

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Proteomik und Bioanalytik

Development of phospho- and chemoproteomic methods to study cellular signaling

Benjamin L. R. Ruprecht

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. M. Klingenspor

Prüfer der Dissertation: 1. Prof. Dr. B. Küster
2. Prof. Dr. S. Lemeer
3. Prof. Dr. J. Olsen

Die Dissertation wurde am 18.02.2016 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 19.04.2016 angenommen.

Abstract

Kinase mediated protein phosphorylation on serine, threonine and tyrosine side chains is an important post-translational modification which affects and governs a large body of cellular signaling in health and disease. In recent years, mass spectrometry-based proteomics has emerged as the prime technology for the large scale, explorative and proteome-wide analysis of the kinome and associated phosphorylation events.

Owing to their substoichiometric nature, phosphorylated peptides need to be enriched prior to measurement. In this regard, chapter 2 of this thesis describes the development of a comprehensive, unbiased, selective and reproducible means to isolate phosphopeptides out of complex proteome digests using a high performance liquid chromatography-based system including an iron immobilized metal affinity chromatography (IMAC) column. This chromatographic method enabled efficient phosphopeptide elution, scaled linearly with the amount of input material and outperformed all tested state of the art enrichment methods (titanium-IMAC and titanium dioxide). Together with an increased analytical depth provided by downstream hydrophilic strong anion exchange chromatography, the workflow can be flexibly applied to a wide range of experiments. Moreover, the study challenged the common notion that different enrichment materials are capable of purifying complementary phosphopeptide species. Instead, the data suggested that the observed complementarity mainly stemmed from a combination of format inadequacies, inefficient elution and insufficient data acquisition speed.

Despite extensive efforts, the phosphorylation site coverage of proteomes is still incomplete and thus merits the exploration of alternative methods complementing existing workflows. In chapter 3, matrix-assisted laser desorption/ionization (MALDI) and nano-electrospray ionization (nESI) tandem mass spectrometry (MS/MS) were evaluated for their ability to identify different phosphopeptide species. MALDI-MS/MS favored the detection of acidic and phosphotyrosine bearing phosphopeptides and was in its orthogonal nature comparable to the use of alternative proteases such as Asp-N, Arg-C, chymotrypsin, Glu-C or Lys-C when employing only nESI-MS/MS. Hence, MALDI-MS/MS can be an useful complement to nESI-MS/MS for phosphorylation site mapping.

Since kinase activity can only be indirectly inferred from phosphoproteomic data, chapter 4 contains a detailed investigation of the question if immobilized kinase inhibitors are capable of capturing kinases in an activity dependent fashion, a concept that would enable the direct measurement of activity. Cellular kinase activity changes caused by pervanadate treatment were determined by quantitative phosphoproteomics and compared to differential kinase binding in a chemoproteomics setup. The obtained results suggested that activity-based kinase binding was a rare event which was dependent on (i) the individual kinase (ii) the applied inhibitor bead matrix and (iii) the kinase's cellular activation status. Hence, rather than being of generic use as a direct read out of kinase activity, chemoproteomic methods should be confined to a limited number of well characterized cases.

The last chapter of this thesis combines methodologies developed, implemented and examined in previous chapters and describes a proteomic, chemoproteomic and phosphoproteomic mode of action and resistance analysis of the kinase inhibitor lapatinib in an ERBB2 overexpressing breast cancer cell line. Obtained results showed that the acquisition of resistance was accompanied by changes in several different, therapeutically relevant targets (such as enhanced CDK1/2 activity) and phenotypes (such as increased invasive potential). A comparison of the data from this study to previously described mechanisms of resistance against ERBB2 targeted therapy revealed an extensive heterogeneity, but concomitantly suggested that glycolytic addiction might be a common trait. Mechanistically, phosphorylation mediated metabolic reprogramming of LDHA and PDHA1 activity led to an increased dependence on anaerobic glycolysis which rendered the resistant cells more sensitive to glycolysis inhibition. Collectively, this comprehensive multi proteomic approach enabled deep insight into signaling recovery and the molecular consequences of resistance development.

Zusammenfassung

Die von Kinasen vermittelte Proteinphosphorylierung von Serin-, Threonin- und Tyrosin-Seitenketten ist eine essentielle, posttranslationale Modifikation, die einen Großteil der physiologischen und krankheitsassoziierten, zellulären Signalverarbeitung steuert. In den letzten Jahren hat sich die Massenspektrometrie-basierte Proteomik zur Kerntechnologie für die großangelegte, explorative und Proteom-weite Untersuchung des Kinoms und assoziierter Phosphorylierungsvorgänge entwickelt.

Aufgrund der geringen Stöchiometrie müssen phosphorylierte Peptide vor der eigentlichen Messung angereichert werden. In diesem Zusammenhang beschreibt Kapitel 2 dieser Arbeit die Entwicklung einer Methode, die auf einer mit Eisenionen beladenen Metallchelate-Affinitätschromatographie (IMAC) Säule basiert. Dadurch war es möglich, die Gesamtheit an Phosphopeptiden chromatographisch, ohne Bevorzugung bestimmter physikochemischer Eigenschaften, in selektiver Art und mit hoher Reproduzierbarkeit aus komplexen, proteomischen Verdauen zu isolieren. Die Methode zeichnete sich durch eine effiziente Elution der gebundenen Phosphopeptide aus, skalierte linear mit der Menge an geladenem Ausgangsmaterial und übertraf alle anderen getesteten Anreicherungsverfahren (Titanium-IMAC, Titaniumdioxid). In Verbindung mit der erhöhten analytischen Tiefe durch eine Fraktionierung angereicherter Phosphopeptide mithilfe der hydrophilen Anionenaustausch-Chromatographie ist die Methode flexibel auf viele experimentelle Fragestellungen anwendbar. Darüber hinaus stellte die Studie die verbreitete Meinung, dass unterschiedliche Anreicherungsmaterialien komplementäre Phosphopeptidspezies isolieren, in Frage. Stattdessen legten die Daten nahe, dass die beobachtete Komplementarität durch eine Kombination aus Formatunzulänglichkeiten, ineffizienter Elution und unzureichender massenspektrometrischer Datenaufnahmegeschwindigkeit hervorgerufen wurde.

Trotz intensiver Bemühungen ist die Proteom-weite Kartierung von Phosphorylierungsstellen noch immer unvollständig. Daher lohnt es sich nach Alternativen zur Ergänzung bestehender Methoden zu suchen. In Kapitel 3 werden Matrix-unterstützte Laser-Desorption/Ionisierung (MALDI)- und Nano-Elektrosprayionisation (nESI)-Tandem Massenspektrometrie (MS/MS) hinsichtlich ihrer Fähigkeit, unterschiedliche Phosphopeptidspezies zu identifizieren, verglichen. MALDI-MS/MS begünstigte die Detektion saurer und phosphotyrosin-haltiger Phosphopeptide und wies dabei eine Orthogonalität auf, die mit der Verwendung alternativer Proteasen wie Asp-N, Arg-C, Chymotrypsin, Glu-C oder Lys-C in Kombination mit nESI vergleichbar war. Im Hinblick auf Phosphoproteomstudien stellt MALDI-MS/MS daher eine nützliche Ergänzung zu nESI-MS/MS dar.

Da die Kinaseaktivität in phosphoproteomischen Experimenten nur indirekt abgeleitet werden kann, beschäftigt sich Kapitel 4 dieser Arbeit detailliert mit der Frage, ob immobilisierte Kinaseinhibitoren dazu geeignet sind, Kinasen aktivitätsabhängig zu binden. Dieses Konzept würde eine direkte Aktivitätsbestimmung ermöglichen. Dazu wurden durch Pervanadatbehandlung hervorgerufene, zelluläre Aktivitätsänderungen von Kinasen zunächst mithilfe der quantitativen Phosphoproteomik bestimmt und dann mit der differentiellen Kinasebindung in einem chemoproteomischen Ansatz verglichen. Es hat sich gezeigt, dass eine aktivitätsabhängige Anreicherung von Kinasen nur sehr selten vorkam und dabei (i) von der individuellen Kinase, (ii) der eingesetzten immobilisierten Inhibitormatrix und (iii) dem zellulären Kinaseaktivitätsstatus abhing. Daher sollte die Bestimmung der Kinaseaktivität über chemoproteomische Methoden auf gut charakterisierte Fälle beschränkt bleiben und kann nicht generisch auf das komplette Kinom angewendet werden.

Das letzte Kapitel dieser Arbeit kombiniert Methoden, die in vorangegangenen Kapiteln entwickelt, implementiert und untersucht wurden. Es beschreibt eine proteomische, chemoproteomische und phosphoproteomische Untersuchung der Wirkweise des Kinaseinhibitors Lapatinib und der Entwicklung von Resistenz gegen Lapatinib in einer ERBB2 überexprimierenden Brustkrebszelllinie. Die Ergebnisse deckten mehrere deregulierte, therapeutisch relevante Zielmoleküle (wie zum Beispiel eine verstärkte Aktivität von CDK1/2) und Phänotypen (wie zum Beispiel eine erhöhte Invasivität) auf. Ein Vergleich der Daten aus dieser Studie mit zuvor beschriebenen Resistenzmechanismen gegen ERBB2 gerichtete Therapien, legte eine große Heterogenität nahe, zeigte aber zugleich, dass die Abhängigkeit von der Glycolyse ein gemeinsames Resistenzmerkmal darstellt. Mechanistisch gesehen ging die Resistenz mit einer Veränderung der Phosphorylierung und in Folge dessen auch der Aktivität von LDHA und PDHA1 einher. Dies führte zu einer Erhöhung der anaeroben Glycolyse und einer gesteigerten Sensitivität der resistenten Zellen gegenüber Glycolyseinhibitoren. Zusammenfassend lässt sich festhalten, dass dieser umfassende, multiproteomische Ansatz eine tiefgehende Analyse der Signalwiederherstellung und einen Einblick in die molekularen Konsequenzen der Lapatinib-Resistenzentstehung ermöglicht hat.

Table of contents

Abstract		iii
Zusammenfassung		v
Chapter 1	General introduction	1
Chapter 2	Comprehensive and reproducible phosphopeptide enrichment using iron immobilized metal ion affinity chromatography (Fe-IMAC) columns	44
Chapter 3	MALDI-TOF- and nESI-Orbitrap-MS/MS identify orthogonal parts of the phosphoproteome	71
Chapter 4	Evaluation of kinase activity profiling using chemical proteomics	94
Chapter 5	Multi proteomic dissection of lapatinib mode of action and resistance in ERBB2 overexpressing breast cancer	120
Chapter 6	General discussion	149
Appendix		164

Chapter 1

General introduction

1 Cellular signaling in health and disease

1.1 Protein phosphorylation and kinases

The central dogma of molecular biology, first coined by Francis Crick in 1957 (1), describes the flow of information in any given cellular system. In very simple terms it means: “DNA makes RNA makes protein”. More elaborately put, it describes the directional flow of biological information which is originally encoded in DNA and contains the instructions for the synthesis of messenger RNA (a process called transcription), which in turn is translated into functional proteins. Over the course of this process, the molecular complexity and number of potential molecules originating from one protein coding gene vastly increases. As an example, alternative splicing of mRNA during transcription causes one gene to express multiple different proteins, so called isoforms. Yet, the biggest explosion in diversity occurs only after translation, when proteins are proteolytically processed and amino acid side chains are decorated by a diverse set of modifications (Figure 1). This causes the proteome to be almost three orders of magnitude more complex than the genome would suggest. The recently introduced term “proteoform” tries to capture this complexity and is defined as “all of the different molecular forms in which the protein product of a single gene can be found, including changes due to genetic variations, alternatively spliced RNA transcripts and post-translational modifications” (2).

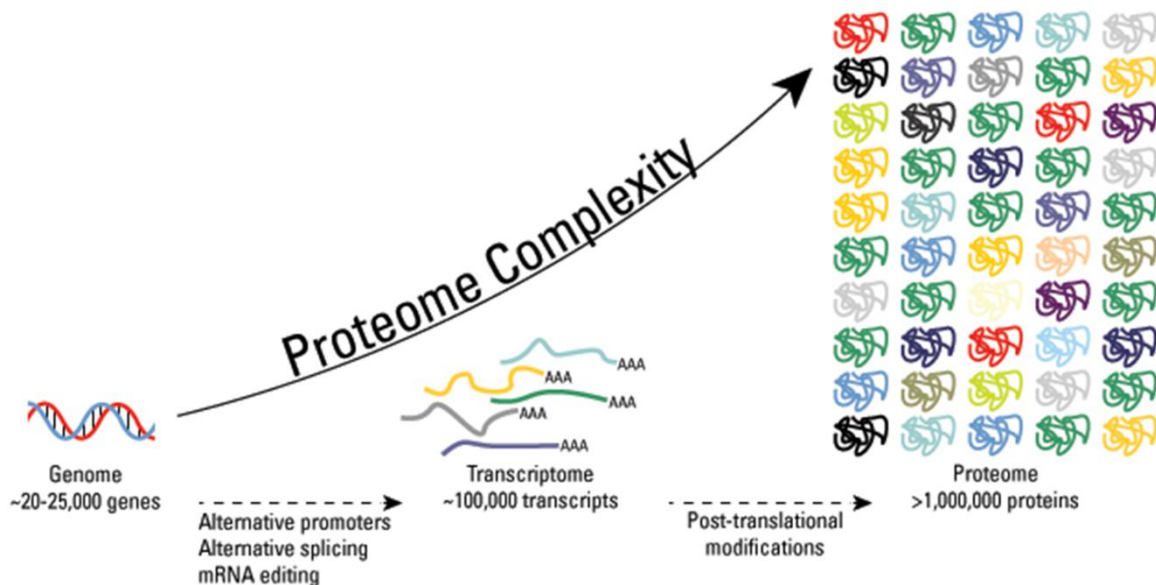


Figure 1. Genes are transcribed into mRNA which is subsequently translated into proteins. During this process, the molecular complexity vastly increases which is predominantly caused by posttranslational modification of amino acid side chains (©Thermo Fisher Scientific).

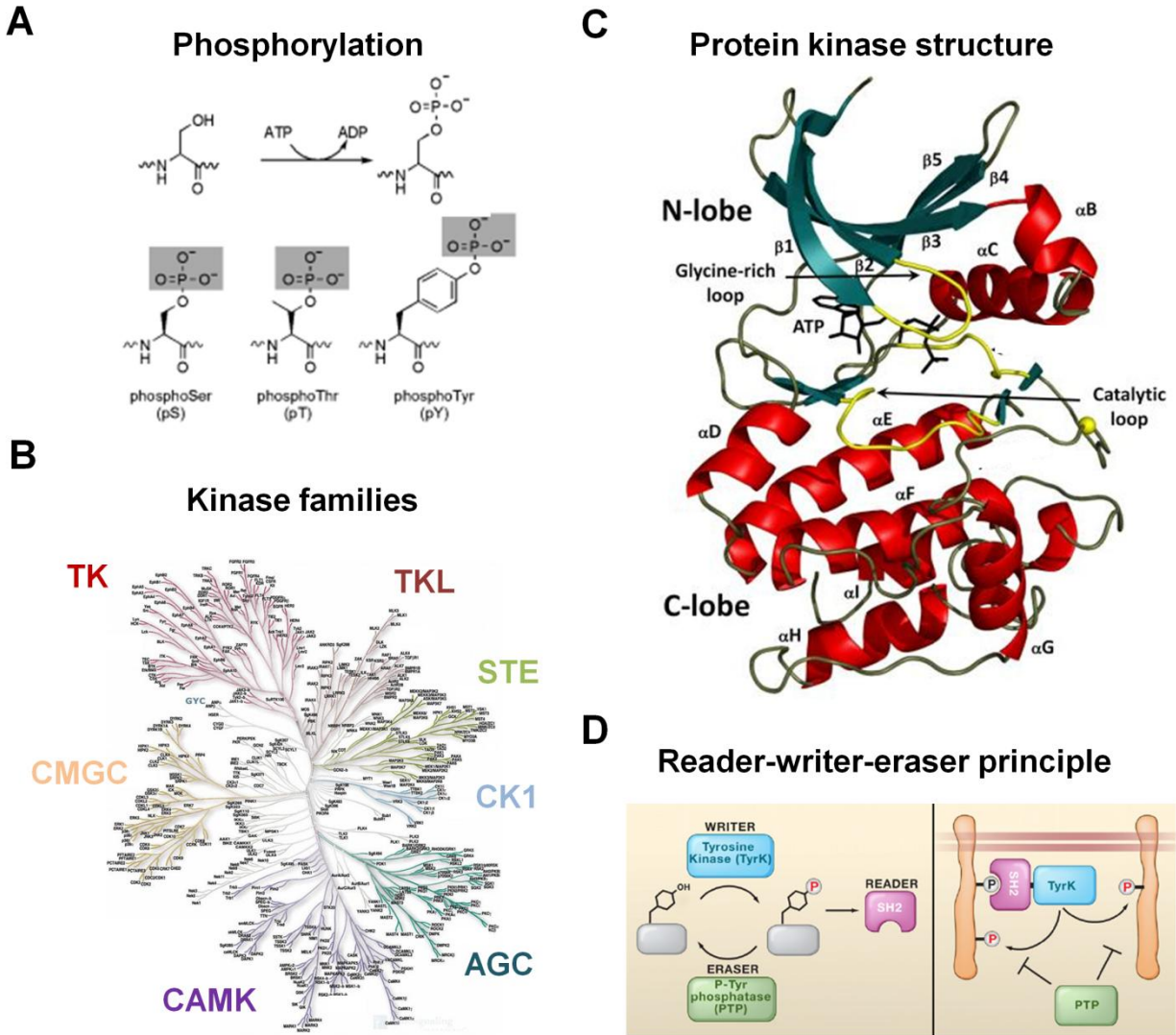


Figure 2. Phosphorylation and kinases. (A) The upper panel shows the transfer of the terminal, negatively charged phosphate group from ATP to the neutral hydroxyl side chain of serine. The lower panel depicts phosphorylated serine, threonine and tyrosine side chains (adapted from Walsh et al. (4)). (B) Genetic homology of the catalytic domain defines seven major protein kinase families which are arranged in a kinome tree (5) (© Cell Signaling Technology). (C) The crystal structure of the PKA kinase domain reveals a bilobal composition (C- and N-lobe) with several universal sequence motifs such as the “glycine-rich loop” and the “catalytic loop”. The phosphate donor ATP is located in the catalytic cleft between the N- and the C-lobe. (D) Reversible phosphorylation signaling follows a simple and modular “reader-writer-eraser” system which is the composing principle of higher order signaling cascades (adapted from (6)).

The number and diversity of the more than 200 currently known PTMs, which are in aggregate capable of modifying 15 out of the 20 proteinogenic amino acid side chains, suggest widespread regulatory roles and underscore the importance to study their molecular function (3, 4).

With more than half of all proteins being phosphorylated at least once (7) and a recent study extrapolating this to be true for over 90% of all proteins (8), phosphorylation is one of the most prevalent and exhaustively studied PTMs in eukaryotes. The main reasons for this are its biological relevance, its stability, the availability of potent affinity reagents for enrichment and suited analytical methods for detection. Phosphorylation is the covalent addition of a charged γ -phosphate group from the high energy donor ATP to the side chain of serine, threonine and tyrosine residues, which occurs in a proportion of roughly 86 : 12 : 2 (Figure 2A) (9). From a chemical perspective, the introduction of a phosphate group transforms a hydrophilic and, in case of serine and threonine, moderately-sized side chain into a negatively-charged and bulky one (4). This in turn affects protein conformation which is frequently accompanied by changes in protein activity, protein-protein interaction, proteolytic stability, protein half-life and localization (10). The phosphate transfer is catalyzed by 518 protein kinases which represent about 1.7 % of the human genome and are divided into 40 atypical and 478 typical kinases (5). Based on sequence comparisons of the catalytic domain, general sequence similarity and overlapping biological functions, the 478 typical protein kinases can be hierarchically arranged and grouped into seven major families which are classified as AGC-, CAMK-, CK1-, CMGC-, STE-, TK- and TKL-kinases (Figure 2B). Despite their diverse biological roles, the basic architecture of protein kinases is remarkably conserved: a catalytic cleft between the smaller amino-terminal lobe (N-lobe) and the carboxy-terminal-lobe (C-lobe) serves as a docking station for the phosphate donor ATP. Both segments are connected by the hinge region which forms hydrogen bonds with ATP (11). The N-lobe comprises a set of common structural kinase features such as the glycine-rich loop which is involved in ATP positioning, a lysine residue which is indispensable for catalytic activity and a flexible C_{α} -helix which connects to different structural parts of the kinase. The C-lobe contains the majority of catalytically important residues, such as the conserved DFG- and APE- motifs located in the activation loop. Usually, phosphorylation of activation loop residues induces a structural rearrangement of the loop which facilitates substrate binding and induces a catalytically active conformation (12).

In contrast to the common capability of phosphate transfer which requires conserved structural features, substrate recognition by kinases necessitates the accurate detection of specific target residues in the background of 100,000s of potential phosphorylation sites. In addition to cellular localization, substrate residue accessibility and spatial proximity, this remarkable substrate specificity is structurally achieved through a variable binding region within the C-lobe (13). The amino acids surrounding the target phosphorylation site are often complementary to this region, which greatly enhances the affinity and thus confers direct specificity. For the serine/threonine

kinase PKA this substrate consensus sequence or motif reads R-R-X-S/T- θ (X denotes a random amino acid and θ a hydrophobic one). Both positively charged side chains at the -2 and -3 position (relative to the phosphorylation site) form ionic bonds with two negatively charged glutamate residues provided by PKA. Moreover, the PKA active site contains a hydrophobic cleft which interacts with the hydrophobic side chain located C-terminal of the substrate target site (13, 14, 15). The physicochemical nature of these motifs broadly classifies kinases into proline-directed, basophilic and acidophilic (16).

