiguous placement of the PTM on the sequence. (d) The mass and list of fragment masses for each peptide are scanned against protein sequence databases, resulting in a list of identified peptides and proteins. These lists of proteins and their quantitative changes are the basis for biological discovery. Reproduced from (Choudhary and Mann 2010) with permission.

Figure 8.

Four different types of O-GlcNAc (G)–phosphate (P) crosstalk on protein substrates. (A) Alternative and competitive occupancy at the same amino acid residue, such as in the transactivation domain (TAD) of the transcription factor Myc, which can be either O-GlcNAcylated or phosphorylated at threonine 58. (B) Alternative and reciprocal occupancy at different sites, such as in the regulatory domain (RD) of the transcription factor C/EBPβ, which is either phosphorylated at threonine 179, serine 184 and threonine 188, or O-GlcNAcylated at serines 180 and 181. (C) Simultaneous occupancy at different sites, such as in IRS-1, which is both O-GlcNAcylated at multiple residues in the C-terminal domain (which contains several docking sites for SH2-domain-containing proteins) and phosphorylated at serines 307, 632 and 635 (which are known to attenuate the insulin signaling pathway). (D) Site-dependent reciprocal or simultaneous occupancy, such as in CaMKIV, which exists either as an O-GlcNAc–phospho active protein (at residues threonine 57/serine 58 for O-GlcNAc and threonine 200 for phosphate) or as an O-GlcNAcylated inactive protein (at serine 189). AD, activation domain; CaMBD, calmodulin-binding domain; DBD, DNA-binding domain; HLH, helix-loop-helix domain; PHD, pleckstrin-homology domain; PKD, protein kinase domain; PRD, proline-rich domain; PTBD, phospho-tyrosine-binding domain; SRD, serine-rich domain. Reproduced from (Zeidan and Hart 2010), *J Cell Science*, with permission.

Table 1**Proteomics: Components and Active Areas of Research**

Proteomic Component	Applications
Determination of the structure and function of all members of the proteome	Protein Identification; PTMs and splice variants identification; Characterization of protein 3D structure
Determination of all protein-protein interactions	Defining protein complexes (stable and transient)
Dynamics (as a function of time and situation) of all proteins and their modifications	Expression profiling – changes in response to stimuli, including drug treatment
Integration of proteomic data with other' omic data	Bioinformatics and Systems Biology

Table 2**What Has Proteomics Accomplished?**

-
- Large-scale surveys of proteins present/expressed in cells/organelles/biological fluids
 - Characterization of cellular protein networks (protein/protein interactions)
 - Identification of the composition and organization of protein machines
 - Technical advances in analytical sensitivity and relative quantification between samples
 - Analyses of clinically-relevant materials to deduce disease-based alterations
-

Table 3**What Hasn't Proteomics Accomplished?**

-
- Assignment of function/s to a significant percentage of the human proteome
 - A useful integration (understanding) of proteomic data with that of transcriptomics and metabolomics
 - Reliable identification and quantification of PTMs
 - Identification of novel biomarkers for hard-to - diagnose disease states
-