In 1992, Edmond Fischer and Edwin Krebs were jointly awarded the Nobel Prize in Physiology or Medicine “for their discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism” (18, 19). This reversibility is a prerequisite for signal modulation and highly dynamic integration and transformation of external and internal stimuli (19). In concert with recognition domains, such as Src Homology2 (SH2) domains, which bind phosphorylated tyrosine residues, the regulation of dynamic phosphorylation processes follows a modular reader-writer-eraser system (21, 22): while kinase “writers” mark residues, binding domains recognize them as “readers”, and phosphatases are the “erasers”, capable of hydrolytically removing those marks (Figure 2D) (21, 17). In humans, 100 tyrosine-, 40 serine/threonine- and 50 dual specificity phosphatases are known (23, 24, 25). Via the complementary interplay of protein kinases and phosphatases - more precisely of their catalytic activity, specificity, affinity for different substrates, their cellular abundance and their spatiotemporal occurrence - the protein phosphorylation status in a cell is tightly controlled (8, 24). An additional level of regulation is indirectly introduced by “reader” domains. They not only recognize the absence or presence of phosphorylated sites, but they are also capable of preventing de-phosphorylation by masking target residues (8, 16, 20). As such, they are also capable of contributing to substrate specificity: a promiscuous kinase might phosphorylate many substrates, but only those that are bound by readers are protected from de-phosphorylation and survive (10). Although the “reader-writer-eraser” principle is conceptually simple, the diverse nature and complex interplay of those three modules creates the basis for higher order signaling circuits (Figure 2D) (20). As a simple example, a protein which is composed of both a “reader” and “writer” domain is capable of recognition, transfer and amplification of the phosphorylation signal which often leads to positive feedback. Similarly, proteins carrying a “reader” and an “eraser” domain are able to confer negative feedback, which restores the physiological state after signal transmission (6).

1.2 Aberrant signaling in cancer

As discussed above, the basic building blocks of the “reader-writer-eraser” system are pivotal for kinase mediated signaling and their combination enables the construction of complex signaling pathways. Such kinase pathways and the inherent propagation of external stimuli shares several, very fundamental principles: receptor tyrosine kinases which are localized at the cell membrane dimerize and phosphorylate each other upon binding of extracellular ligands. Consequently, adaptor proteins which are recruited via their reader domains are able to trigger kinase signaling cascades that propagate the signal via consecutive phosphorylation. Signals frequently converge on transcription factors leading to modified gene expression, or varied activity of proteins involved in protein synthesis, or cell cycle progression which entails direct biological effects such as altered cellular proliferation or cell division (25, 26). Owing to the fundamental impact on a diverse set of biological processes, kinase signaling has to be tightly controlled and orchestrated. Non-physiological activation at the wrong time or at the wrong place frequently entails proliferative diseases such as cancer. Hence, it is not surprising that many of the postulated hallmarks of cancer, such as sustained proliferation, activated invasion or resistance towards cell death are direct phenotypic consequences of derailed kinase signaling (27). Despite the accumulation of many malignant molecular features during cancer development, the activation of one critical oncogene (which is defined as a gene that has the potential to cause cancer, if overexpressed or activated) can be sufficient to sustain the malignant phenotype in certain cancer types, a concept which is known as “oncogene addiction” (28, 29). Kinases constitute a large part of the roughly 100 known oncogenes and clinically relevant receptor tyrosine kinases include *c-KIT*, *BCR/ABL*, *VEGFR*, *EGFR* and *ERBB2* (29). The transforming capability of mutationally activated *ERBB2* (at that time called *neu*) was initially discovered by Schechter and coworkers in 1986 using rat cells (30). Soon after, Di Fiore *et al.* showed that *ERBB2* needs to be overexpressed rather than mutated in order to act as a potent oncogene in human cells (31). The translation into the clinic was promoted by analysis of 189 primary human breast cancers which identified *ERBB2* amplification in roughly 30 % of the patients and additionally identified its prognostic value for reduced overall survival time and time to relapse (32). These initial observations have subsequently been confirmed and extended in many independent studies and “*ERBB2*-positive” is now considered a distinct and clinically relevant molecular subtype of breast cancer (33, 34). Mechanistically, the overexpression of *ERBB2* triggers dimerization and autophosphorylation even in the absence of a stimulating agent. Next, several adaptor proteins are recruited and propagate the signal to

downstream pathways including PIK3/AKT/mTOR and MEK/ERK. Oncogenic activation is, however, not confined to the receptor level, but also frequently affects cytosolic kinases. As an example, mutations in the catalytic subunit of the lipid kinase Phosphatidylinositol 3-kinase (PIK3CA) were found in up to 40 % of all breast cancers and some of them are known to activate PIK3 independent of upstream stimulation (35, 36, 37). Consequences of uncontrolled signaling activation caused by oncogenic transformations include cell growth, cell survival, acquisition of invasive capabilities and rewired energy metabolism, all of which are essential traits of tumor development and maintenance (27).

1.3 Kinase inhibitors

Owing to their structurally conserved features (see 1.1.), their fundamental role in cancer biology (see 1.2.), and their oncogenic gain-of-function potential, kinases are prime targets for the design of small molecule inhibitors. Initially, antagonistic antibodies, such as cetuximab (Erbix[®]) or trastuzumab (Herceptin[®]), which recognize the extracellular domain of EGFR or ERBB2, were used to prevent receptor tyrosine kinase (RTK) dimerization and concomitant oncogenic signaling (38). An alternative strategy is the direct inhibition of catalytic activity by intracellular inhibition of the kinase's active site with small molecule kinase inhibitors. As the intracellular, catalytic domains of all active kinases catalyze phosphate transfer, they share several structural features which enable effective ATP binding. Those include two hydrophobic pockets, a purine binding site and several amino acid side chains which form hydrogen bonds with ATP. These interaction sites can be exploited for the design of competitive inhibitors which mimic ATP and block the active site. Common to most type 1 inhibitors is a heterocyclic core which fits into the purine binding site and forms hydrogen bonds with the hinge region, a structural kinase motif that connects the C- to the N-lobe (Figure 2C and 3A). Attached to the heterocyclic core are additional chemical groups which reach into two distinct hydrophobic pockets (39, 11). This inhibitor class is often referred to as "DFG-in" because it targets an active conformation where the DFG-motif within the activation loop needs to be orientated inwards to enable aspartate side chain interaction with one of the ATP bound Mg²⁺ ions (Figure 3A). Displacement of the activation loop is primarily achieved via autophosphorylation or phosphorylation by another kinase which in turn induces a conformational change and renders the active site fully accessible (12). In contrast, "DFG-out" inhibitors bind and stabilize the inactive kinase conformation, containing an unphosphorylated activation loop which sterically blocks ATP binding. The inactive state is structurally less conserved and features a unique set

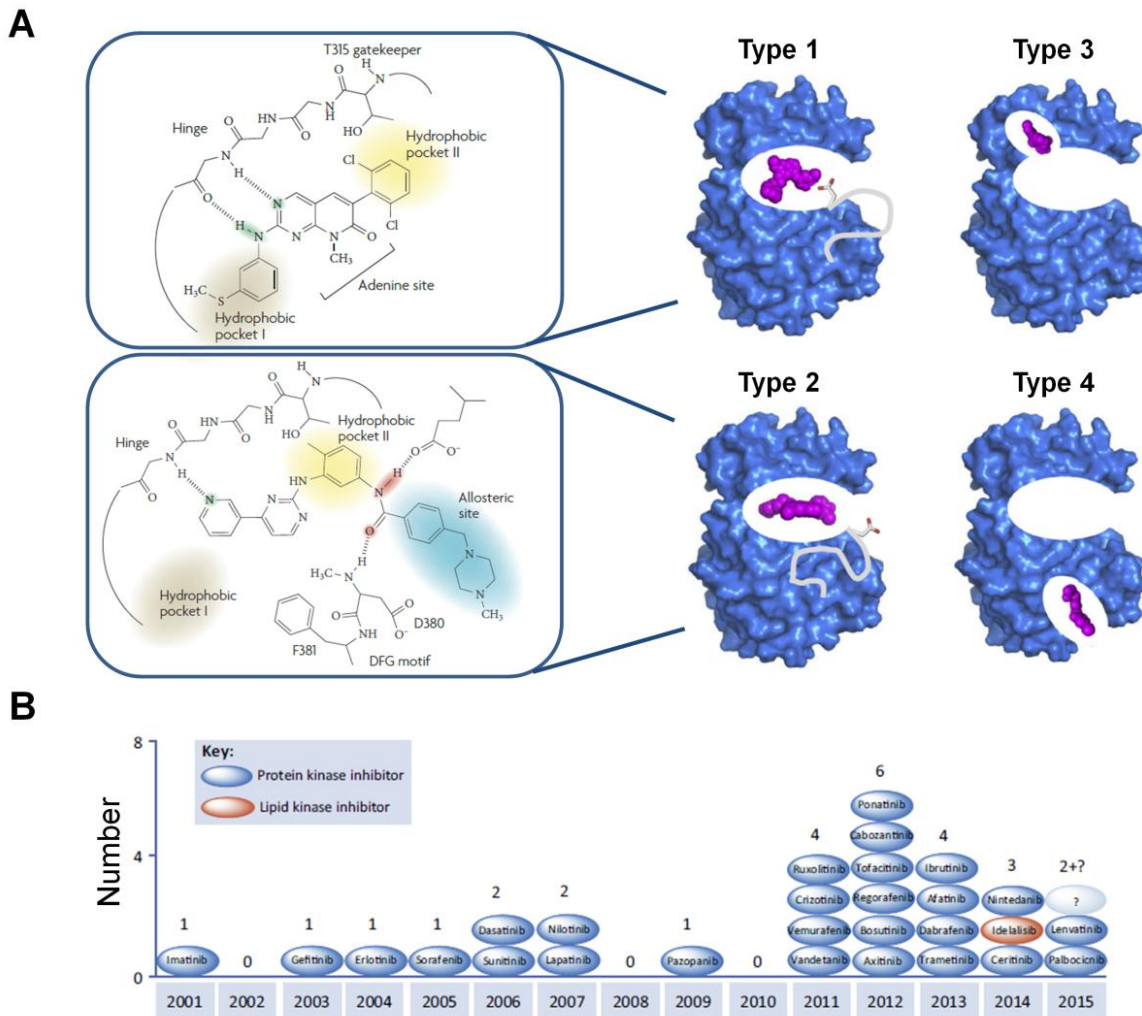


Figure 3. Kinase inhibitors. (A) Reversible kinase inhibitors are classified based on their mode of binding. Type 1 inhibitors bind to a catalytically competent, active kinase conformation where the aspartate side chain of the DFG motif within the activation loop (grey) is oriented inwards (“DFG-in”). In contrast, type 2 inhibitors bind to and stabilize an inactive kinase conformation with a flipped activation loop that sterically blocks accessibility to the active site (“DFG-out”). A type 3 mode of binding is defined as the occupancy of a cleft in close proximity to the active site. If the inhibitor binds to a distant site which influences kinase activity in an allosteric manner, it is classified as type 4. Both insets on the left show unique structural features of the inactive and active kinase conformation which can be exploited for inhibitor design. The chemical structure on the upper left depicts the kinase ABL1 in complex with a type 1 inhibitor (PD166326) that interacts with the hinge region and two different hydrophobic pockets. The conformational change of the activation loop in the inactive kinase state exposes an additional allosteric site (depicted in blue on the lower left of panel A, where the chemical structure shows imatinib binding to ABL1) which can be bound by type 2 inhibitors. In addition, the orientation of the activation loop enables hydrogen bonding with the glycine side chain of the DFG motif (adapted from Zhang et al. (39) and Wu et al. (42)) (B) In April 2015, 28 kinase inhibitors were FDA approved for clinical use. The majority is directed towards tyrosine kinases, one is a lipid kinase inhibitor (idealisib), one is type 3 (trametinib) and two are irreversible (afatinib and ibrutinib). Remarkably, half of those inhibitors have been approved in the last three years (adapted from (42)).

of hydrogen bonding options and an additional hydrophobic pocket which can be utilized for inhibitor design (Figure 3A) (39). Although type 1 inhibitors target a common structure and type 2 inhibitors a flexible one, recent studies suggest that type 2 inhibitors are not *per se* more selective than type 1 inhibitors (40, 41). Type 3 inhibitors bind to a hydrophobic pocket close to the active site and inhibitors of the fourth class target structural features unique to the respective kinases which modulates activity in an allosteric manner (Figure 3A) (42). In contrast to the reversible nature of inhibitor classes 1 to 4, irreversible inhibitors contain a Michael acceptor system and covalently attach to a nucleophilic cysteine side chain within the active site (43). Irreversible and allosteric inhibitors are, once discovered, remarkably selective for the intended target (44, 39). Given the high degree of structural target similarity, this is different for type 1 and 2 inhibitors which are naturally more prone to inhibit additional kinases. Hence, several methods for kinase selectivity profiling have been developed (41). Recombinant kinase panels featuring only the active site of kinases have, for example, been used to screen 72 inhibitors against 442 catalytic kinase domains (45) and 178 kinase inhibitors against 300 catalytic kinase domains (46). A similar approach was used to screen the Published Kinase Inhibitor Set (PKIS) set containing 367 ATP competitive compounds against 224 recombinant kinases (47). Alternatively, mass spectrometry-based chemical proteomics can be used to immobilize the inhibitor of interest on sepharose beads and enrich targets directly out of native cell lysates (see 2.1.) (48). A recent addition to the chemical proteomics arsenal is the cellular thermal shift assay (CETSA) which monitors the melting behavior of proteins and measures the shift in melting temperature which is induced by the stabilizing effect of drug target engagement (49, 50). Using the promiscuous kinase inhibitor staurosporine, CETSA identified and confirmed more than 50 targets in an unbiased way such that neither the inhibitor nor the target needed to be chemically modified (50).

At the time of writing, 28 small molecule kinase inhibitors were FDA approved and more than 200 inhibitors with diverse chemical scaffolds, pharmacological profiles and primary kinase targets were in clinical trials (Figure 3B) (51, 52). Yet, not even one fifth of the kinases are part of the “targeted kinome” which conversely means that the vast majority of inhibitors is directed towards redundant targets (52, 53). Given their biological role and the frequent deregulation in cancer (which is discussed in 1.2.), EGFR and ERBB2 are amongst the most extensively targeted kinases (52, 54). Two EGFR inhibitors, gefitinib and erlotinib, are FDA approved for the treatment of non-small cell lung cancer, whereas the dual EGFR/ERBB2 inhibitor lapatinib is used to treat metastatic, ERBB2 overexpressing breast cancer. Many more single and dual target agents, such as the irreversible afatinib (directed against EGFR/ERBB2) or the

pan-ERBB family inhibitor poziotinib (directed against EGFR/ERBB2/ERBB4), are evaluated in clinical trials for the treatment of various types of cancer (38). Interestingly, ERBB family inhibitors, such as lapatinib which has no other kinase off-targets, are usually remarkably selective. This might either be related to a structurally distinct kinase domain or could simply mean that the ERBB family got more pharmacological attention compared to other kinases (46). In addition to small molecules which inhibit intracellular kinase activity, many monoclonal antibodies which selectively target the extracellular domain of a single ERBB family member or heterodimers thereof have been developed. Clinically approved ones are trastuzumab (blocks ERBB2 homodimerization), pertuzumab (blocks ERBB2 homo- and heterodimerization), cetuximab and panitumumab (both block EGFR dimerization and ligand binding) (38, 55). Targeting pathway nodes downstream of ERBB receptor tyrosine kinases drives additional drug research efforts. Often such nodes acquire independent oncogenic potential due to activating mutations or overexpression, which renders upstream signaling dispensable. As previously mentioned, PIK3 mutations are frequently observed in breast cancer and several PIK3 kinase inhibitors and inhibitors targeting PIK3 downstream nodes such as AKT and mTOR are in clinical development for the treatment of breast cancer (56, 57). Other cancer types are addicted to the RAS/RAF/MEK/ERK pathway which is a second prominent signaling branch located downstream of the ERBB receptor family (38).

1.4 Drug resistance

An important step in fighting cancer with small molecule kinase inhibitors is the identification of targetable and subtype specific oncogenic patterns emerging from the molecular heterogeneity of different tumors (see 1.2). In a personalized medicine setting, patients are screened for those abnormalities and specific inhibitors are used to block the identified, oncogenic signaling nodes in order to stop tumor cell proliferation (see 1.3). Despite an initially high response rate, tumor plasticity and signaling redundancy frequently cause acquisition of resistance against the inhibitor, which in turn leads to tumor relapse (61, 62). Hence, the identification of resistance mechanism, which might prioritize beneficial combination therapies, constitutes another important challenge. In addition to more general mechanisms based on increased drug efflux, decreased drug influx or enhanced drug decomposition, mechanisms of acquired kinase inhibitor resistance can be divided into four major categories, which are hereafter introduced, primarily in the context of resistance to ERBB family targeted therapies (Figure 4) (63, 61, 64). The first mechanism involves mutation or amplification of the primary target. A typical example

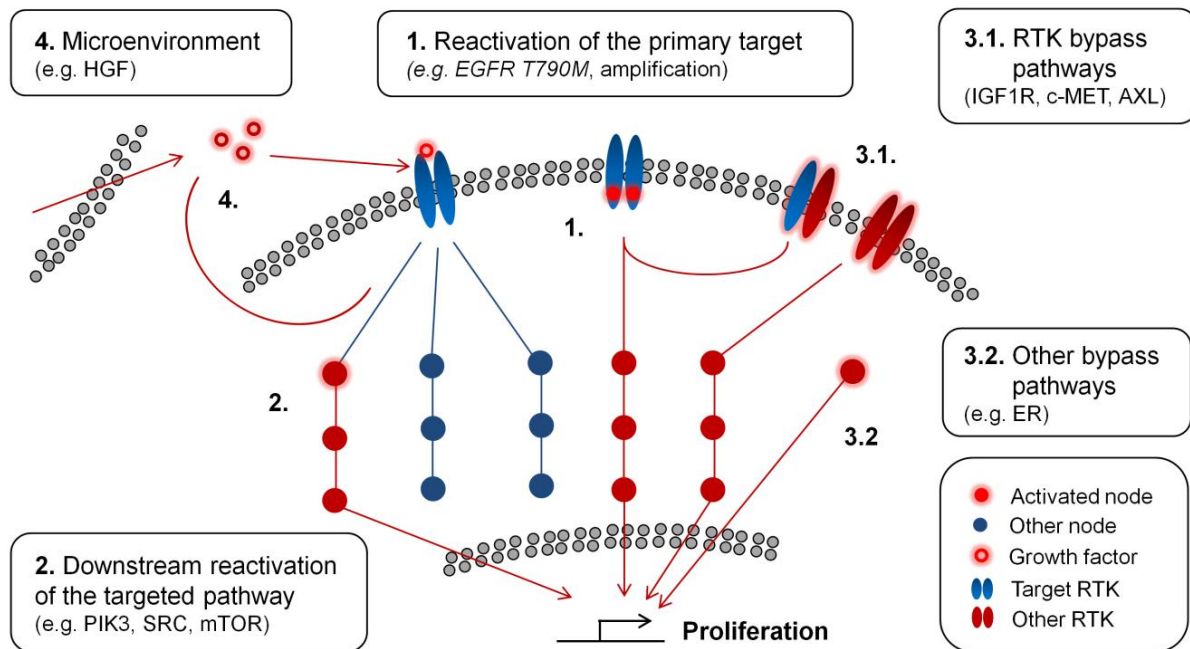


Figure 4. ERBB kinase inhibitor resistance. Despite initial tumor cell response, resistance to kinase targeted therapies frequently occurs. The scheme depicts the four major mechanisms which mediate drug resistance.

is the discovery of a T790M mutation at the “gatekeeper” position of EGFR which confers resistance to gefitinib by increasing the affinity for ATP (65, 66, 67). Since the mutation does not significantly impair primary target signaling, inhibitors with alternative binding modes are often able to break resistance and thus provide an efficient means to overcome primary resistance (68). Other studies have shown that EGFR amplification, ERBB2 T798I mutation and mutation within the EGFR ectodomain (S492R) or alternative splicing of ERBB2 (truncated p95ERBB2), are other possible, albeit less frequently occurring alterations of the primary target (70, 71, 72, 73). The second major cause of resistance is the reactivation of upstream or downstream effectors within the initially targeted pathway. As an example, the mutation of PIK3 or the loss of PTEN causes resistance to ERBB2 targeting agents by activation of the downstream PIK3/AKT/mTOR pathway which is able to fuel cellular proliferation independently of extracellular stimuli (74, 75). Other ERBB2 related examples include the downregulation of cyclin-dependent kinase inhibitor CDKN1B, upregulation of SRC or activation of mTOR, all of which are important downstream mediators of initial response (76, 77, 78). The compensatory upregulation of alternative signaling pathways is the third and probably the most heterogeneous way of resistance acquisition. Receptor tyrosine kinases can be activated and drive parallel pathways leading to sustained proliferation independent of the initial target, a process known as

“oncogenic drift”. The activation of AXL, c-MET and EPHA2 in ERBB2 and EGFR driven tumors are just a few of the known examples (79, 80, 81, 82, 83, 84). In addition to homodimerization, resistance mediating RTKs such as IGFR or other structurally related ERBB family members are also frequently transactivated by heterodimerization with EGFR or ERBB2 (85, 86). Further downstream of RTKs, resistance can be mediated by the compensatory activation of a parallel pathway. Indeed, extensive cross talk between the PIK3/AKT/mTOR and the MEK/ERK pathway often allows the tumor cells to switch their proliferative dependency in case one of the two branches is inhibited (87, 88). Moreover, the activation of bypass signaling is not confined to kinases: a common feature of resistance in ERBB2 driven breast cancer is the activation of estrogen receptor (ER) signaling (89, 90). Upon ligand induced activation, the ER mediates transcription of genes involved in proliferation which establishes a ERBB2/ER codependence (88). Similarly, the transmembrane protein integrin-b1 is a non-kinase which has been shown to cause resistance by stimulation of FAK and SRC (90). The fourth and last way of kinase inhibitor resistance is mediated by the tumor microenvironment. Two recent studies outline how paracrine or autocrine secretion of certain growth factors can immediately rescue tumor cells from kinase inhibitor induced cell death (92, 93). One such mechanism is the activation of c-MET-RTK signaling by HGF, which causes resistance to BRAF inhibitors (91).

Together these findings also highlight the inherent redundancy of RTK signaling which provides a flavor of how widespread the means of resistance acquisition might be. Once the underlying resistance mechanism is identified, the therapeutic strategy is to target the aberrantly activated signaling node in order to restore sensitivity towards the initial target or kill the cells in case they are solely addicted to the bypass pathway.

2 Mass spectrometry-based proteomics

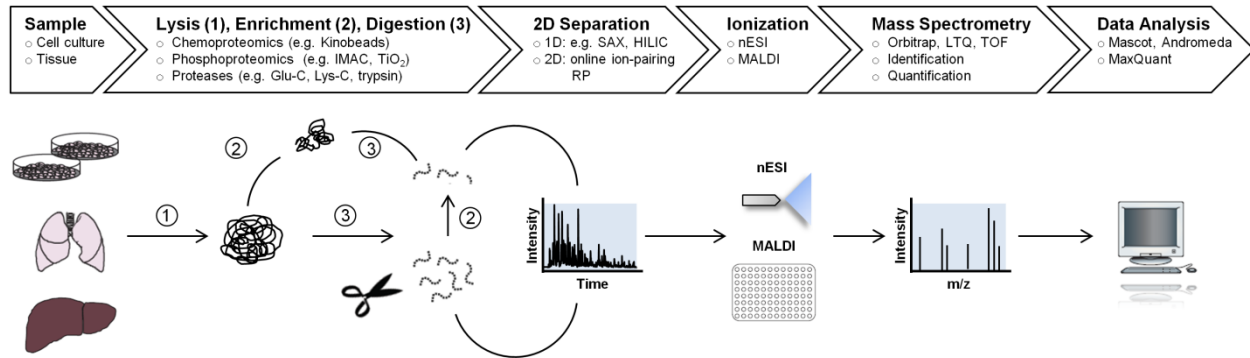


Figure 5. Typical bottom-up proteomics workflow (inspired by (93) and (94)).

The proteome is defined as the full complement of proteins which is present in a biological system at any given point in time. As such, the proteome is not static but the protein abundance, their post-translational modifications, their structure and interactions are dynamically regulated in response to a variety of external and internal stimuli. Proteomics, the large-scale study of the proteome, complements genomic and transcriptomic approaches, but is unique in the sense that it enables the direct analysis of the executing molecules within a cell and is in turn capable of providing a different level of biological insight and understanding (95). Since its inception in the 1990s, staggering technological advances have established mass spectrometry as the central tool for the qualitative and quantitative study of proteomes and have recently cumulated in first draft maps of the entire proteome, cataloguing the origin and abundance of more than 17,000 proteins (5, 94, 95). Strengths of mass spectrometry include automatization possibilities (online peptide separation and measurement), exquisite sensitivity (< femtomol) and rapid sequencing speeds (> 10,000 sequences per hour), which in combination enable accurate measurements of a large part of the proteome within only a few hours (96, 97, 98). In bottom-up (or shotgun) proteomics, mass spectrometers measure the mass-to-charge ratio of intact peptides which are subsequently isolated, fragmented and identified by matching the experimentally observed to theoretically calculated fragment masses. A typical workflow covers protein extraction from cells or tissue, proteolytic digestion into peptides (typically using trypsin), high-resolution peptide separation (one or two dimensional), peptide ionization (by electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI)), tandem mass spectrometry and bioinformatic data analysis (Figure 5). As the complexity of whole proteome digests is still too high for the analytical capabilities of current mass spectrometers, the standard

workflow can be modified such that an orthogonal, second dimension of peptide separation is introduced, or such that sub-proteomes are isolated either prior to (e.g. by protein affinity enrichments; e.g. chemoproteomics) or after digestion (e.g. by the selective enrichment of different PTMs; e.g. phosphoproteomics). The following paragraphs provide a more detailed introduction into the techniques employed in the respective workflow steps and put special emphasis on methods which are particularly relevant for this thesis.

2.1 Peptide separation

Peptide separation

The bottom-up proteomic analysis of cell line and tissue samples to a depth greater than 10,000 proteins still represents an analytical challenge because of the sheer number of peptides generated by proteolytic digestions and the high dynamic range of protein expression. High resolution, two dimensional peptide separation/fractionation is an efficient means to boost proteome coverage, sequence coverage, peak capacity and quantification performance. The fundamental chromatographic principle involves reciprocal interactions of an analyte, a stationary phase and a mobile phase. Peptide analytes are dissolved in a mobile phase and passed through a column which contains a stationary phase, usually in the form of polymerized monoliths or small particles which carry functional groups. If the analyte does have no attraction to the stationary phase, it will pass the column essentially at the speed of the mobile phase. However, if the analyte-stationary phase interaction is stronger than the analyte-mobile phase interaction, the analyte will be attracted to the stationary phase which causes retention. In practical terms, the stationary phase is kept constant, whereas the mobile phase composition is varied in a gradient over time. Hence, physicochemically distinct peptide species can be separated in timely manner by sequential elution off the column.

Given the high resolving power and volatile solvent components compatible with online coupling to the mass spectrometer, the stationary phase used in nanoflow-LC-MS/MS setups is almost exclusively comprised of reversed-phase (RP) material which separates peptides mainly based on hydrophobicity. In the typical nanoflow ion-pairing reversed-phase chromatography setup, μm -columns (I.D. of 50-100 μm) are packed with silica beads (1-5 μm) which are functionalized with hydrophobic C18 alkyl chains. The mobile phase (usually a mixture of water and acetonitrile) is supplemented with an amphiphilic acid (usually formic acid or trifluoroacetic acid) that attaches to the hydrocarbon chains of the stationary phase with its hydrophobic part. The hydrophilic group confers interactions with polar side chains of the peptide analytes causing

additional retention and increased resolution. Finally, elution based on hydrophobicity is achieved by increasing the concentrations of the organic solvent acetonitrile in the mobile phase.

The first separation dimension should ideally be orthogonal to ion-pairing reversed-phase and offer high chromatographic resolution. Widely applied methods include peptide separation by charge, as in strong cation exchange (SCX) (98) or strong anion exchange (SAX) (99), or peptide separation by hydrophilicity/hydrophobicity such as in hydrophilic interaction liquid chromatography (HILIC (100)) or high-pH reversed-phase separation (101). But also electrostatic repulsion hydrophilic interaction chromatography (ERLIC) (102), zwitterionic HILIC (ZIC-HILIC) (103), weak anion exchange (WAX) (104) and hydrophilic strong anion exchange (hSAX) (105) have proven their merit for the separation of complex digests.

2.2 Sub-proteome enrichment

Protein enrichment

In addition to multidimensional separation, the selective enrichment of sub-proteomes is another facile means to reduce complexity and to study a subset of proteins or peptides which might otherwise perish in the background of a complete proteome. In immuno-affinity purification, antibodies, either directed against a specific bait protein or against an epitope artificially attached to the bait, are immobilized on beads and used to enrich proteins out of native lysates. Coupled to LC-MS/MS, this methodology enables the identification, quantification and stoichiometry determination of high affinity protein complex members which are co-purified alongside the protein of interest (107, 108, 109). Two recent studies applied this principle in a large scale setup in order to systematically map and explore a large part of the human protein interaction landscape within a single cell line (110, 111).

Chemical compounds which are conjugated to biochemical handles (e.g. biotin) or directly immobilized on sepharose beads can also be used as affinity tools in combination with protein identification by mass spectrometry. Since such chemical proteomics approaches often require the synthesis of linkable compound analogues, proper information about structure activity relationship (SAR) is important to ensure retained biological activity and a target engagement which is comparable to the original compound (111). However, once accomplished, native targets of a variety of immobilized probes can be enriched out of lysate which provides invaluable information for inhibitor target and mode-of-action analysis (48, 112, 113). Moreover, the application of unselective compounds or a mixture thereof enables the application of the

chemoproteomic principle to the enrichment of whole protein families including histone deacetylases (HDACs) (113), ADP/ATP binders (114) and kinases (116, 117). If the latter is configured as a competition binding assay, the kinome is enriched by broad specific inhibitor beads and targets are dose dependently outcompeted by the free compound which is spiked into the lysate (117, 118). This generic approach enables screening of inhibitors against the full length, cofactor bound, posttranslationally modified kinases and other ATP binding molecules (116, 117, 119, 120, 121).

Phosphopeptide enrichment

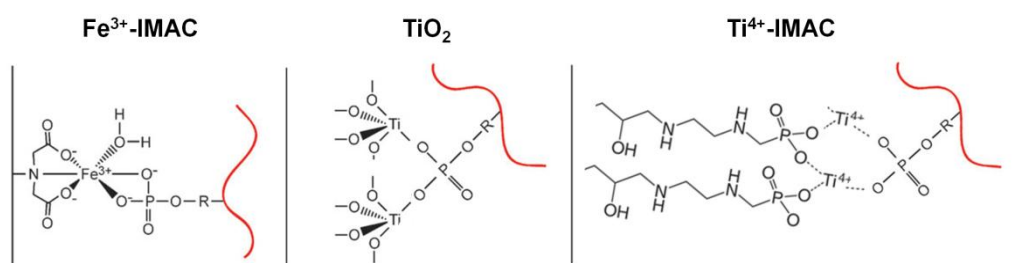


Figure 6. Scheme of phosphopeptide interaction with different, generic enrichment materials (adapted from (121), (122) and (123)).

With the rationale of increased selectivity, the enrichment of posttranslational modifications is almost exclusively conducted on the peptide level. Enrichment is required due to the low abundance and sub-stoichiometric levels of most posttranslational modification events. A viable strategy, which is in principle applicable to a wide variety of stable PTMs and only limited by the availability of suitable reagents, is the use of antibody-based affinity enrichments. These have, for instance, been successfully applied to the large scale enrichment of ubiquitin, acetyl and tyrosine phosphorylation modifications (124, 125, 126). Also, peptide-centric antibody approaches targeting a specific kinase motif have recently been employed (122, 123, 124) and led to the identification of kinase substrates that were not identified in conventional, large scale studies (129). Although immunoaffinity reagents are also available for phosphoserine and -threonine modifications, more efficient methods, which make use of the coordination/chelation capabilities of immobilized metal ions like Fe^{3+} or Ga^{3+} towards phosphate groups, exist. Notably, these methods also enrich phosphotyrosine containing peptides, but the proportionally low occurrence (~1-2 %) occludes efficient detection in the presence of serine/threonine phosphorylation.

In immobilized metal affinity chromatography (IMAC), metal ions are chelated either via iminodiacetic acid (IDA) or nitrilotriacetic acid (NTA) and phosphopeptides are retained by the formation of a coordinative bond with the free orbitals of the bound metal ion (125, 126) (Figure 6). The use of low pH solvents causes protonation of negatively charged carboxyl side chains which reduces unspecific binding to the metal ions. Still, the co-purification of acidic peptides cannot be prevented entirely which leads to decreased selectivity of IMAC enrichments. Subsequent elution of the phosphate group is typically achieved by competitive displacement with hydroxide ions. Metal oxide affinity chromatography (MOAC), which uses titaniumdioxide (TiO_2 (132)) or zirconiumdioxide (ZrO_2 (133)) is the other widely applied method for the generic enrichment of phosphopeptides (Figure 6). High selectivity towards phosphopeptides can be achieved by adding organic acids such as 2,5-dihydroxy-benzoic acid (DHB) or lactic acid (129, 130). Finally, a new generation of IMAC material, where metal (IV) ions (Zr^{4+} , Ti^{4+}) are coordinated by phosphate or phosphonate groups has been developed over the last years (131, 132) (Figure 6). It is generally assumed that there is a high degree of complementarity between these different enrichment materials and that the combination enables the purification of phosphopeptide species with distinct physicochemical characteristics (132, 133, 134, 135). Combinations of phosphopeptide enrichment methods and two dimensional fractionation techniques now enable routine identification of several 10,000s of non-redundant phosphorylated peptides per sample. Initially, enrichment was preceded by a fractionation step, which decreases the sample complexity. In strong cation exchange chromatography (SCX), performed at low pH, peptides are separated based on their solution net charge and coordination to the negatively charged stationary phase. The majority of peptides generated from a tryptic digest carry a 2+ net charge (138, 139). Under low pH conditions, unlike acidic amino acids, phosphate groups on the peptide carry a negative charge, reducing the net charge of a peptide. Consequently, the phosphopeptides elute earlier which separates them from the majority of non-phosphorylated peptides. However, a large number of phosphopeptides contain multiple basic residues and therefore still elute in the later fractions, together with non-phosphorylated peptides. Therefore the SCX separation is often followed by affinity chromatography such as IMAC and TiO_2 . Kettenbach *et al.* (134) introduced an alternative approach, in which TiO_2 batch enrichment is performed prior to phosphopeptide enrichment, directly on a complex tryptic digest. This single stage purification enhanced the reproducibility of the workflow at similar phosphoproteome coverage and concomitantly reduced wet lab time. Other fractionation approaches which have been combined with phosphoproteomics include HILIC (142), ERLIC (143) and high-pH reversed-phase separation (144).

2.3 Mass spectrometry

Figuratively speaking, a mass spectrometer is comparable to a molecular scale which is used to weigh the mass of proteins and peptides. It utilizes the fundamental principle that charges can be manipulated by electromagnetic fields in the gas phase and is composed of three main parts: an ion source where analyte molecules are charged and transferred into the gas phase, a mass analyzer where ions are separated based on their mass-to-charge (m/z) ratio and a detection system which measures the resulting ion current (145).

Soft ionization techniques

The ability to ionize intact biomolecules by soft ionization techniques such as matrix-assisted laser desorption/ionization (MALDI) (144, 145, 146) and electrospray ionization (ESI) (150) has initiated the field of mass spectrometry-based proteomics. In MALDI, analyte molecules are co-crystallized with an excess of aromatic matrix molecules which contain a delocalized π -electron system and efficiently absorb laser light of a specific wavelength. Short and pulsed irradiation with such lasers stimulates resonant excitation of the π -electron system and the subsequent relaxation of the energy back into the crystal lattice leads to rapid gas phase desorption of both the matrix molecule and the analytes. The concomitant cooling process prevents thermal decomposition and preserves large biomolecules (151). Regarding the ionization process, which is still not fully understood, two models exist: the “lucky survivor”- model states that ions are

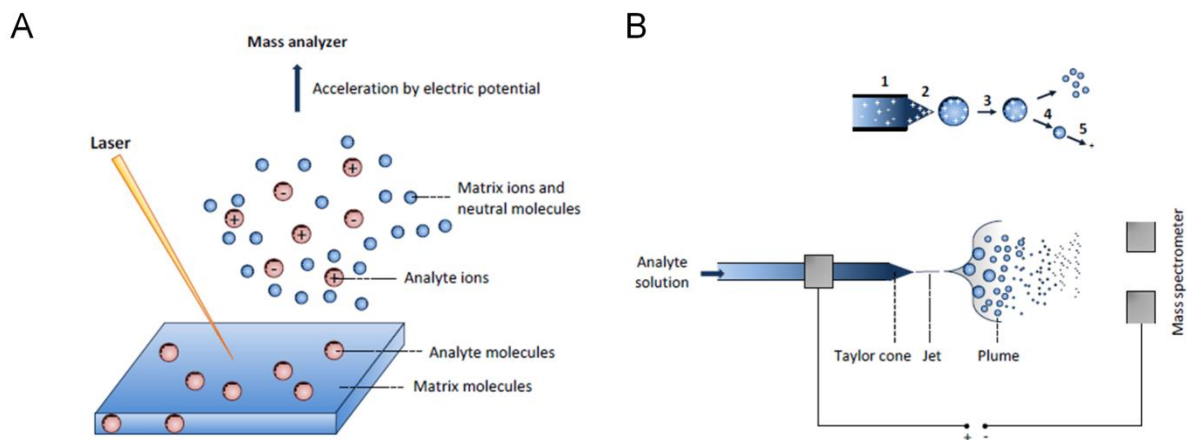


Figure 7. *Soft ionization techniques applied in mass spectrometry. (A) MALDI process. A short laser pulse leads to rapid desorption of the matrix and analyte molecules into the gas phase. (B) ESI process. Upper panel: (1) analytes dissolved in volatile solvent, (2) Taylor cone, droplet formation, (3) solvent evaporation, (4) droplet fission and Coulomb explosion, (5) formation of gas phase ions (adapted from (146) and (93)).*

already present in the crystal lattice, whereas the “gas-phase ionization” model postulates that ionization occurs only in the gas phase after collisions with e.g. matrix molecules. Rather than being exclusive, recent research suggest that both processes are equally responsible for ionization (152).

Owing to the ionization from liquid phase and its efficient, robust and highly automatable online-coupling to peptide separation, ESI has become the most prevalent ionization technique in modern mass spectrometry (Figure 7). Here, the analyte is dissolved in a volatile solvent and passed through a thin metal capillary to which an electrostatic potential is applied. Once the solvent reaches the tip of the capillary, charges are electrophoretically separated and, if the electrostatic potential is greater than the surface tension, a stable Taylor cone appears and emits a continuous liquid jet towards the counter electrode. During this process, the liquid jet becomes unstable and bursts into a spray of multiply charged droplets which contain the dissolved peptide molecules. As the droplets continue to migrate through the electrostatic field, the liquid evaporates and the surface charge increases until the Raleigh limit is reached. In a process called Coulomb explosion, the repulsion of like-charges induces the uneven fission of droplets into nano-droplets. The final analyte ionization is either achieved in a passive process where the droplet charges end up on the analyte once the solvent completely evaporated (charge residue model (153)) or in an active process where accumulated peptide molecules on the surface of the droplets are extracted by field desorption (ion emission model (154)). The majority of the emerging analyte ions is multiply charged. In addition to conventional ESI, which is usually operated with microliter flows, nanoESI introduced nanoliter flow rates and smaller emitter sizes, which cumulatively enhances the efficiency of analyte ionization due to smaller initial droplet sizes. Consequently, analytical sensitivity is increased and less analyte molecules are required for detection (155).

Mass analyzers

Once the peptides are ionized, they enter the mass analyzer which acquires their m/z values and separates them accordingly. Mass analyzers come in different forms and shapes and each type has its unique characteristics regarding resolution (ability to separate two adjacent m/z signals), mass accuracy (how close is the measured value to the true mass), dynamic range (difference between the smallest and the biggest signal within one spectrum), scan speed (how fast can a spectrum be acquired) and covered m/z range (see Table 1). Time-of-flight (TOF) mass analyzers measure how long accelerated ions with the same initial kinetic energy but different masses need to travel a fixed distance within a field-free vacuum. The measured

time-of-flight is directly proportional to the square root of the analytes m/z value. Quadrupole mass analyzers consist of four rods which create a quadrupolar electrical field. Application of a radio frequency facilitates oscillation of the field and allows only ions with a specific m/z value (or a defined range thereof) to have stable trajectories through the rods. In addition to the use as a plain mass filter, application of potential barriers at the end of the rods enables ions to be stored in the quadrupole, a concept which is implemented in linear (2D) ion trap mass analyzers.

Table 1. Acquisition characteristics of frequently employed mass analyzers.

	TOF	Quadrupole	2D ion trap	Orbitrap
Mass accuracy	2 - 10 ppm	0.2 - 0.5 Da	100 - 1,000 ppm	< 1 ppm
Resolution	> 30,000	< 2,000	200 - 20,000	> 500,000
m/z range	> 500,000	< 4,000	< 4,000	< 2,000
Acquisition				
speed	< MHz	1 Hz	< 50 Hz	< 50 Hz
Dynamic range	1 : 10^3	1 : 10^4	1 : 10^3	1 : 10^4
Sensitivity	++	+	+++	+++

The orbitrap is a compact mass analyzer which combines high scanning speeds with high resolution and high mass accuracy. It consists of three essential components: two outer, cup-shaped electrodes and a central spindle electrode which holds the orbitrap together. Application of a voltage to outer and inner electrodes creates a radial electrical field which causes ion attraction to the center of the trap. Initial tangential velocity creates an opposing centrifugal force which cumulatively leads to circular ion movement around the central rod (156). An additional axial field which is caused by the canonical shape of the electrodes causes the orbiting ions to oscillate alongside the central rod. Of the complex ion trajectories only the axial movement is independent on initial ion energies, angles or positions, but is directly proportional to the ion's m/z value (157). Hence, the frequency of axial movement is recorded in an image current and is subsequently translated from the time domain into the m/z domain by Fourier transformation (158). Modern mass spectrometers are not necessarily restricted to only one mass analyzer. Popular combinations include quadrupole-TOF, linear ion trap-orbitrap and the more recently introduced quadrupole-orbitrap instruments.

Tandem mass spectrometry

Tandem mass spectrometry is the arrangement of multiple consecutive scan and selection events with intermediate fragmentation. In a typical schedule of operation, a MS1 scan, where the m/z values and intensities of intact peptide precursors are recorded, is followed by isolation and fragmentation of a precursor and the acquisition of the resulting fragment ions in a MS2 scan. Combined information about accurate precursor and fragment ion masses enables determination of the peptide sequence. Either one mass analyzer is used for both operations in a consecutive manner (“tandem-in-time”) or two mass analyzers within the same mass spectrometer acquire spectra in a parallelized fashion (“tandem-in-space”). A widely used, automated and real-time strategy to decide on which precursor to pick for fragmentation, is data-dependent acquisition (DDA). In this mode of data collection, a predefined number of precursors is picked based on their abundance in the preceding MS1 scan. The implementation of exclusion and inclusion lists allows to fine tune the acquisition, to exclude contaminants and to preferentially pick peptide species of interest. However, a clear drawback of DDA is the stochastic nature involved in precursor selection which limits reproducibility of detection.

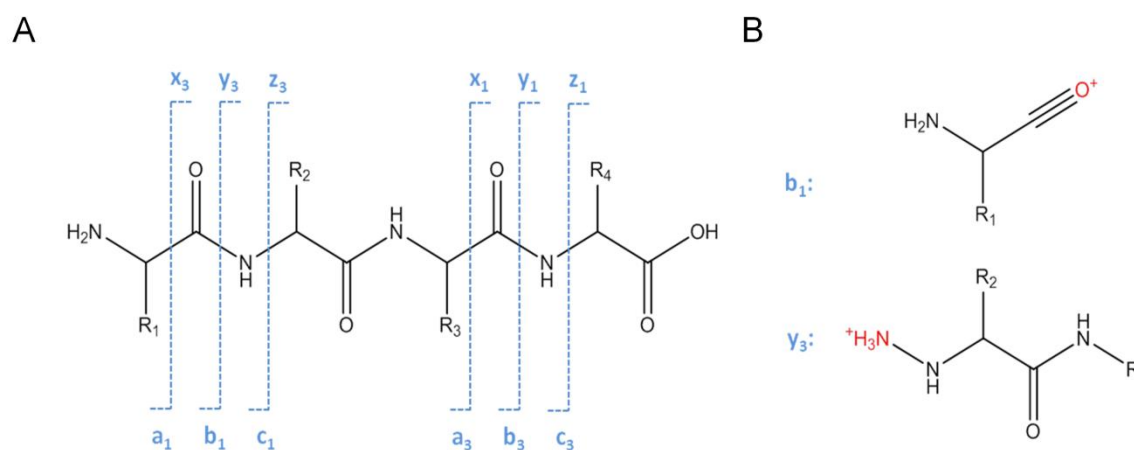


Figure 8. (A) Roepstorff Fohlmann nomenclature (159) of sequence specific peptide fragments. *a*-, *b*- and *c*- ions contain an intact N-terminus; *x*-, *y*- and *z*-ions contain an intact C-terminus. (B) Example for *y*- and *b*- fragment ions resulting from peptide bond breakage.

As the principle of tandem mass spectra crucially depends on peptide fragmentation, several methods such as the commonly applied collision induced dissociation (CID), higher energy C-trap dissociation (HCD (160)) and electron transfer dissociation (ETD (161)) have been developed. In the conceptually similar CID and HCD type fragmentation, isolated peptide species acquire vibrational energy by multiple collisions with inert gas molecules (e.g. N₂ or He).

This energy leads to evenly distributed breakage of the peptide bond and produces sequence-informative b- and y-ions (see Roepstorff Fohlmann nomenclature (159), Figure 8). Additional information is provided by internal fragments, neutral losses (e.g. ammonium and water) or immonium-ions originating from amino acid side chains (162). In modern hybrid mass spectrometers of the linear ion trap-orbitrap type, CID and concomitant spectrum acquisition is performed in a low resolution linear ion trap, whereas HCD is performed in a dedicated collision cell which enables flexible transfer of ions to an orbitrap mass analyzer where high resolution fragment spectra are acquired (163).

The mechanism of ETD fragmentation is fundamentally different compared to CID and HCD. An anion radical donates an electron to the analyte which reduces the charge of the peptide. The unpaired electron is highly unstable and mediates rapid decay of the peptide along the N-C_α-bonds, producing c- and z- type ions (164). A major advantage of ETD for the study of phosphorylation events is that it preserves the labile modification, which would usually detach upon CID and HCD fragmentation. Hence, no intensity consuming and spectrum dominating neutral loss of phosphoric acid (H₃PO₄) is observed and the phosphorylation site localization is improved. However, doubly charged precursors which constitute the majority of all tryptic peptides are often only inefficiently fragmented. Recently, EThcD was introduced as a hybrid fragmentation method which combines the complementary ion series of ETD and HCD and thus leads to higher sequence coverage, more site determining ions and improved phosphorylation site assignment (165).

2.4 Protein identification

As discussed above, the sequence of peptides is derived from fragment ion spectra generated by tandem mass spectrometry. Rather than *de novo* sequencing (166) or spectral library searching (167), the typical mass spectrometry workflow relies on comparison of measured fragment masses to theoretically calculated ones, an approach referred to as "database searching" (168). The theoretical search space is generated by a proteolytic *in silico* digest of proteins collected in a database. In addition to the "naked" peptides, versions containing delta masses of specified modifications are included. Several parameters such as precursor and fragment mass tolerance or number of proteolytic missed cleavages introduce additional constraints and keep the search space within limits. After noise reduction, de-isotoping and charge state deconvolution, the experimentally determined fragment masses are compared to those calculated from the search space. A score which is based on metrics such as number and type of matching fragments or fragment and precursor mass deviation is used to probabilistically

describe the quality and similarity of identified spectra. Usually only the highest scoring hit is reported as the peptide spectrum match (PSM). Frequently employed search engines for the identification and scoring of tandem mass spectra are Andromeda (169), Mascot (170) and Sequest (171). Importantly, bottom-up mass spectrometry is peptide- and not protein-centric. Hence, it infers the presence of a protein solely based on the identification of one or several unique peptides. If the same peptide sequence occurs in more than one protein, the assignment might be ambiguous (“protein-inference problem” (172)), which can complicate correct quantification and biological conclusions. Due to this ambiguity, often protein groups rather than single protein identifications are reported.

A similar problem is encountered when dealing with PTMs such as phosphorylation. Isobaric peptides often contain several potential phosphorylation sites. In conjunction with the mainly incomplete fragment ion information, confident phosphorylation site assignment has proven to be challenging. Several different computation approaches and scoring systems which facilitate automated phosphorylation site assignment and quality assessment of localization have been developed. Examples include the PTM score (169) (Andromeda), the MD score (173) (Mascot) and phosphoRS (174) (Proteome Discoverer). A recent study benchmarked these algorithms by the use of synthetic phosphopeptides which also enabled the calculation of false localization rates (FLR) by comparison of the assigned to the true site of phosphorylation. Besides, the study showed that all these methods share a substantial amount of orthogonality which suggests room for further improvement (175).

Given the sheer amount of spectra acquired in each bottom-up experiment and the size of the *in silico* search space, the occurrence of false positive spectrum matches is a common problem of scoring algorithms (168). The score is just a measure of how likely it is that the assignment occurs by random chance (i.e. the score is a transformed *p*-value) and thus even high scoring matches have a certain probability of being a false positive hit. As it is practically impossible to manually validate all spectra, other approaches, such as the automated and widely applied target-decoy search, have been developed (176). Briefly, a decoy database is derived by reversing or scrambling the protein sequences present in the target database. Although the size and general composition of both databases is similar, spectral matches with sequences derived from the decoy database are per definition wrong. Scoring spectra against both search spaces enables the calculation of a false discovery rate (FDR in percent) which reflects the ratio of decoy and target hits and thus the tolerated (albeit not visible) number of false positive assignments in the target space. Importantly, FDR filters are usually applied consecutively at the PSM, peptide and protein level (168).

2.5 Protein quantification

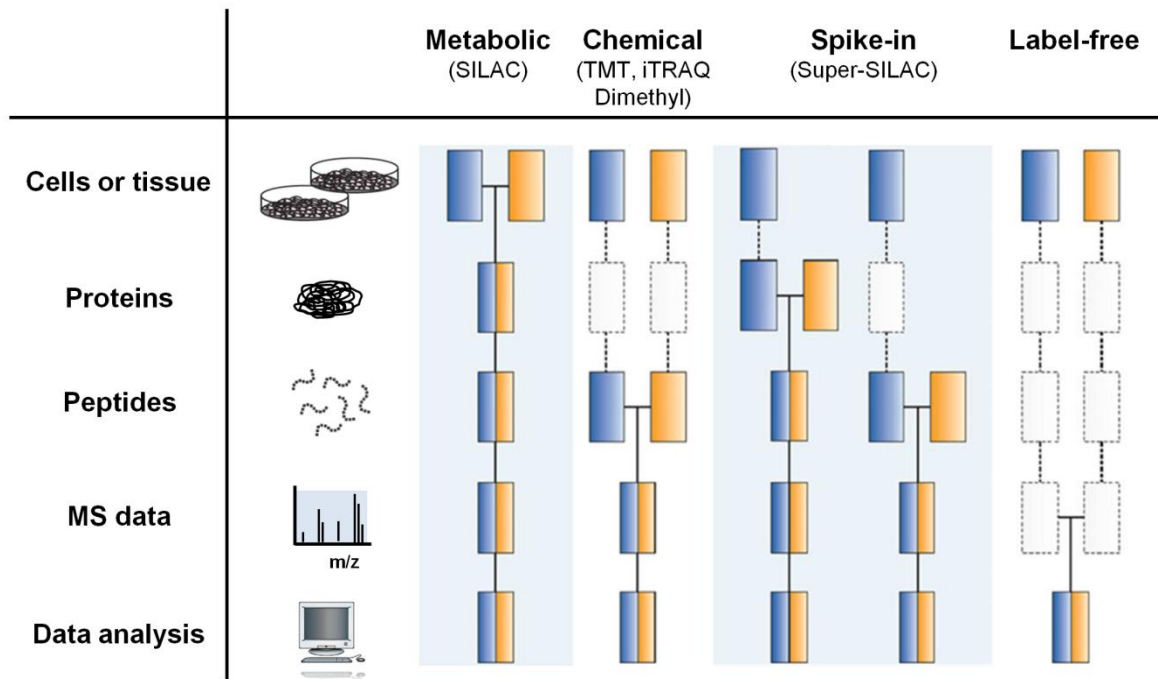


Figure 9. Summary of common labeling and sample combination approaches for relative quantification in bottom-up proteomics (adapted from (177)). Blue and yellow boxes reflect different experimental conditions and the lines indicate at which point in the workflow the samples are combined.

Very broadly speaking, one can distinguish relative and absolute quantification approaches. The latter aims to determine the exact concentration and copy number of a given peptide or protein present in a biological sample whereas the former tries to quantify relative changes. Quantification by mass spectrometry is based on the fact that the response the analyte triggers in the mass detector is proportional to its abundance. Mainly due to differences in ionization efficiency, which are caused by differences in physicochemical peptide characteristics, and an bias introduced by sample handling (e.g. preferential digestion of certain peptides over others), only the signals from the same peptides can be compared to each other. In label-based quantification techniques, the introduction of isotopes or chemical tags enables the mass-encoded multiplexing of peptides from different samples. The resulting mass increment of the labeled over the non labeled version can be discriminated in the mass analyzer and is used for relative quantification within the same spectrum. Depending on the employed labeling technique, samples are combined at different steps of the workflow (see Figure 9) (175, 176). Early combination decreases independent accumulation of technical variation and in turn enhances the resolution of biological changes.

MS1-based quantification

In metabolic labeling (the most commonly used method is stable isotope labeling by amino acids in cell culture (SILAC)), isotopically labeled amino acids are introduced at the level of living cells which makes it the most accurate of all available quantification methods. Since all tryptic peptides carry either an arginine or a lysine, SILAC media are usually supplemented with heavy lysine (e.g. $^{13}\text{C}_6$ -lysine) and arginine (e.g. $^{13}\text{C}_6^{15}\text{N}_4$ -arginine) to ensure comprehensive labeling (179). However, clear drawbacks of SILAC are the increased complexity of MS1 spectra, its limited applicability (e.g. labeling of human tissue samples is not possible), the time spent for label incorporation (usually at least five cell doublings are required) and the restriction to the comparison of only three experimental states. The latter point was recently addressed by NeuCode, an extension of the SILAC principle which makes use of the high resolution of modern mass spectrometers to resolve the subtle mass change caused by nuclear binding energy differences of a neutron present in heavy nitrogen compared to a neutron present in heavy carbon (the so-called “mass defect”). The use of isotopologues rather than isotopes increases the multiplexing capacity considerably (e.g. $^{13}\text{C}_6^{15}\text{N}_2$ -lysine and $^2\text{H}_8$ -lysine differ by 36 mDa which can be resolved at a resolution of 200,000) (180). The use of Super-SILAC mixes is a mean to harness SILAC for the analysis of tissues. Here, SILAC encoded proteins which are derived from cells with a similar biological background are spiked into different tissue proteomes. Although those peptides serve as a common reference, both the introduction only at the level of extracted proteins and the subsequent calculation of ratios from ratios severely impairs the quantitative accuracy compared to metabolic SILAC labeling (181).

A cheaper, universally applicable, but also less accurate alternative to SILAC is the introduction of isotopically labeled dimethyl tags at the peptide level. Primary amines at the N-termini and lysine side chains are converted to dimethylamines by different combinations of formaldehyde and cyanoborohydrides isotopes. This method enables multiplexing of three states that differ in mass by 4 Da (182).

Peptides can also be quantified entirely label-free using separate mass spectrometry runs. Label free quantification is the least accurate approach because samples are combined only *in silico*. Moreover, each sample has to be measured individually which requires both an increased amount of MS time and a highly reproducible sample preparation and peptide measurement. Still, clear benefits are the possibility to compare an unlimited number of samples, the “native” nature of the mass spectra (no artificial chemical noise or increased complexity) and the highest analytical depth out of all available quantification methods.

MS2-based quantification

Up to this point, all discussed methods are based on MS1 intensity quantification of the intact peptide ions. In contrast, MS2-based methods employ chemical labeling with primary amine targeting tags which are attached via succinimide chemistry. Due to their isobaric nature, differentially labeled species cannot be distinguished in MS1 scans. However, upon fragmentation they lose isotopically encoded reporter ions which are detected on the low mass region of the MS2 spectrum and can be used for relative quantification. The most widely used approaches make use of tandem mass tags (TMT) (183) or isobaric tags for absolute and relative quantification (iTRAQ) (184) which enable comparison of up to six (TMT) or eight (iTRAQ) different experimental states without an increase in MS1 complexity. Recently, the resolution of mass differences between isotopologues has increased the multiplexing capability of TMT tags up to ten (185). A drawback of such MS2-based approaches is the decreased dynamic quantification range which is mainly caused by ratio compression due to co-eluting and co-isolated peptide species (186).

The choice of which method to use ultimately depends on the specific experiment and the biological question at hand. If one wants to compare many different samples with large differences in peptide and protein abundance, label-free quantification might be employed. On the other hand, quantification of very subtle quantitative changes upon perturbation of the same biological system might require the superior precision of metabolic labeling.

2.6 Bioinformatic data analysis

Given the throughput and quantitative data points state-of-the-art proteomics experiments deliver these days, bioinformatic data analysis is gaining tremendous momentum. Fortunately, several software suites, such as MaxQuant (187) with its implemented search algorithm Andromeda (169), have been developed to automatize the data processing steps and to integrate useful features at the same time. Exemplified for MaxQuant, the user defines the experimental design, loads the acquired raw files and specifies the desired parameters (e.g. labeling method, digestion enzyme, modifications or mass tolerance). First, features which represent the mass and the intensity of the measured peptides are extracted from the MS1 spectra. The m/z values are recalibrated and if several samples are processed together, the total intensity of peptides present in different samples is normalized to each other. A useful addition is the match-between-runs option which increases quantitative proteome coverage by alignment of features measured in different runs based on retention time and accurate mass.

Even if only one of those aligned features triggers an identification, the others can still be matched by inference. Subsequently, the MS2 spectra are processed (e.g. noise reduction, de-isotoping, charge state deconvolution), recalibrated and prepared for searches against the specified database. The integrated search engine scores the identifications and filters the dataset in a target-decoy approach according to user defined PSM, peptide and protein FDR levels. If phosphopeptides were recorded, an additionally computed metric indicates the likelihood of phosphorylation site localization at the annotated peptide side chain (“localization probability”). Alternatively, one can judge the quality of site assignment based on the delta score between the highest ranking phosphoisomers, a measure that recently was benchmarked, adjusted and translated into FLRs (175). Next, peptides are classified in unique and razor peptides (which are peptides shared by two or more proteins/protein isoforms) and are used to infer protein identity and abundance. Once the proteins, peptides and phosphorylation sites in a sample have been measured, identified and quantified, higher order information can be retrieved from the dataset. The first goal is usually the identification of quantitative changes between experimental states. Often, p -values, which is the percentual chance of making the wrong call on a classified change, are computed for each observation and a user defined cut-off is chosen (e.g. 1 % or 5 %). Changing (phospho)proteins can then be connected with tools that score protein-protein interaction based on literature derived knowledge such as STRING (188). Associated interaction modules can be further classified according to GO terms or matched to different pathways collected in the KEGG database (187, 188). Tools enabling automated annotation and enrichment analysis of large-scale datasets include PANTHER (191) and DAVID (192).

With the increasing number of phosphoproteomic studies performed and the growing number of phosphopeptides identified per study, bioinformatic tools to validate the localization of the phosphorylation site have become indispensable (see also 2.3). At the same time, phosphorylation events have to be linked to the kinases and phosphatases involved in the (de)phosphorylation process. In large-scale phosphoproteomics, phosphorylation sites are often described without having knowledge about the responsible kinase, which is very much different from the classical approach in which substrates were identified from a specific kinase as starting point. Prediction of kinase motifs is essential to link phosphorylation events to the upstream kinase and various tools are available that subtract and predict motifs from large-scale datasets (193, 194, 195, 196). Tools that assess the complexity of signaling events and integrate data by computational approaches have become evenly important (197, 198). Finally, collection of the identified phosphosites in public repositories allows for the easy retrieval of experimental

verified phosphorylation sites and, if known, their implications on protein function (199, 200, 201).

3 Objectives

Kinase mediated phosphorylation is an important post translational modification which affects and governs a large body of cellular signaling in health and disease. The mass spectrometry-based study of the kinome and associated phosphorylation events has emerged as the prime tool for the explorative study of systematic signaling changes upon perturbation or aberrant rewiring and has enabled tremendous and invaluable biological insights. Although extensive effort has been made, phosphorylation site mapping is still incomplete and thus merits exploration of alternative methods complementing existing workflows. Moreover, unbiased and reproducible methods for the efficient and robust enrichment of phosphopeptides on a large scale and in a short time frame are needed. Such phosphoproteomic characterization can for example also be used to infer a kinase's activation state based on substrate changes and thus reveal clinically actionable targets. However, substrate changes are only an indirect measure, require knowledge of the underlying motifs and are frequently not directly connected to a single kinase. Hence, a method which allows for a direct and unbiased read out of kinase activity would be of tremendous value.

Given the aforementioned points, the main objectives of this thesis were to develop, implement and benchmark a new phosphopeptide enrichment technique embedded in an efficient workflow for routine phosphoproteomic analysis (chapter 2), to increase phosphoproteome coverage by means of different ionization techniques (chapter 3), to evaluate if chemoproteomic methods are suited for direct, selective and unbiased determination of kinase activity (chapter 4) and to apply the established methodology to the system-wide dissection of chemo- and phosphoproteomic signaling changes upon the development of drug resistance (chapter 5).

Abbreviations

ADP	Adenosine diphosphate
APE-motif	Alanine-phenylalanine-glutamate-motif
ATP	Adenosine triphosphate
CETSA	Cellular thermal shift assay
CID	Collision-induced dissociation
CRM	Charge residue model
DDA	Data dependent acquisition
DFG-motif	Aspartate-phenylalanine-glycine-motif
DNA	Deoxyribonucleic acid
ERLIC	Electrostatic repulsion hydrophilic interaction chromatography
ESI	Electro-spray ionization
ETD	Electron transfer dissociation
FDA	Food and Drug Administration
FDR	False discovery rate
FLR	False localization rate
FT	Fourier transformation
HCD	Higher energy C-trap dissociation
HILIC	Hydrophilic interaction chromatography
hSAX	Hydrophilic strong anion exchange
I.D.	Inner diameter
IDA	Iminodiacetic acid
IMAC	Immobilized metal affinity chromatography
iTRAQ	Isobaric tags for relative and absolute quantification
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Liquid chromatography
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry
LTD	Linear trap quadrupole
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption/ionization
MDscore	Mascot delta score
MOAC	Metal oxide affinity chromatography
mRNA	Messenger ribonucleic acid

MS	Mass spectrometer
MS	Mass spectrometry
MS/MS	Tandem mass spectromerty
MS1	Precursor mass spectrum
MS2	Fragment mass spectrum
NTA	Nitrilotriacetic acid
ppm	Parts per million
pS, pT, pY	Phosphoserine, -threonine, -tyrosine
PSM	Peptide spectrum match
PTM	Post-translational modification
RNA	Ribonucleic acid
RP	Reversed phase
RTK	Receptor tyrosine kinase
SAR	Structure activity relationship
SAX	Strong anion exchange
SCX	Strong cation exchange
SH2	Src Homology 2
SILAC	Stable isotope labelling with amino acids in cell culture
TiO ₂	Titaniumdioxide
TMT	Tandem mass tag
TOF	Time-of-flight
WAX	Weak anion exchange
XIC	Extracted ion chromatogram
ZIC-HILIC	Zwitterionic hydrophilic interaction chromatography
ZrO ₂	Zirconiumdioxide

4. References

1. Crick, F. (1970) Central dogma of molecular biology. *Nature* 227, 561–563
2. Smith, L. M., Kelleher, N. L., and Proteomics, T. C. for T. D. (2013) Proteoform: a single term describing protein complexity. *Nat. Methods* 10, 186–187
3. Walsh, C. (2006) *Posttranslational modification of proteins: expanding nature's inventory* (Roberts and Co. Publishers, Englewood, Colo)
4. Walsh, C. T., Garneau-Tsodikova, S., and Gatto, G. J. (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed Engl.* 44, 7342–7372
5. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The Protein Kinase Complement of the Human Genome. *Science* 298, 1912–1934
6. Lim, W. A., and Pawson, T. (2010) Phosphotyrosine Signaling: Evolving a New Cellular Communication System. *Cell* 142, 661–667
7. Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J.-H., Bantscheff, M., Gerstmair, A., Faerber, F., and Kuster, B. (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509, 582–587
8. Sharma, K., D'Souza, R. C. J., Tyanova, S., Schaab, C., Wiśniewski, J. R., Cox, J., and Mann, M. (2014) Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.* 8, 1583–1594
9. Olsen, J. V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127, 635–648
10. Lim, W., Mayer, B., and Pawson, T. (2014) *Cell Signaling: principles and mechanisms* (Taylor & Francis)
11. Alton, G. R., and Lunney, E. A. (2008) Targeting the unactivated conformations of protein kinases for small molecule drug discovery. *Expert Opin. Drug Discov.* 3, 595–605
12. Taylor, S. S., and Kornev, A. P. (2011) Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* 36, 65–77
13. Ubersax, J. A., and Ferrell Jr, J. E. (2007) Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* 8, 530–541
14. Taylor, S. S., Kim, C., Vigil, D., Haste, N. M., Yang, J., Wu, J., and Anand, G. S. (2005) Dynamics of signaling by PKA. *Biochim. Biophys. Acta* 1754, 25–37
15. Zheng, J., Trafny, E. A., Knighton, D. R., Xuong, N. H., Taylor, S. S., Ten Eyck, L. F., and Sowadski, J. M. (1993) 2.2 A refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr. D Biol. Crystallogr.* 49, 362–365
16. Pinna, L. A., and Ruzzene, M. (1996) How do protein kinases recognize their substrates? *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1314, 191–225
17. Fischer, E. H., and Krebs, E. G. (1955) Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J. Biol. Chem.* 216, 121–132
18. Fischer, E. H. (2013) Cellular regulation by protein phosphorylation. *Biochem. Biophys. Res. Commun.* 430, 865–867
19. Cohen, P. (2002) The origins of protein phosphorylation. *Nat. Cell Biol.* 4, E127–E130
20. Pawson, T. (1995) Protein modules and signalling networks. *Nature* 373, 573–580
21. Yaffe, M. B. (2002) Phosphotyrosine-binding domains in signal transduction. *Nat. Rev. Mol. Cell Biol.* 3, 177–186
22. Tonks, N. K. (2006) Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* 7, 833–846
23. Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J., and Mustelin, T. (2004) Protein Tyrosine Phosphatases in the Human Genome. *Cell* 117, 699–711
24. Shi, Y. (2009) Serine/Threonine Phosphatases: Mechanism through Structure. *Cell* 139, 468–484

25. Choudhary, C., and Mann, M. (2010) Decoding signalling networks by mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* 11, 427–439
26. Blume-Jensen, P., and Hunter, T. (2001) Oncogenic kinase signalling. *Nature* 411, 355–365
27. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674
28. Weinstein, I. B. (2002) Addiction to Oncogenes--the Achilles Heal of Cancer. *Science* 297, 63–64
29. Weinstein, I. B., and Joe, A. (2008) Oncogene Addiction. *Cancer Res.* 68, 3077–3080
30. Schechter, A. L., Stern, D. F., Vaidyanathan, L., Decker, S. J., Drebin, J. A., Greene, M. I., and Weinberg, R. A. (1984) The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature* 312, 513–516
31. Fiore, P. D., Pierce, J. H., Kraus, M. H., Segatto, O., King, C. R., and Aaronson, S. A. (1987) erbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. *Science* 237, 178–182
32. Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., and McGuire, W. L. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235, 177–182
33. West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R., Zuzan, H., Olson, J. A., Marks, J. R., and Nevins, J. R. (2001) Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11462–11467
34. Perou, C. M., Sørlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, Ø., Pergamenschikov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A.-L., Brown, P. O., and Botstein, D. (2000) Molecular portraits of human breast tumours. *Nature* 406, 747–752
35. Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S. M., Riggins, G. J., Willson, J. K. V., Markowitz, S., Kinzler, K. W., Vogelstein, B., and Velculescu, V. E. (2004) High frequency of mutations of the PIK3CA gene in human cancers. *Science* 304, 554
36. Campbell, I. G., Russell, S. E., Choong, D. Y. H., Montgomery, K. G., Ciavarella, M. L., Hooi, C. S. F., Cristiano, B. E., Pearson, R. B., and Phillips, W. A. (2004) Mutation of the PIK3CA Gene in Ovarian and Breast Cancer. *Cancer Res.* 64, 7678–7681
37. Samuels, Y., Diaz Jr., L. A., Schmidt-Kittler, O., Cummins, J. M., DeLong, L., Cheong, I., Rago, C., Huso, D. L., Lengauer, C., Kinzler, K. W., Vogelstein, B., and Velculescu, V. E. (2005) Mutant PIK3CA promotes cell growth and invasion of human cancer cells. *Cancer Cell* 7, 561–573
38. Yarden, Y., and Pines, G. (2012) The ERBB network: at last, cancer therapy meets systems biology. *Nat. Rev. Cancer* 12, 553–563
39. Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* 9, 28–39
40. Zhao, Z., Wu, H., Wang, L., Liu, Y., Knapp, S., Liu, Q., and Gray, N. S. (2014) Exploration of Type II Binding Mode: A Privileged Approach for Kinase Inhibitor Focused Drug Discovery? *ACS Chem. Biol.* 9, 1230–1241
41. Müller, S., Chaikuad, A., Gray, N. S., and Knapp, S. (2015) The ins and outs of selective kinase inhibitor development. *Nat. Chem. Biol.* 11, 818–821
42. Wu, P., Nielsen, T. E., and Clausen, M. H. (2015) FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol. Sci.* 36, 422–439
43. Liu, Q., Sabnis, Y., Zhao, Z., Zhang, T., Buhrlage, S. J., Jones, L. H., and Gray, N. S. (2013) Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chem. Biol.* 20, 146–159
44. Leproult, E., Barluenga, S., Moras, D., Wurtz, J.-M., and Winssinger, N. (2011) Cysteine Mapping in Conformationally Distinct Kinase Nucleotide Binding Sites: Application to the Design of Selective Covalent Inhibitors. *J. Med. Chem.* 54, 1347–1355
45. Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., Hocker, M., Treiber, D. K., and Zarrinkar, P. P. (2011) Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1046–1051
46. Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H., and Peterson, J. R. (2011) Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1039–1045
47. Elkins, J. M., Fedele, V., Szklarz, M., Abdul Azeez, K. R., Salah, E., Mikolajczyk, J., Romanov, S., Sepetov, N., Huang, X.-P., Roth, B. L., Al Haj Zen, A., Fouches, D., Muratov, E., Tropsha, A.,

- Morris, J., Teicher, B. A., Kunkel, M., Polley, E., Lackey, K. E., Atkinson, F. L., Overington, J. P., Bamborough, P., Müller, S., Price, D. J., Willson, T. M., Drewry, D. H., Knapp, S., and Zuercher, W. J. (2015) Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* advance online publication,
48. Rix, U., and Superti-Furga, G. (2009) Target profiling of small molecules by chemical proteomics. *Nat. Chem. Biol.* 5, 616–624
 49. Molina, D. M., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., Sreekumar, L., Cao, Y., and Nordlund, P. (2013) Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science* 341, 84–87
 50. Savitski, M. M., Reinhard, F. B. M., Franken, H., Werner, T., Savitski, M. F., Eberhard, D., Molina, D. M., Jafari, R., Dovega, R. B., Kläeger, S., Kuster, B., Nordlund, P., Bantscheff, M., and Drewes, G. (2014) Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* 346, 1255784
 51. Wu, P., Nielsen, T. E., and Clausen, M. H. Small-molecule kinase inhibitors: an analysis of FDA-approved drugs. *Drug Discov. Today*,
 52. Fedorov, O., Müller, S., and Knapp, S. (2010) The (un)targeted cancer kinome. *Nat. Chem. Biol.* 6, 166–169
 53. Fabbro, D. (2015) 25 Years of Small Molecular Weight Kinase Inhibitors: Potentials and Limitations. *Mol. Pharmacol.* 87, 766–775
 54. Tebbutt, N., Pedersen, M. W., and Johns, T. G. (2013) Targeting the ERBB family in cancer: couples therapy. *Nat. Rev. Cancer* 13, 663–673
 55. Maija Hollmén, K. E. (2010) Potential of ErbB4 antibodies for cancer therapy. *Future Oncol. Lond. Engl.* 6, 37–53
 56. Baselga, J. (2011) Targeting the Phosphoinositide-3 (PI3) Kinase Pathway in Breast Cancer. *The Oncologist* 16, 12–19
 57. Engelman, J. A. (2009) Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat. Rev. Cancer* 9, 550–562
 58. Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J., and Norton, L. (2001) Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *N. Engl. J. Med.* 344, 783–792
 59. Geyer, C. E., Forster, J., Lindquist, D., Chan, S., Romieu, C. G., Pienkowski, T., Jagiello-Gruszfeld, A., Crown, J., Chan, A., Kaufman, B., Skarlos, D., Campone, M., Davidson, N., Berger, M., Oliva, C., Rubin, S. D., Stein, S., and Cameron, D. (2006) Lapatinib plus Capecitabine for HER2-Positive Advanced Breast Cancer. *N. Engl. J. Med.* 355, 2733–2743
 60. Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J., and Haber, D. A. (2004) Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer to Gefitinib. *N. Engl. J. Med.* 350, 2129–2139
 61. Ramos, P., and Bentires-Alj, M. (2015) Mechanism-based cancer therapy: resistance to therapy, therapy for resistance. *Oncogene* 34, 3617–3626
 62. Lavi, O. (2015) Redundancy: A Critical Obstacle to Improving Cancer Therapy. *Cancer Res.* 75, 808–812
 63. Gillet, J.-P., and Gottesman, M. M. (2010) Mechanisms of multidrug resistance in cancer. *Methods Mol. Biol. Clifton NJ* 596, 47–76
 64. Chong, C. R., and Jänne, P. A. (2013) The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat. Med.* 19, 1389–1400
 65. Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., Kris, M. G., and Varmus, H. (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* 2, e73
 66. Kobayashi, S., Boggon, T. J., Dayaram, T., Jänne, P. A., Kocher, O., Meyerson, M., Johnson, B. E., Eck, M. J., Tenen, D. G., and Halmos, B. (2005) EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 352, 786–792

67. Yun, C.-H., Mengwasser, K. E., Toms, A. V., Woo, M. S., Greulich, H., Wong, K.-K., Meyerson, M., and Eck, M. J. (2008) The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci.* 105, 2070–2075
68. Kobayashi, S., Ji, H., Yuza, Y., Meyerson, M., Wong, K.-K., Tenen, D. G., and Halmos, B. (2005) An Alternative Inhibitor Overcomes Resistance Caused by a Mutation of the Epidermal Growth Factor Receptor. *Cancer Res.* 65, 7096–7101
69. Ercan, D., Zejnullahu, K., Yonesaka, K., Xiao, Y., Capelletti, M., Rogers, A., Lifshits, E., Brown, A., Lee, C., Christensen, J. G., Kwiatkowski, D. J., Engelman, J. A., and Jänne, P. A. (2010) Amplification of EGFR T790M causes resistance to an irreversible EGFR inhibitor. *Oncogene* 29, 2346–2356
70. Trowe, T., Boukouvava, S., Calkins, K., Cutler, R. E., Fong, R., Funke, R., Gendreau, S. B., Kim, Y. D., Miller, N., Woolfrey, J. R., Vysotskaia, V., Yang, J. P., Gerritsen, M. E., Matthews, D. J., Lamb, P., and Heuer, T. S. (2008) EXEL-7647 inhibits mutant forms of ErbB2 associated with lapatinib resistance and neoplastic transformation. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 14, 2465–2475
71. Montagut, C., Dalmases, A., Bellosillo, B., Crespo, M., Pairet, S., Iglesias, M., Salido, M., Gallen, M., Marsters, S., Tsai, S. P., Minoche, A., Seshagiri, S., Serrano, S., Himmelbauer, H., Bellmunt, J., Rovira, A., Settleman, J., Bosch, F., and Albanell, J. (2012) Identification of a mutation in the extracellular domain of the Epidermal Growth Factor Receptor conferring cetuximab resistance in colorectal cancer. *Nat. Med.* 18, 221–223
72. Scaltriti, M., Chandarlapaty, S., Prudkin, L., Aura, C., Jimenez, J., Angelini, P. D., Sánchez, G., Guzman, M., Parra, J. L., Ellis, C., Gagnon, R., Koehler, M., Gomez, H., Geyer, C., Cameron, D., Arribas, J., Rosen, N., and Baselga, J. (2010) Clinical benefit of lapatinib-based therapy in patients with human epidermal growth factor receptor 2-positive breast tumors coexpressing the truncated p95HER2 receptor. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 16, 2688–2695
73. Berns, K., Horlings, H. M., Hennessy, B. T., Madiredjo, M., Hijmans, E. M., Beelen, K., Linn, S. C., Gonzalez-Angulo, A. M., Stenke-Hale, K., Hauptmann, M., Beijersbergen, R. L., Mills, G. B., van de Vijver, M. J., and Bernards, R. (2007) A Functional Genetic Approach Identifies the PI3K Pathway as a Major Determinant of Trastuzumab Resistance in Breast Cancer. *Cancer Cell* 12, 395–402
74. Nagata, Y., Lan, K.-H., Zhou, X., Tan, M., Esteva, F. J., Sahin, A. A., Klos, K. S., Li, P., Monia, B. P., Nguyen, N. T., Hortobagyi, G. N., Hung, M.-C., and Yu, D. (2004) PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* 6, 117–127
75. Nahta, R., Takahashi, T., Ueno, N. T., Hung, M.-C., and Esteva, F. J. (2004) P27kip1 Down-Regulation Is Associated with Trastuzumab Resistance in Breast Cancer Cells. *Cancer Res.* 64, 3981–3986
76. Zhang, S., Huang, W.-C., Li, P., Guo, H., Poh, S.-B., Brady, S. W., Xiong, Y., Tseng, L.-M., Li, S.-H., Ding, Z., Sahin, A. A., Esteva, F. J., Hortobagyi, G. N., and Yu, D. (2011) Combating trastuzumab resistance by targeting SRC, a common node downstream of multiple resistance pathways. *Nat. Med.* 17, 461–469
77. Brady, S. W., Zhang, J., Tsai, M.-H., and Yu, D. (2015) PI3K-independent mTOR activation promotes lapatinib resistance and IAP expression that can be effectively reversed by mTOR and Hsp90 inhibition. *Cancer Biol. Ther.* 16, 402–411
78. Zhang, Z., Lee, J. C., Lin, L., Olivas, V., Au, V., LaFramboise, T., Abdel-Rahman, M., Wang, X., Levine, A. D., Rho, J. K., Choi, Y. J., Choi, C.-M., Kim, S.-W., Jang, S. J., Park, Y. S., Kim, W. S., Lee, D. H., Lee, J.-S., Miller, V. A., Arcila, M., Ladanyi, M., Moonsamy, P., Sawyers, C., Boggon, T. J., Ma, P. C., Costa, C., Taron, M., Rosell, R., Halmos, B., and Bivona, T. G. (2012) Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. *Nat. Genet.* 44, 852–860
79. Liu, L., Greger, J., Shi, H., Liu, Y., Greshock, J., Annan, R., Halsey, W., Sathe, G. M., Martin, A.-M., and Gilmer, T. M. (2009) Novel mechanism of lapatinib resistance in HER2-positive breast tumor cells: activation of AXL. *Cancer Res.* 69, 6871–6878
80. Shattuck, D. L., Miller, J. K., Carraway, K. L., and Sweeney, C. (2008) Met Receptor Contributes to Trastuzumab Resistance of Her2-Overexpressing Breast Cancer Cells. *Cancer Res.* 68, 1471–1477

81. Engelman, J. A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J. O., Lindeman, N., Gale, C.-M., Zhao, X., Christensen, J., Kosaka, T., Holmes, A. J., Rogers, A. M., Cappuzzo, F., Mok, T., Lee, C., Johnson, B. E., Cantley, L. C., and Jänne, P. A. (2007) MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 316, 1039–1043
82. Koch, H., Busto, M. E. D. C., Kramer, K., Médard, G., and Kuster, B. (2015) Chemical Proteomics Uncovers EPHA2 as a Mechanism of Acquired Resistance to Small Molecule EGFR Kinase Inhibition. *J. Proteome Res.* 14, 2617–2625
83. Zhuang, G., Brantley-Sieders, D. M., Vaught, D., Yu, J., Xie, L., Wells, S., Jackson, D., Muraoka-Cook, R., Arteaga, C., and Chen, J. (2010) Elevation of receptor tyrosine kinase EphA2 mediates resistance to trastuzumab therapy. *Cancer Res.* 70, 299–308
84. Nahta, R., Yuan, L. X. H., Zhang, B., Kobayashi, R., and Esteva, F. J. (2005) Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. *Cancer Res.* 65, 11118–11128
85. Sergina, N. V., Rausch, M., Wang, D., Blair, J., Hann, B., Shokat, K. M., and Moasser, M. M. (2007) Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature* 445, 437–441
86. Wee, S., Jagani, Z., Xiang, K. X., Loo, A., Dorsch, M., Yao, Y.-M., Sellers, W. R., Lengauer, C., and Stegmeier, F. (2009) PI3K Pathway Activation Mediates Resistance to MEK Inhibitors in KRAS Mutant Cancers. *Cancer Res.* 69, 4286–4293
87. Carracedo, A., Ma, L., Teruya-Feldstein, J., Rojo, F., Salmena, L., Alimonti, A., Egia, A., Sasaki, A. T., Thomas, G., Kozma, S. C., Papa, A., Nardella, C., Cantley, L. C., Baselga, J., and Pandolfi, P. P. (2008) Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J. Clin. Invest.* 118, 3065–3074
88. Xia, W., Bacus, S., Hegde, P., Husain, I., Strum, J., Liu, L., Paulazzo, G., Lyass, L., Trusk, P., Hill, J., Harris, J., and Spector, N. L. (2006) A model of acquired autoresistance to a potent ErbB2 tyrosine kinase inhibitor and a therapeutic strategy to prevent its onset in breast cancer. *Proc. Natl. Acad. Sci.* 103, 7795–7800
89. Giuliano, M., Hu, H., Wang, Y.-C., Fu, X., Nardone, A., Herrera, S., Mao, S., Contreras, A., Gutierrez, C., Tao, W., Hilsenbeck, S. G., Angelis, C. D., Wang, N. J., Heiser, L., Gray, J. W., Lopez-Tarruella, S., Pavlick, A., Trivedi, M. V., Chamness, G. C., Chang, J. C., Osborne, C. K., Rimawi, M. F., and Schiff, R. (2015) Upregulation of ER signaling as an adaptive mechanism of cell survival in HER2-positive breast tumors treated with anti-HER2 therapy. *Clin. Cancer Res., clincanres.2728.2014*
90. Huang, C., Park, C. C., Hilsenbeck, S. G., Ward, R., Rimawi, M. F., Wang, Y., Shou, J., Bissell, M. J., Osborne, C. K., and Schiff, R. (2011) β 1 integrin mediates an alternative survival pathway in breast cancer cells resistant to lapatinib. *Breast Cancer Res.* 13, R84
91. Harbinski, F., Craig, V. J., Sanghavi, S., Jeffery, D., Liu, L., Sheppard, K. A., Wagner, S., Stamm, C., Bunes, A., Chatenay-Rivauday, C., Yao, Y., He, F., Lu, C. X., Guagnano, V., Metz, T., Finan, P. M., Hofmann, F., Sellers, W. R., Porter, J. A., Myer, V. E., Graus-Porta, D., Wilson, C. J., Buckler, A., and Tiedt, R. (2012) Rescue Screens with Secreted Proteins Reveal Compensatory Potential of Receptor Tyrosine Kinases in Driving Cancer Growth. *Cancer Discov.* 2, 948–959
92. Wilson, T. R., Fridlyand, J., Yan, Y., Penuel, E., Burton, L., Chan, E., Peng, J., Lin, E., Wang, Y., Sosman, J., Ribas, A., Li, J., Moffat, J., Sutherland, D. P., Koeppen, H., Merchant, M., Neve, R., and Settleman, J. (2012) Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature* 487, 505–509
93. Steen, H., and Mann, M. (2004) The abc's (and xyz's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* 5, 699–711
94. Kim, M.-S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabudde, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L. D. N., Patil, A. H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T., Raju, R., Kumar, M., Sreenivasamurthy, S. K., Marimuthu, A., Sathe, G. J., Chavan, S., Datta, K. K., Subbannayya, Y., Sahu, A., Yelamanchi, S. D., Jayaram, S., Rajagopalan, P., Sharma, J., Murthy, K. R., Syed, N., Goel, R., Khan, A. A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T.-C., Zhong, J., Wu, X., Shaw, P. G., Freed, D., Zahari, M. S., Mukherjee, K. K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C. J., Shankar, S. K., Satishchandra, P., Schroeder, J. T., Sirdeshmukh, R.,

- Maitra, A., Leach, S. D., Drake, C. G., Halushka, M. K., Prasad, T. S. K., Hruban, R. H., Kerr, C. L., Bader, G. D., Iacobuzio-Donahue, C. A., Gowda, H., and Pandey, A. (2014) A draft map of the human proteome. *Nature* 509, 575–581
95. Mallick, P., and Kuster, B. (2010) Proteomics: a pragmatic perspective. *Nat. Biotechnol.* 28, 695–709
 96. Hebert, A. S., Richards, A. L., Bailey, D. J., Ulbrich, A., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2014) The One Hour Yeast Proteome. *Mol. Cell. Proteomics* 13, 339–347
 97. Mann, M., Kulak, N. A., Nagaraj, N., and Cox, J. (2013) The Coming Age of Complete, Accurate, and Ubiquitous Proteomes. *Mol. Cell* 49, 583–590
 98. Wolters, D. A., Washburn, M. P., and Yates, J. R. (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* 73, 5683–5690
 99. Zhou, F., Sikorski, T. W., Ficarro, S. B., Webber, J. T., and Marto, J. A. (2011) Online nanoflow reversed phase-strong anion exchange-reversed phase liquid chromatography-tandem mass spectrometry platform for efficient and in-depth proteome sequence analysis of complex organisms. *Anal. Chem.* 83, 6996–7005
 100. Alpert, A. J. (1990) Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J. Chromatogr.* 499, 177–196
 101. Gilar, M., Olivova, P., Daly, A. E., and Gebler, J. C. (2005) Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J. Sep. Sci.* 28, 1694–1703
 102. Hao, P., Guo, T., Li, X., Adav, S. S., Yang, J., Wei, M., and Sze, S. K. (2010) Novel Application of Electrostatic Repulsion-Hydrophilic Interaction Chromatography (ERLIC) in Shotgun Proteomics: Comprehensive Profiling of Rat Kidney Proteome. *J. Proteome Res.* 9, 3520–3526
 103. Boersema, P. J., Divecha, N., Heck, A. J. R., and Mohammed, S. (2007) Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. *J. Proteome Res.* 6, 937–946
 104. Hennrich, M. L., Groenewold, V., Kops, G. J. P. L., Heck, A. J. R., and Mohammed, S. (2011) Improving Depth in Phosphoproteomics by Using a Strong Cation Exchange-Weak Anion Exchange-Reversed Phase Multidimensional Separation Approach. *Anal. Chem.* 83, 7137–7143
 105. Ritorto, M. S., Cook, K., Tyagi, K., Pedrioli, P. G. A., and Trost, M. (2013) Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *J. Proteome Res.* 12, 2449–2457
 106. Fabrizio, P., Esser, S., Kastner, B., and Luhrmann, R. (1994) Isolation of *S. cerevisiae* snRNPs: comparison of U1 and U4/U6.U5 to their human counterparts. *Science* 264, 261–265
 107. Neubauer, G., Gottschalk, A., Fabrizio, P., Séraphin, B., Luhrmann, R., and Mann, M. (1997) Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc. Natl. Acad. Sci.* 94, 385–390
 108. Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S.-L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jepsen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sørensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W. V., Figeys, D., and Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415, 180–183
 109. Hein, M. Y., Hubner, N. C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., Hyman, A. A., and Mann, M. (2015) A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. *Cell* 163, 712–723
 110. Huttlin, E. L., Ting, L., Bruckner, R. J., Gebreab, F., Gygi, M. P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., Dong, R., Guarani, V., Vaites, L. P., Ordureau, A., Rad, R., Erickson, B. K., Wühr, M., Chick, J., Zhai, B., Kolippakkam, D., Mintseris, J., Obar, R. A., Harris, T., Artavanis-Tsakonas, S., Sowa, M. E., De Camilli, P., Paulo, J. A., Harper, J. W., and Gygi, S. P. (2015) The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* 162, 425–440
 111. Schirle, M., Bantscheff, M., and Kuster, B. (2012) Mass spectrometry-based proteomics in preclinical drug discovery. *Chem. Biol.* 19, 72–84
 112. Cuatrecasas, P., Wilchek, M., and Anfinsen, C. B. (1968) Selective enzyme purification by affinity chromatography. *Proc. Natl. Acad. Sci. U. S. A.* 61, 636–643

113. Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A.-M., Schlegl, J., Abraham, Y., Becher, I., Bergamini, G., Boesche, M., Delling, M., Dümpelfeld, B., Eberhard, D., Huthmacher, C., Mathieson, T., Poeckel, D., Reader, V., Strunk, K., Sweetman, G., Kruse, U., Neubauer, G., Ramsden, N. G., and Drewes, G. (2011) Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat. Biotechnol.* 29, 255–265
114. Graves, P. R., Kwiek, J. J., Fadden, P., Ray, R., Hardeman, K., Coley, A. M., Foley, M., and Haystead, T. A. J. (2002) Discovery of novel targets of quinoline drugs in the human purine binding proteome. *Mol. Pharmacol.* 62, 1364–1372
115. Brehmer, D., Godl, K., Zech, B., Wissing, J., and Daub, H. (2004) Proteome-wide identification of cellular targets affected by bisindolylmaleimide-type protein kinase C inhibitors. *Mol. Cell. Proteomics MCP* 3, 490–500
116. Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* 25, 1035–1044
117. Médard, G., Pachi, F., Ruprecht, B., Klaeger, S., Heinzlmeir, S., Helm, D., Qiao, H., Ku, X., Wilhelm, M., Kuehne, T., Wu, Z., Dittmann, A., Hopf, C., Kramer, K., and Kuster, B. (2015) Optimized Chemical Proteomics Assay for Kinase Inhibitor Profiling. *J. Proteome Res.* 14, 1574–1586
118. Pachi, F., Plattner, P., Ruprecht, B., Médard, G., Sewald, N., and Kuster, B. (2013) Characterization of a Chemical Affinity Probe Targeting Akt Kinases. *J. Proteome Res.* 12, 3792–3800
119. Ku, X., Heinzlmeir, S., Helm, D., Médard, G., and Kuster, B. (2014) New Affinity Probe Targeting VEGF Receptors for Kinase Inhibitor Selectivity Profiling by Chemical Proteomics. *J. Proteome Res.* 13, 2445–2452
120. Daub, H. (2015) Quantitative Proteomics of Kinase Inhibitor Targets and Mechanisms. *ACS Chem. Biol.* 10, 201–212
121. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jørgensen, T. J. D. (2005) Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Mol. Cell. Proteomics* 4, 873–886
122. Rogers, L. D., and Foster, L. J. (2009) Phosphoproteomics—finally fulfilling the promise? *Mol. Biosyst.* 5, 1122–1129
123. Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A. J. R., Zou, H., and Mohammed, S. (2013) Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat. Protoc.* 8, 461–480
124. Xu, G., Paige, J. S., and Jaffrey, S. R. (2010) Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* 28, 868–873
125. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840
126. Salomon, A. R., Ficarro, S. B., Brill, L. M., Brinker, A., Phung, Q. T., Ericson, C., Sauer, K., Brock, A., Horn, D. M., Schultz, P. G., and Peters, E. C. (2003) Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 100, 443–448
127. Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., and Elledge, S. J. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316, 1160–1166
128. Moritz, A., Li, Y., Guo, A., Villén, J., Wang, Y., MacNeill, J., Kornhauser, J., Sprott, K., Zhou, J., Possemato, A., Ren, J. M., Hornbeck, P., Cantley, L. C., Gygi, S. P., Rush, J., and Comb, M. J. (2010) Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Sci. Signal.* 3, ra64
129. Giansanti, P., Stokes, M. P., Silva, J. C., Scholten, A., and Heck, A. J. R. (2013) Interrogating cAMP-dependent Kinase Signaling in Jurkat T Cells via a Protein Kinase A Targeted Immuno-precipitation Phosphoproteomics Approach. *Mol. Cell. Proteomics MCP* 12, 3350–3359

130. Andersson, L., and Porath, J. (1986) Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal. Biochem.* 154, 250–254
131. Dunn, J. D., Watson, J. T., and Bruening, M. L. (2006) Detection of Phosphopeptides Using Fe(III)–Nitrilotriacetate Complexes Immobilized on a MALDI Plate. *Anal. Chem.* 78, 1574–1580
132. Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. R. (2004) Selective Isolation at the Femtomole Level of Phosphopeptides from Proteolytic Digests Using 2D-NanoLC-ESI-MS/MS and Titanium Oxide Precolumns. *Anal. Chem.* 76, 3935–3943
133. Kweon, H. K., and Håkansson, K. (2006) Selective Zirconium Dioxide-Based Enrichment of Phosphorylated Peptides for Mass Spectrometric Analysis. *Anal. Chem.* 78, 1743–1749
134. Kettenbach, A. N., and Gerber, S. A. (2011) Rapid and Reproducible Single-Stage Phosphopeptide Enrichment of Complex Peptide Mixtures: Application to General and Phosphotyrosine-Specific Phosphoproteomics Experiments. *Anal. Chem.* 83, 7635–7644
135. Zhou, H., Ye, M., Dong, J., Han, G., Jiang, X., Wu, R., and Zou, H. (2008) Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. *J. Proteome Res.* 7, 3957–3967
136. Zhou, H., Low, T. Y., Hennrich, M. L., Toorn, H. van der, Schwend, T., Zou, H., Mohammed, S., and Heck, A. J. R. (2011) Enhancing the Identification of Phosphopeptides from Putative Basophilic Kinase Substrates Using Ti (IV) Based IMAC Enrichment. *Mol. Cell. Proteomics* 10, M110.006452
137. Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., and Aebersold, R. (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* 4, 231–237
138. Tsai, C.-F., Hsu, C.-C., Hung, J.-N., Wang, Y.-T., Choong, W.-K., Zeng, M.-Y., Lin, P.-Y., Hong, R.-W., Sung, T.-Y., and Chen, Y.-J. (2013) Sequential Phosphoproteomic Enrichment through Complementary Metal-Directed Immobilized Metal Ion Affinity Chromatography. *Anal. Chem.*,
139. Thingholm, T. E., Jensen, O. N., Robinson, P. J., and Larsen, M. R. (2008) SIMAC (Sequential Elution from IMAC), a Phosphoproteomics Strategy for the Rapid Separation of Monophosphorylated from Multiply Phosphorylated Peptides. *Mol. Cell. Proteomics* 7, 661–671
140. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villén, J., Li, J., Cohn, M. A., Cantley, L. C., and Gygi, S. P. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12130–12135
141. Hennrich, M. L., van den Toorn, H. W. P., Groenewold, V., Heck, A. J. R., and Mohammed, S. (2012) Ultra Acidic Strong Cation Exchange Enabling the Efficient Enrichment of Basic Phosphopeptides. *Anal. Chem.* 84, 1804–1808
142. McNulty, D. E., and Annan, R. S. (2008) Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol. Cell. Proteomics MCP* 7, 971–980
143. Loroch, S., Zahedi, R. P., and Sickmann, A. (2015) Highly Sensitive Phosphoproteomics by Tailoring Solid-Phase Extraction to Electrostatic Repulsion-Hydrophilic Interaction Chromatography. *Anal. Chem.* 87, 1596–1604
144. Batth, T. S., Francavilla, C., and Olsen, J. V. (2014) Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics. *J. Proteome Res.* 13, 6176–6186
145. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198–207
146. Hahne, H. (2012) Studies towards the proteome-wide detection, identification and quantification of protein glycosylation.
147. Karas, M., Bachmann, D., Bahr, U., and Hillenkamp, F. (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int. J. Mass Spectrom. Ion Process.* 78, 53–68
148. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 60, 2299–2301
149. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 2, 151–153
150. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64–71
151. Knochenmuss, R. (2006) Ion formation mechanisms in UV-MALDI. *The Analyst* 131, 966–986

152. Jaskolla, T. W., and Karas, M. (2011) Compelling evidence for Lucky Survivor and gas phase protonation: the unified MALDI analyte protonation mechanism. *J. Am. Soc. Mass Spectrom.* 22, 976–988
153. Wilm, M. S., and Mann, M. (1994) Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Process.* 136, 167–180
154. Iribarne, J. V., and Thomson, B. A. (1976) On the evaporation of small ions from charged droplets. *J. Chem. Phys.* 64, 2287–2294
155. Wilm, M., and Mann, M. (1996) Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* 68, 1–8
156. Zubarev, R. A., and Makarov, A. (2013) Orbitrap Mass Spectrometry. *Anal. Chem.* 85, 5288–5296
157. Scigelova, M., and Makarov, A. (2006) Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics* 6 Suppl 2, 16–21
158. Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., and Graham Cooks, R. (2005) The Orbitrap: a new mass spectrometer. *J. Mass Spectrom. JMS* 40, 430–443
159. Roepstorff, P., and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* 11, 601
160. Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* 4, 709–712
161. Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9528–9533
162. Michalski, A., Neuhauser, N., Cox, J., and Mann, M. (2012) A Systematic Investigation into the Nature of Tryptic HCD Spectra. *J. Proteome Res.* 11, 5479–5491
163. Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E. R., Senko, M., Makarov, A., Mann, M., and Horning, S. (2009) A Dual Pressure Linear Ion Trap Orbitrap Instrument with Very High Sequencing Speed. *Mol. Cell. Proteomics MCP* 8, 2759–2769
164. McAlister, G. C., Phanstiel, D., Good, D. M., Berggren, W. T., and Coon, J. J. (2007) Implementation of Electron-Transfer Dissociation on a Hybrid Linear Ion Trap–Orbitrap Mass Spectrometer. *Anal. Chem.* 79, 3525–3534
165. Frese, C. K., Zhou, H., Taus, T., Altelaar, A. F. M., Mechtler, K., Heck, A. J. R., and Mohammed, S. (2013) Unambiguous Phosphosite Localization using Electron-Transfer/Higher-Energy Collision Dissociation (ET_hCD). *J. Proteome Res.* 12, 1520–1525
166. Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003) PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom. RCM* 17, 2337–2342
167. Yates, J. R., Morgan, S. F., Gatlin, C. L., Griffin, P. R., and Eng, J. K. (1998) Method To Compare Collision-Induced Dissociation Spectra of Peptides: Potential for Library Searching and Subtractive Analysis. *Anal. Chem.* 70, 3557–3565
168. Nesvizhskii, A. I., Vitek, O., and Aebersold, R. (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat. Methods* 4, 787–797
169. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* 10, 1794–1805
170. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567
171. Eng, J. K., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 5, 976–989
172. Nesvizhskii, A. I., and Aebersold, R. (2005) Interpretation of shotgun proteomic data: the protein inference problem. *Mol. Cell. Proteomics MCP* 4, 1419–1440
173. Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., and Kuster, B. (2011) Confident phosphorylation site localization using the Mascot Delta Score. *Mol. Cell. Proteomics MCP* 10, M110.003830

174. Taus, T., Köcher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., and Mechtler, K. (2011) Universal and Confident Phosphorylation Site Localization Using phosphoRS. *J. Proteome Res.* 10, 5354–5362
175. Marx, H., Lemeer, S., Schliep, J. E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A. J. R., and Kuster, B. (2013) A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nat. Biotechnol.* 31, 557–564
176. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* 4, 207–214
177. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* 389, 1017–1031
178. Bantscheff, M., Lemeer, S., Savitski, M. M., and Kuster, B. (2012) Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* 404, 939–965
179. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics MCP* 1, 376–386
180. Hebert, A. S., Merrill, A. E., Bailey, D. J., Still, A. J., Westphall, M. S., Strieter, E. R., Pagliarini, D. J., and Coon, J. J. (2013) Neutron-encoded mass signatures for multiplexed proteome quantification. *Nat. Methods* 10, 332–334
181. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., and Mann, M. (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* 7, 383–385
182. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. R. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* 4, 484–494
183. Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K. A., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 75, 1895–1904
184. Wiese, S., Reidegeld, K. A., Meyer, H. E., and Warscheid, B. (2007) Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics* 7, 340–350
185. McAlister, G. C., Huttlin, E. L., Haas, W., Ting, L., Jedrychowski, M. P., Rogers, J. C., Kuhn, K., Pike, I., Grothe, R. A., Blethrow, J. D., and Gygi, S. P. (2012) Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* 84, 7469–7478
186. Ow, S. Y., Salim, M., Noirel, J., Evans, C., Rehman, I., and Wright, P. C. (2009) iTRAQ underestimation in simple and complex mixtures: “the good, the bad and the ugly.” *J. Proteome Res.* 8, 5347–5355
187. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372
188. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., Mering, C. von, and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808–D815
189. Gene Ontology Consortium (2008) The Gene Ontology project in 2008. *Nucleic Acids Res.* 36, D440–444
190. Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30
191. Mi, H., Muruganujan, A., and Thomas, P. D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* 41, D377–386
192. Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4, P3
193. Schwartz, D., and Gygi, S. P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* 23, 1391–1398
194. Obenaus, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* 31, 3635–3641

195. Yaffe, M. B., Leparc, G. G., Lai, J., Obata, T., Volinia, S., and Cantley, L. C. (2001) A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nat. Biotechnol.* 19, 348–353
196. Miller, M. L., Jensen, L. J., Diella, F., Jørgensen, C., Tinti, M., Li, L., Hsiung, M., Parker, S. A., Bordeaux, J., Sicheritz-Ponten, T., Olhovsky, M., Pasculescu, A., Alexander, J., Knapp, S., Blom, N., Bork, P., Li, S., Cesareni, G., Pawson, T., Turk, B. E., Yaffe, M. B., Brunak, S., and Linding, R. (2008) Linear motif atlas for phosphorylation-dependent signaling. *Sci. Signal.* 1, ra2
197. Linding, R., Jensen, L. J., Pasculescu, A., Olhovsky, M., Colwill, K., Bork, P., Yaffe, M. B., and Pawson, T. (2008) NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* 36, D695–699
198. Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., Palma, A., Cesareni, G., Jensen, L. J., and Linding, R. (2014) KinomeExplorer: an integrated platform for kinome biology studies. *Nat. Methods* 11, 603–604
199. Gnad, F., Ren, S., Cox, J., Olsen, J. V., Macek, B., Oroshi, M., and Mann, M. (2007) PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol.* 8, R250
200. Dinkel, H., Chica, C., Via, A., Gould, C. M., Jensen, L. J., Gibson, T. J., and Diella, F. (2011) Phospho.ELM: a database of phosphorylation sites--update 2011. *Nucleic Acids Res.* 39, D261–267
201. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270

Chapter 2

Comprehensive and reproducible phosphopeptide enrichment using iron immobilized metal ion affinity chromatography (Fe-IMAC) columns

Abstract

Advances in phosphopeptide enrichment methods enable the identification of thousands of phosphopeptides from complex samples. Current offline enrichment approaches using TiO_2 , Ti-IMAC and Fe-IMAC material in batch or microtip format are widely used, but suffer from irreproducibility and compromised selectivity. To address these shortcomings, the merits of performing phosphopeptide enrichments in an high performance liquid chromatography (HPLC) column format were revisited. It was found that Fe-IMAC columns enable the selective, comprehensive and reproducible enrichment of phosphopeptides out of complex lysates. Column enrichment did not suffer from bead-to-sample ratio issues and scaled linearly from 100 μg to 5 mg digest. Direct measurements on an Orbitrap Velos mass spectrometer identified >7,500 unique phosphopeptides with 90 % selectivity and good quantitative reproducibility (median CV of 15 %). The number of unique phosphopeptides could be increased to over 14,000 when subjecting the IMAC eluate to a subsequent hydrophilic strong anion exchange (hSAX) separation. Fe-IMAC columns outperformed Ti-IMAC and TiO_2 in batch or tip mode in terms of phosphopeptide identification and intensity. Permutation enrichments of flow troughs show that all materials largely bind the same phosphopeptide species, independent of physicochemical characteristics. However, binding capacity and elution efficiency did profoundly differ among the enrichment materials and formats. As a result, the often quoted orthogonality of the materials has to be called into question. My results strongly suggest that insufficient capacity, inefficient elution and the stochastic nature of data dependent acquisition in mass spectrometry are the cause of the experimentally observed complementarity. The Fe-IMAC enrichment workflow using an HPLC format developed here enables rapid and comprehensive phosphoproteome analysis which can be applied to a wide range of biological systems.

Introduction

Protein phosphorylation is a reversible post translational modification with pivotal roles in cellular signaling. It is not only implicated in many essential biological processes, but aberrant protein phosphorylation is causally linked to numerous diseases (1). At least one third to half of all human proteins are thought to be phosphorylated at some point (2, 3) and hundreds of thousands of different phosphorylation sites are believed to exist. Due to this molecular complexity and the often substoichiometric extent of phosphorylation specific enrichment prior to analysis by liquid chromatography tandem mass spectrometry (LC-MS/MS) is generally required (4). Most current such strategies exploit the affinity of phosphate groups to metals immobilized on carrier resins. These include Fe^{3+} -(5), Ga^{3+} -(6) or Zr^{4+} -(7)- immobilized metal ion affinity chromatography (IMAC), metal oxide affinity chromatography (TiO_2 (8), ZrO_2 (9) and others) and the recently introduced Ti-IMAC material which is a hybrid of the two (10, 11). It is often stated if not generally accepted that the above enrichment methods are capable of purifying complementary parts of the phosphoproteome each with unique physiochemical characteristics (10, 12, 13). For example, Ti-IMAC is thought to be better at purifying basophilic phosphopeptides compared to TiO_2 (10) whereas Fe-IMAC is attributed with the more efficient enrichment of multiply phosphorylated peptides (14). The current consensus in the field is that no method is superior over others and neither of them alone is sufficient for the comprehensive purification of the phosphoproteome, a view that is challenged in this study. Most phosphopeptide enrichments are conducted either in a batch format or in self-constructed microcolumns, packed into pipette tips. Both formats have been found to suffer from considerable variability introduced by various manual steps in the process. Moreover, differences in loading conditions including additives (15, 16), acid concentration (16), incubation time (16, 17) and wash volume (17) can lead to different results. One prominent parameter that was found to have considerable impact on the enrichment quality is the ratio of bead amount to the protein digest quantity and finding the optimal ratio is often a tradeoff between comprehensiveness and selectivity (12, 17-20). Consequently, careful *a priori* evaluation and optimization of enrichment conditions is generally required on a case by case basis for every experimental system. Even then, the use of such formats often comes at the expense of intra-experimental and, even more so, of inter-experimental accuracy. As a means to overcome these issues, direct coupling of the chromatographic enrichment step with the LC-MS/MS system have been applied. Although these online systems increase reproducibility, sensitivity

and robustness, they suffer from limited capacity which is why they have been primarily used for the analysis of limited sample amounts or samples of rather low complexity (8, 21-26).

In order to address the latter, phosphopeptide enrichment workflows often employ an upstream peptide separation step followed by phosphopeptide enrichment from each fraction. Examples for first dimension separations include HILIC (27), ERLIC (28), SAX (29) and SCX (10, 30). Although powerful and frequently used, enrichment after fractionation can be irreproducible and time consuming. As a consequence, reversing the order (i.e. enriching phosphopeptides first followed by fractionation of the phosphopeptide pool) has recently become popular (16, 17, 31). This, however, necessitates that the phosphopeptide enrichment step is of exquisite selectivity, has sufficient capacity and allows complete elution (16, 17). Phosphopeptides purified in this way can be directly analyzed by LC-MS/MS (32, 33) typically using shallow reversed phase LC gradients (34). However, it turns out that current mass spectrometers still lack the scan speed and dynamic range required to reach complete (phospho)peptide sampling in direct LC-MS/MS measurements (35, 36).

In this study, I describe a robust and flexible workflow based on offline chromatographic enrichment of phosphopeptides using a commercially available Fe-IMAC column. This approach offers selective, comprehensive and reproducible enrichment and scales over wide range of sample quantities. I show that this method outperforms all Ti-IMAC and TiO₂ methods tested and my data argues that the apparent orthogonality of all three methods is caused by a combination of format inadequacies, inefficient elution and insufficient data acquisition speed rather than exploiting different physicochemical characteristics of phosphopeptides.

Material and methods

Cell culture and protein digest

Human epidermoid A431 cells were grown in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10 % (v/v) Fetal Bovine Serum and 1 % Antibiotic/Antimycotic solution. For all phosphopeptide enrichment optimization experiments, cells were treated with 1 mM pervanadate for 5min prior to lysis. After harvesting, cells were washed two times with ice cold PBS and lysed in 8 M Urea, 40 mM Tris/HCl (pH 7.6), 1 x EDTA free protease inhibitor mixture (complete mini, Roche) and 1 x Phosphatase inhibitor cocktail (Sigma). The lysate was centrifuged at 20,000 rpm for 1 hour at 4 °C. Protein concentration was determined using the Bradford method (Coomassie (Bradford) Protein Assay Kit, Thermo Scientific). The supernatant was reduced with 10 mM DTT at 56 °C for 1 h and alkylated with 25 mM iodoacetamide for 45 min at room temperature in the dark. The protein mixture was diluted with 40 mM Tris/HCl to a final Urea concentration of 1.6 M. Digestion was performed by adding sequencing grade trypsin (Promega, 1:100 enzyme:substrate ratio) and incubation at 37 °C for 4 h. Subsequently, another 1:100 trypsin was added for overnight digestion at 37 °C. Samples were acidified with TFA to a pH of 2 in order to stop trypsin activity. SepPack columns (C18 cartridges Sep-Pak Vac 1cc (50 mg), Waters Corp., solvent A: 0.07 % TFA, solvent B: 0.07 % TFA, 50 % ACN) were used for peptide desalting according to manufacturer's instructions and eluates were dried down and stored at -80 °C. Smaller sample amounts of up to 100 µg were desalted as described (37).

Fe-IMAC batch enrichment

Fe-IMAC batch enrichments were essentially performed as described (30). Briefly, 100 µl Fe-IMAC beads (PhosSelect iron affinity gel, Sigma) beads were washed four times with 1 ml of binding solvent (25 mM FA, 40 % ACN) and a 50 % slurry in binding solvent was prepared thereafter. Dried down peptides were resuspended in binding solvent at a concentration of 1 µg/µl. After combination of beads (at the amount specified in the respective figure) and dissolved peptides, samples were incubated for 1 h at room temperature under vigorous shaking. Subsequently, Fe-IMAC beads were transferred on top of previously equilibrated C18 StageTips and washed by sequentially passing through 50 µl of Fe-IMAC binding solvent (twice) and 40 µl 1 % FA. Elution was achieved by application of 70 µl of 500 mM K₂HPO₄, pH 7 (twice). Peptides which were retained on the C18 material were washed with 40 µl 1% FA and

eluted directly into an MS plate using 40 μ l 60 % ACN, 0.1 % FA. Eluates were dried down and stored at -80 °C.

Fe-IMAC column phosphopeptide enrichment

The desalted digest was reconstituted in 0.5 ml Fe-IMAC solvent A (30 % ACN, 0.07 % (v/v) TFA) and loaded onto an analytical Fe-IMAC column (4 mm I.D. x 50 mm or 9 mm I.D. x 50 mm ProPac IMAC-10, Thermo Fisher Scientific) connected to an HPLC system (ÄKTA explorer HPLC system, Amersham Pharmacia) with a 1 ml sample loop. The column was charged with iron according to manufacturer's instructions. Briefly, the column was rinsed with 3 column volumes of 20 mM formic acid and charged using 3 column volumes of 25 mM FeCl₃, 100 mM acetic acid. To wash out unbound iron-ions the column was flushed with 20 column volumes of 20 mM formic acid and subsequently removed from the HPLC system. The HPLC lines were first flushed with 20 ml ddH₂O water, followed by 20 ml 50 mM EDTA and 10 ml ddH₂O to remove remaining iron-ions. After column re-connection and baseline equilibration, the gradient was started. Sample was loaded (0.1 ml/min, 10 min) and unbound peptides were washed out with Fe-IMAC solvent A (0.3 ml/min, 16 min). Subsequent phosphopeptide elution was achieved by a linear gradient from 0 % to 45 % Fe-IMAC solvent B (0.5 % (v/v) NH₄OH) (0.2 ml/min, 60 min). After increase to 100 % B and a holding step for 5 min the column was re-equilibrated by switching to Fe-IMAC solvent A (30 min, 0.5 ml/min). Flow through and a phosphopeptide fraction were collected according to the UV signal (280 nm), dried down and stored at -80°C. The column was recharged after a maximum of three enrichments.

Ti-IMAC synthesis and phosphopeptide enrichment

Ti-IMAC beads were synthesized as previously described by Zhou et al. (38) with some modifications and additional quality control steps. A detailed description can be found in the supplementary material and methods. Whereas enrichment of up to 250 μ g of sample was performed as described (38), the phosphopeptide enrichment protocol was modified for sample amounts exceeding 250 μ g. Sep-Pak cartridges (C18 cartridges Sep-Pak Vac 1cc (50 mg), Waters Corp.) were attached to a vacuum manifold and flushed with 1 ml Ti-IMAC loading solvent (80 % ACN, 6 % TFA). Subsequently, 50 mg of Ti-IMAC beads were loaded on top of the C18 material and equilibrated with 5 x 1 ml loading solvent. Dried down digests were dissolved in 1ml of binding solvent and slowly passed through the cartridges. After reapplication of the flow through, columns were washed with 10 ml of washing solution 1 (50 % (v/v) ACN,

0.5 % (v/v) TFA, 200 mM NaCl) and 10 ml of washing solution 2 (50 % (v/v) ACN, 0.1 % (v/v) TFA). Bound peptides were eluted sequentially with 0.8 ml 10 % (v/v) NH₄OH and 0.2 ml 80 % (v/v) ACN, 2 % (v/v) FA. Flow-through and elution fractions were dried down and stored at -80 °C.

TiO₂ batch enrichment

TiO₂ batch enrichment was performed as described by Kettenbach et al. (16) with some modifications. TiO₂ beads (5 µM, GL Sciences Inc.) were washed twice with 1 ml of washing solvent (50 % ACN, 0.1 % TFA) and four times with 1 ml of binding solvent (2 M lactic acid, 50 % ACN, 0.1 % TFA). In between beads were spun down and the supernatant was discarded. Peptides were dissolved in 0.5 ml of binding solvent and after addition of 0.25 ml equilibrated bead slurry, the mixture was incubated for 1 h at room temperature under vigorous shaking. Subsequently, beads were washed four times with 0.2 ml binding solvent and five times with 1 ml washing solution. Bound peptides were eluted by two 10 min incubation steps with 200 µl elution solvent (50 mM KH₂PO₄, 0.5 % (v/v) NH₄OH, pH 11.3). The supernatant was quenched by addition of 30 µl 100 % FA, dried down and stored at -80 °C.

Phosphopeptide enrichment with TiO₂ and Ti-IMAC columns

For chromatographic separation of phosphorylated peptides, TiO₂ beads (5 µm, GL Sciences Inc) were packed into a column (4 mm I.D. x 10 mm, Dr. Maisch, Ammersham) and connected to a HPLC system (ÄKTA explorer HPLC, Amersham Pharmacia). Dried samples were reconstituted in 0.5 ml TiO₂ solvent A (80 % ACN, 6 % TFA) and injected. The system was operated at a flow rate of 0.2 ml/min with linear gradient from 0 % to 85 % solvent B (1.5 % (v/v) NH₄OH, pH 11.8) in 30 min, followed by an increase to 95 % in 15 min. For elution of bound peptides the concentration of solvent B was kept at 95 % for 50 min. In total, 13 fractions were collected. In house synthesized Ti-IMAC beads were also packed into a column (4 mm I.D. x 50 mm, Dr. Maisch, Ammersham). The gradient corresponded to the one used for Fe-IMAC column enrichments, except for some modifications: due to strong metal ion peptide interaction 10 % (v/v) NH₄OH was used for elution.

Hydrophilic strong anion exchange separation (hSAX)

A Dionex Ultimate 3000 HPLC system (Dionex Corp.) equipped with an IonPac AG24 guard column (2 mm I.D. x 50 mm, Thermo Fisher Scientific) and a IonPac AS24 SAX-column

(2 mm I.D. x 250 mm, Thermo Fisher Scientific) was used for hydrophilic strong anion exchange separation of both enriched phosphopeptides and digested full proteome samples. The system was operated at 30 °C with a flow rate of 0.25 ml/min and an initial 3 min equilibration step with 100 % hSAX solvent A (5 mM Tris/HCl, pH=8.5) followed by elution with a linear 17 min gradient up to 40 % hSAX solvent B (5 mM Tris/HCl, pH=8.5, 1 M NaCl). Solvent B was increased to 100 % in 10 min and held constant for another 10 min. A subsequent switch to 100 % solvent A in 3 min was followed by column re-equilibration with 100 % solvent A for 10 min. 24 fractions were collected in 1 min intervals (starting 2 min into the gradient), dried down and measured after on-trap desalting.

LC-MS/MS analysis

Nanoflow LC-MS/MS was performed by coupling an Eksigent nanoLC-Ultra 1D+ (Eksigent, Dublin, CA) to an Oribtrap Velos (Thermo Scientific, Bremen, Germany). Peptides were delivered to a trap column (100 µm I.D. x 2 cm, packed with 5 µm C18 resin, Reprosil PUR AQ, Dr. Maisch, Ammerbuch, Germany) at a flow rate of 2 µL/minute for 12 min and 5 µl/min for 13 min in 100 % solvent A (0.1 % FA in HPLC grade water). After 25 min of loading and washing, peptides were transferred to an analytical column (75 µm I.D. x 40 cm C18 column Reprosil PUR AQ, 3 µm, Dr. Maisch, Ammerbuch, Germany) and separated using a linear gradient from 2 % to 27 % (4 % - 32 % for not previously enriched sample) solvent B (0.1 % FA, 5 % DMSO in ACN) at a flow rate of 300 nl/min. For one direct measurement, a 210 min gradient (220 min turnaround) was applied whereas hSAX fractions were measured for 110 min per fraction (135 min turnaround including an on-trap desalting step). TiO₂ column fractions were measured for 45 min each (60 min turnaround including an on-trap desalting step). To prevent potential Fe³⁺ and Ti⁴⁺ ion adsorption to the reversed phase stationary phase, 5 µl of 1 µM deferoxamine dissolved in ddH₂O was injected in between phosphopeptide direct measurements (39).

Peptides were ionized using 2.2 kV spray voltage and a capillary temperature of 275 °C. The mass spectrometer was operated in data dependent acquisition mode, automatically switching between MS and MS². Full scan MS spectra (m/z 360 – 1300) were acquired in the Orbitrap at 30,000 (m/z 400) resolution and an AGC target value of 1e6. For internal calibration the signal at m/z 401.922718 was used as a lock mass (40). High resolution HCD MS² spectra were generated for up to 10 precursors with a normalized collision energy of 30 % (for non enriched samples) and 35 % (for phosphopeptide samples). The precursor ion count for triggering an MS² event was set to 500 with a dynamic exclusion of 20 s. Fragment ions were read out in the

Orbitrap mass analyzer at a resolution of 7,500 (isolation window 2 Th). For phosphopeptide samples, the MS2 AGC target value was set to 4e4 with a maximum ion injection time of 250 ms whereas not previously enriched samples were measured with 3e4 and 200 ms.

Peptide and protein identification and quantification

For peptide and protein identification, peak lists were extracted from raw files using Mascot Distiller v2.2.1 (Matrix Science, UK) and subsequently searched against the Human IPI database (v3.68) using Mascot (v2.3.0) with the following parameters: carbamidomethyl cysteine as a fixed modification, phosphorylation of serine threonine and tyrosine, oxidation of methionin and N-terminal protein acetylation as variable modifications. Precursor tolerance was set to 10 ppm and fragment ion tolerance to 0.05 Da. Trypsin/P was specified as the proteolytic enzyme with up to two missed cleavage sites allowed. Identified phosphopeptides were filtered using two criteria: (i) PSMs were filtered for 1 % FDR using Rockerbox (41) in combination with percolator (v1.0) (42), (ii) remaining peptides corresponding to accepted PSMs were filtered according to a 5 % global peptide FDR cutoff based on the formula published by Marx et al. (43) (HCD, global FDR). This translates to a Mascot ion score > 13.96 for phosphopeptides and > 27.65 for non phosphorylated peptides. The peptide table created after Rockerbox processing was used for further analysis. Three different categories for phosphopeptide identifications are established: (i) unique phosphopeptides are based on non redundant phosphosequences including phosphoisomers, (ii) unique phosphosequences are based on the sequence combined with the number of phosphomodifications (phosphoisomers are not considered), (iii) confidently localized phosphopeptides are unique phosphopeptides which were additionally filtered with a 5 % false localization cutoff based on the Mascot delta score which was calculated according to the formula introduced by Marx et al. (43) (HCD). In order to calculate phosphopeptide selectivity the global 5 % FDR cutoff was separately applied to identified phospho and non-phosphopeptides. For label-free quantification Progenesis (v4.1, Nonlinear Dynamics, Newcastle, UK) was used essentially as described previously (44). Briefly, after selection of one sample as a reference, the retention times of all eluting precursor m/z values within the experiment were aligned. Precursors between two and ten charges were included and those with two isotopes or less excluded. After filtering, precursor abundances were normalized and replicate samples were grouped together. The corresponding MS2 spectra were exported and searched against the Human IPI database (v3.68) using Mascot (v2.3.0). The parameters corresponded to those used for previously mentioned peptide identification. The search results were exported in a .dat file format and filtered for 1 % PSM FDR and Mascot ion scores > 13.96

(or > 27.65 for non-phosphopeptides) using Rockerbox in combination with percolator (v1.0). Filtered PSMs were re-imported into Progenesis and matched to the respective features. Intensities < 2048 within the control dataset were filtered out. All data associated with protein and peptide quantification and identification are deposited as supplementary tables included in the ProteomeXchange upload with the identifier PXD001060

Results and discussion

Chromatographic phosphopeptide enrichment using an Fe-IMAC column

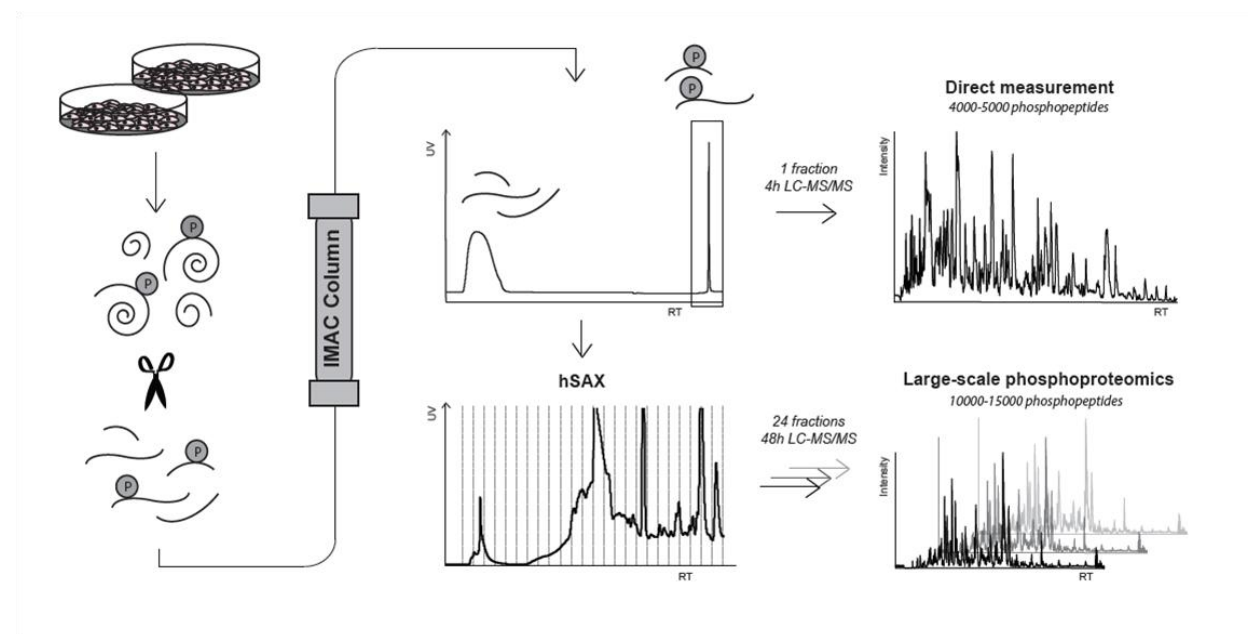


Figure 1. Workflow utilizing an Fe-IMAC HPLC column setup for the purification of phosphopeptides from complex digests. Fe-IMAC eluates can either be directly measured by LC-MS/MS providing considerable throughput or further fractionated by hSAX to provide increased analytical depth.

As discussed above, batch and tip-based enrichment protocols have been found to suffer from irreproducibility due to a number of factors including significant manual intervention. In addition enrichment selectivity and efficiency in such formats is also strongly dependent on the ratio of the amounts of enrichment material and protein digest used and which is hard to predict for any specific experimental context (12, 17-20). It was hypothesized that much of the quoted complementarity between different enrichment materials originates from practical shortcomings rather than differences in molecular properties of phosphopeptides. A series of experiments outlined below was conducted that support this hypothesis and the data collectively show that these practical issues can be substantially reduced by using a conventional Fe-IMAC HPLC column (Figure 1). First, batch and tip-based enrichments using TiO_2 , Ti-IMAC and Fe-IMAC beads with different bead-to-sample ratios were re-evaluated. The results confirm previous observations showing that comprehensiveness and selectivity is directly related to the bead-to-sample ratio (Figure 2A). Next, the use of a column format employing a commercially available IMAC HPLC column charged with Fe^{3+} ions was revisited. The UV trace shown in Figure 2B (1 mg of a trypsin digested A431 cell lysate) showed a clear separation of two different peptide

pools. As expected, LC-MS/MS analysis only detected non-phosphorylated peptides in the column breakthrough. In contrast 5,494 non redundant phosphopeptides were identified in the retained, peptide pool with > 90 % selectivity in a single 4 hour LC-MS/MS run on an Orbitrap Velos (HCD fragmentation). No peptides were detected in any of the other fractions of the chromatogram. The performance of the Fe-IMAC column was further characterized with respect to: (i) comprehensiveness, (ii) scalability and (iii) reproducibility. To address comprehensiveness, the flow through of the initial enrichment was dried down and reapplied to a second round of chromatographic enrichment on the same Fe-IMAC column. Both the weak UV trace and the very low number of identified phosphopeptides (five, of which four were also identified in the first enrichment) indicate virtually complete phosphopeptide depletion in a single

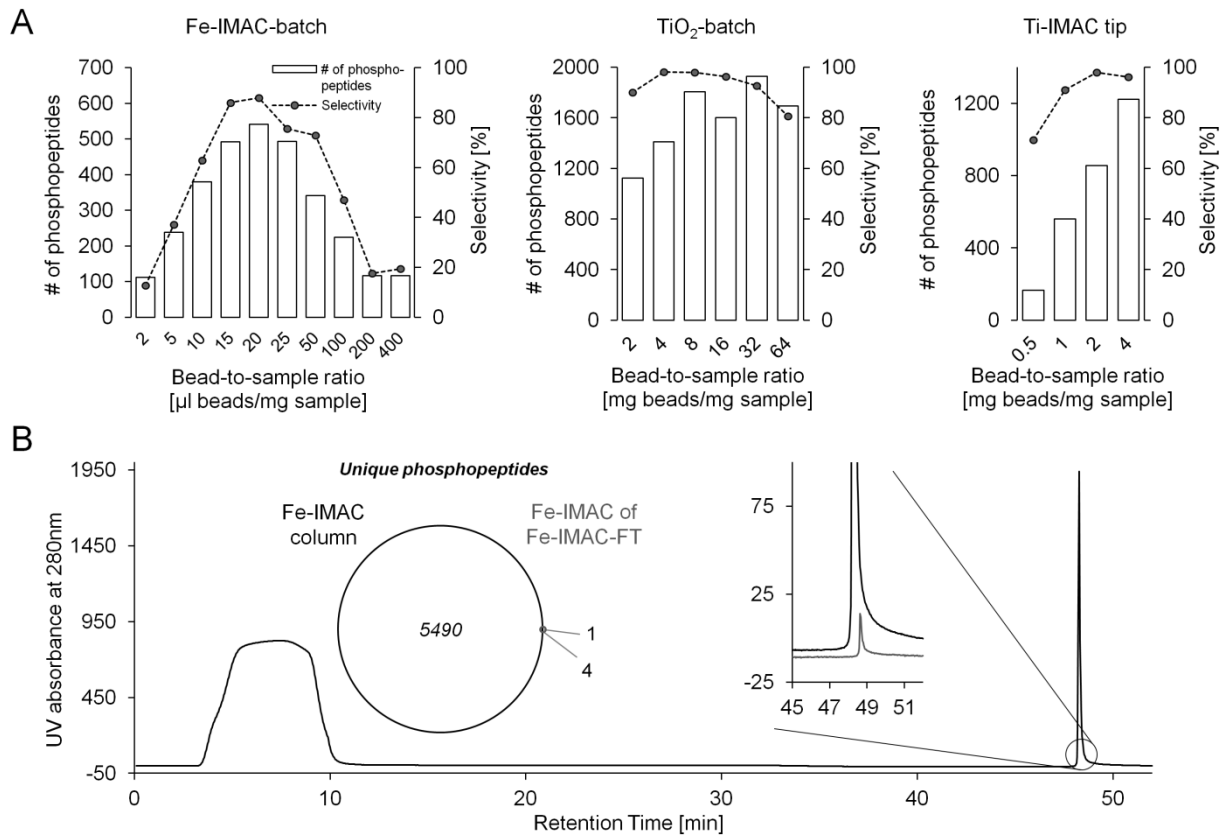


Figure 2. (A) Phosphopeptide number and selectivity in relation to the bead-to-sample ratio for Fe-IMAC batch, TiO₂ batch and Ti-IMAC tip enrichments. (B) Fe-IMAC UV chromatogram of 1 mg A431 cell digest. The phosphopeptide fraction (RT ~49 min) is well separated from the column breakthrough that contains the non-phosphorylated peptides. The inset on the left shows the number of phosphopeptides identified from the IMAC eluate and the number of phosphopeptides identified from a second IMAC enrichment of the column flow through (FT) of the first Fe-IMAC enrichment. The inset on the right shows the respective UV traces. Both indicate that the A431 digest was essentially depleted of phosphopeptides.

enrichment step (Figure 2B). Next, the scalability of the Fe-IMAC column (4 x 50 mm) was evaluated. Figure 3A shows that the chromatographic peak area scales linearly between 100 µg and 5 mg digest on column, indicating sufficient capacity within that range. Importantly, phosphopeptide selectivity exceeded 90 % in all cases (Figure 3B) demonstrating that the use of a Fe-IMAC column solves the issue of capacity without compromising on selectivity. As a result, there was no apparent overrepresentation of multiply phosphorylated peptides over other methods (8 % in this cell line, table 1), which would be expected to occur when exceeding the capacity of the enrichment material (14, 49).

The reproducibility of the Fe-IMAC column format was evaluated by the analysis and quantification of phosphopeptides from three independent Fe-IMAC enrichments from the same A431 cell lysate digest. Collectively, 7,510 unique phosphopeptides were identified (direct LC-MS/MS of IMAC eluates) of which 3,586 were identified in all three replicates (Table 1, Figure 3C). To account for the fact that many phosphopeptides are not picked for fragmentation due to the stochastic nature of data dependent acquisition (see also further below), the retention times of all peptide precursors were aligned using Progenesis LCMS. This boosted the number of unique phosphopeptides quantified across all replicates to 5,009, still indicating under-sampling by the mass spectrometer. To analyze quantitative reproducibility, quantifiable phosphopeptides were binned according to their coefficient of variation (CV).

Table 1. Summary of phosphopeptide identifications purified via Fe-IMAC column, Ti-IMAC tip, or TiO₂ batch format.

	Fe-IMAC column	Fe-IMAC column	Fe-IMAC column	Ti-IMAC tip	Ti-IMAC tip	Ti-IMAC tip	TiO ₂ batch	TiO ₂ batch	TiO ₂ batch
	R1	R2	R3	R1	R2	R3	R1	R2	R3
Unique phosphopeptides	5494	5585	5169	5397	4727	3706	4187	4173	3709
Unique phosphosequences	4467	4597	4206	4427	3879	3055	3617	3572	3135
Confidently localized phosphopeptides	3987	4127	3799	3956	3511	2694	3192	3168	2765
Non-p	561	681	856	26	77	377	267	217	156
Selectivity [%]	91	89	86	100	98	91	94	95	96
1p [%]	92	93	94	90	91	92	96	95	94
> 2p [%]	8	7	6	10	9	8	4	5	6
Unique phosphopeptides across triplicates		7510			6599			5763	
Unique phosphosequences across triplicates		5552			4906			4518	
Confidently localized phosphopeptides across triplicates		5490			4822			4315	

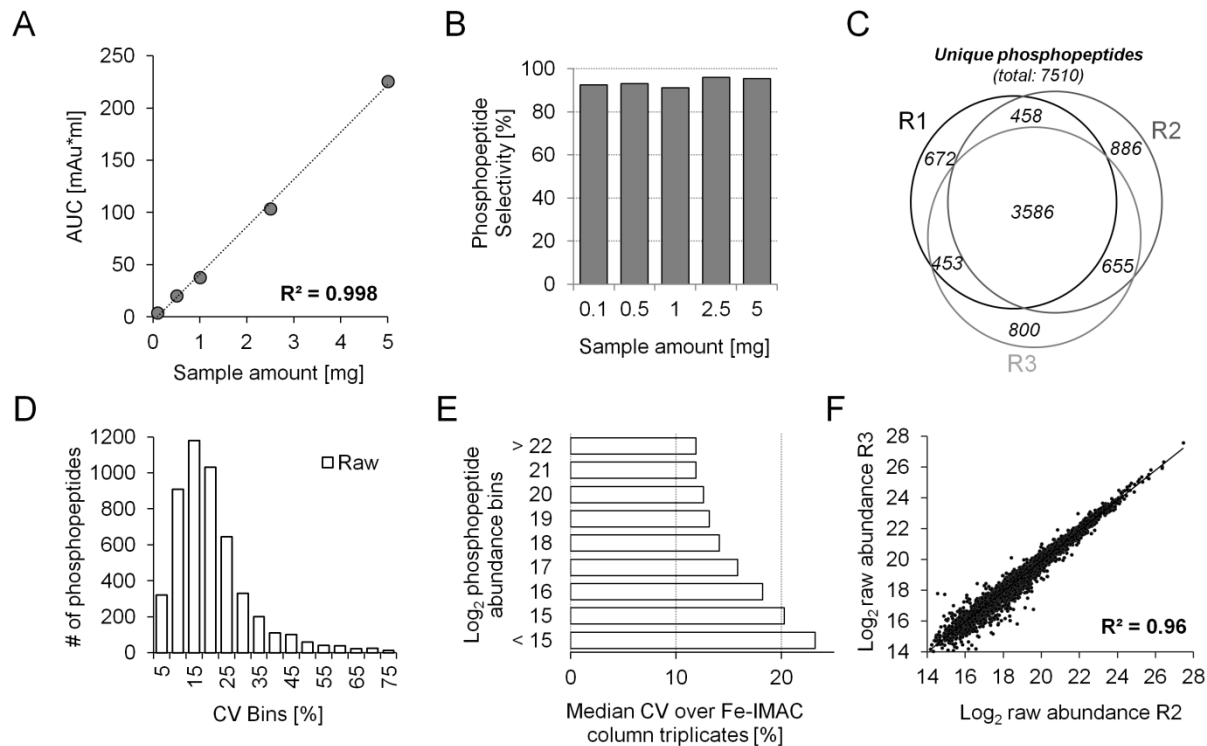


Figure 3. *Fe-IMAC column characterization. (A) UV signal-based quantification of Fe-IMAC enriched phosphopeptides as a function of the amount of cellular digest applied. It is apparent that the column captures phosphopeptides over a wide linear range. (B) Phosphopeptide selectivity for different input amounts (C) Venn diagram of the number of unique phosphopeptide identifications across three technical replicates of Fe-IMAC purification of phosphopeptides from 1 mg A431 lysate digests. (D) Reproducibility of raw and normalized technical Fe-IMAC column triplicates depicted as the number of quantifiable phosphopeptides per CV bin. (E) Median CV for triplicate Fe-IMAC column enrichments per \log_2 phosphopeptide intensity bin. (F) Raw \log_2 transformed phosphopeptide abundances of two technical Fe-IMAC column enrichment replicates.*

The median CV of phosphopeptide quantification was 15 % even for raw feature abundances (Figure 3D) and, as expected, variation was a function of feature abundance itself (Figure 3E). The excellent reproducibility between the replicates is further underscored by an R^2 value of 0.96 when plotting the raw phosphopeptide feature abundances of the respective replicates (Figure 3F). The above results indicate that Fe-IMAC HPLC column enrichments enable the comprehensive, reproducible and selective isolation of phosphopeptides over a wide range of input amounts. The workflow does not require special additives often used for e. g. TiO_2 methods to enhance selectivity and benefits from sample loading and elution conditions that are downstream compatible (e.g. for full proteome measurements or enrichment of different PTMs from the same digest (50)) without time consuming and sample losing desalting steps.

Apparent complementarity of methods is primarily caused by inefficient elution and insufficient mass spectrometric capacity

The Fe-IMAC column enrichment characterized above was benchmarked against Ti-IMAC in tip and TiO₂ batch enrichment. Ti-IMAC beads were synthesized according published procedures (51). One mg of A431 digest and optimized bead-to-sample ratios were used to perform triplicate enrichments for each method (Figure 2A). Both methods identified competitive numbers: 6,599 phosphopeptides (cumulative) for Ti-IMAC tip and 5,763 (cumulative) for TiO₂ batch (Table1). The overlap between replicates, phosphopeptide selectivity and the percentage of multiply phosphorylated peptides did also not differ profoundly compared to Fe-IMAC column triplicates (Table1). I next extracted phosphopeptide abundance (intensity) information using Progenesis and found that the intensity of triplicate Fe-IMAC column enrichments was seven and four times higher compared to the respective Ti-IMAC tip and TiO₂ batch triplicates (Figure 4A). This striking finding suggests that the Fe-IMAC column not only depleted phosphopeptides from complex digests (see above), but that it was also possible to elute the peptides very efficiently (see also further below). It should be noted that the Ti-IMAC beads were synthesized in house and hence their quality may not be entirely comparable to those published in the original paper (38, 51). Still, this observation provoked the question if phosphopeptide binding to or elution from Ti-IMAC or TiO₂ material is as efficient as for Fe-IMAC. A further observation was that the overlap between phosphopeptides identified by the different materials did not exceed 20 % (Figure 4B). This strong apparent complementarity led us to perform sequential enrichments from the same sample but using different materials. Surprisingly, neither Ti-IMAC tip nor TiO₂ batch purification recovered a considerable amount of phosphopeptides from Fe-IMAC column flow throughs (cumulative 19 for Ti-IMAC tip and 69 for TiO₂ batch; Figure 4C). I then reversed the enrichment order such that desalted TiO₂ batch and Ti-IMAC tip enrichment flow-throughs were subjected to Fe-IMAC column enrichment. Although several hundred phosphopeptides were recovered in this way (676 for Ti-IMAC and 501 for TiO₂; triplicate experiments; Figure 4C), the respective Fe-IMAC UV traces indicate that the enrichment efficiency of the Ti-IMAC tip and TiO₂ batch methods was likely still at least 90%. Interestingly, when comparing the recovered phosphopeptides from the Fe-IMAC flow through enrichments to those identified in the respective original triplicate analysis, a considerable overlap of 79 % for Ti-IMAC tip and 66 % for TiO₂ batch was found (Figure 4D) which is much higher than the 20 % figure quoted above, indicating that there is much less complementarity in the methods than previously thought.

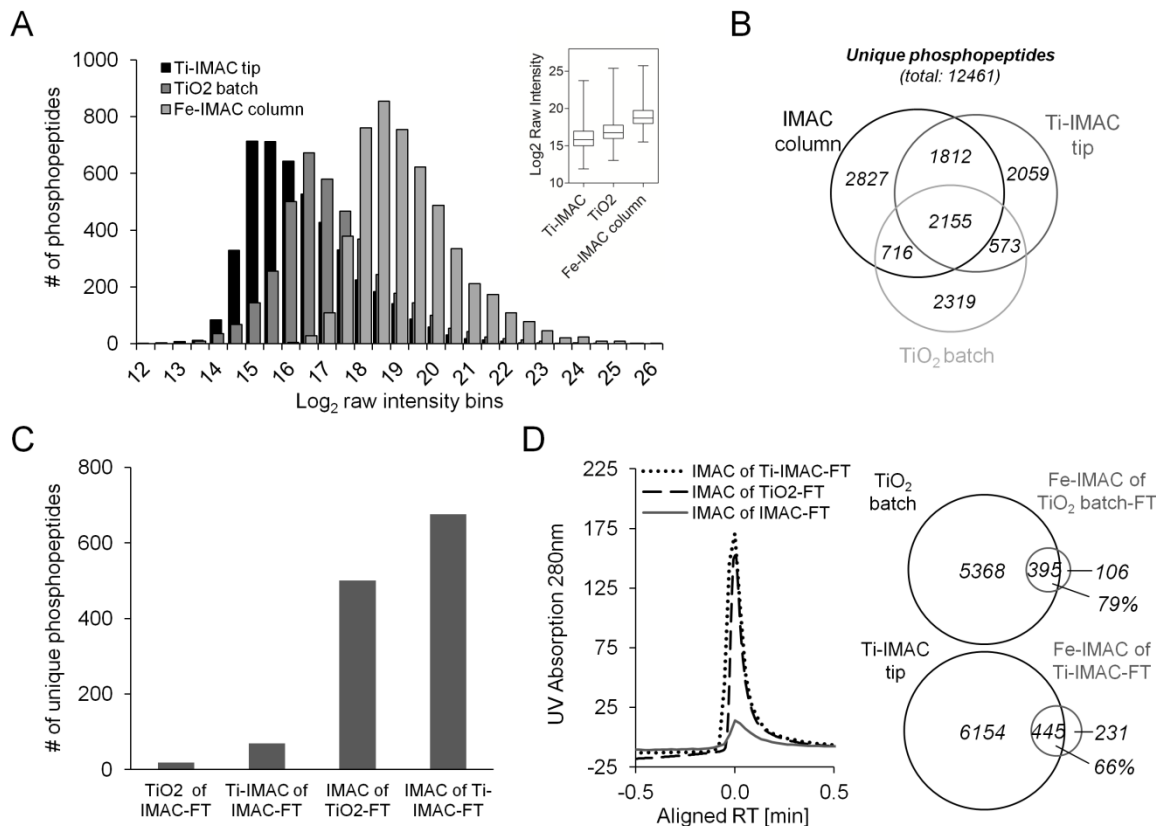


Figure 4. Comparison of Fe^{3+} -IMAC, Ti-IMAC tip and TiO_2 batch enrichments. (A) MS signal intensity distribution of unique phosphopeptides for each method showing the superiority of Fe-IMAC for phosphopeptide recovery from complex digests. The box plot (inset) summarizes the same data for a triplicate measurement. (B) A Venn diagram depicting the overlap of unique phosphopeptides of the combined Fe-IMAC column, Ti-IMAC tip and TiO_2 batch enrichment triplicate datasets. The fact that only ~20 % of the phosphopeptides are identified in all three datasets indicates apparent complementarity. (C) Summary of the number of unique phosphopeptide identifications (in triplicate) from experiments aiming at recovering phosphopeptides from the unbound fractions of one enrichment method using an alternative enrichment method. The data shows that Fe-IMAC recovers a noticeable number of phosphopeptides from the unbound fraction of TiO_2 batch and Ti-IMAC tip enrichments. TiO_2 batch and Ti-IMAC tip purifications from the unbound Fe-IMAC fraction do so to a much lesser extent. (D) UV chromatogram of Fe-IMAC enrichments from the unbound fractions of TiO_2 batch and Ti-IMAC tip experiments. The Venn diagrams (triplicate experiments) show that the phosphopeptides recovered by Fe-IMAC are largely overlapping with those identified in the original TiO_2 batch and Ti-IMAC tip enrichments, indicating insufficient capacity of the amount of TiO_2 and Ti-IMAC used.

Next, in order to distinguish material from format, a Ti-IMAC column was packed by a professional service and a TiO_2 HPLC column containing the very same materials as used in tip or batch format was purchased. For the Ti-IMAC column, the same loading conditions as for Fe-IMAC columns were used. Phosphopeptides eluted in a single peak but over a wider retention time window than for Fe-IMAC columns (Figure 5A). The flow through of the Ti-IMAC column was then subjected to Fe-IMAC column enrichment and the low UV signal and the identification of just 18 phosphopeptides in this analysis indicate a more efficient enrichment of

phosphopeptides by Ti-IMAC in column than tip format. For the TiO₂ column, very acidic (6 % TFA) loading conditions were needed in order to separate phosphopeptides from non-phosphorylated peptides. In addition, ten-fold higher ammonia concentrations had to be used in order to elute any phosphopeptides from both TiO₂ and Ti-IMAC HPLC columns. Surprisingly, phosphopeptide eluted from the TiO₂ column over a range of nearly 50 minutes (or 10 ml of solvent; Figure 5B), indicating rather inefficient elution presumably owing to very strong binding to the TiO₂ stationary phase. Fraction 4 contained most of the unique phosphopeptides and in subsequent fractions, the extent of redundant phosphopeptide identifications increased, another sign of inefficient elution (Figure 5B). Further evidence for this comes from an analysis of the amino acid composition of phosphopeptides identified in the different TiO₂ fractions (Figure 5C). Between the first and last fraction, the absolute percentage of acidic and hydrophobic amino acids within each fraction decreased by roughly 7 %. In contrast, the frequency of hydrophilic amino acids increased by 11 %. In other words: whereas acidic and hydrophobic phosphopeptides elute relatively easily off the column, otherwise hydrophilic peptides are more strongly retained. Some peptides may in fact never elute using ammonia as a solvent and, as a consequence, the observed bias in identified phosphopeptides between Fe-IMAC and TiO₂ can, at least in part, be attributed to inefficient elution from the TiO₂ material. In contrast, for Fe-IMAC columns both the sharp elution peak (Figure 5A) and the absence of phosphopeptide identifications in collected fractions before and after the main peak indicate efficient elution. It is surmised that the elution bias for Ti-IMAC is less pronounced than for TiO₂ because the amino acid compositions of phosphopeptides eluting from Ti-IMAC and Fe-IMAC are largely the same. To address if complete elution of phosphopeptides can be achieved from Fe-IMAC columns using ammonia, I performed two consecutive phosphopeptide enrichments using 1 mg A431 digest each and subsequently stripped the iron off the column using EDTA and analysed the phosphopeptide content of the EDTA fraction. In total, 208 unique phosphopeptides were identified of which 172 were high abundance peptides present in the prior ammonia elution (Figure 5D). The vast majority of these were 100-times less abundant in the EDTA fraction demonstrating that phosphopeptide elution with ammonia from Fe-IMAC columns is virtually complete.

A separate aspect that somewhat complicates the dissection of the various factors determining the phosphopeptide enrichment efficiency of the different methods tested is the fact that, in complex samples, the mass spectrometer cannot keep up with the number of peptides entering the instrument at any one time. For data dependent acquisition, peptide abundance is of course

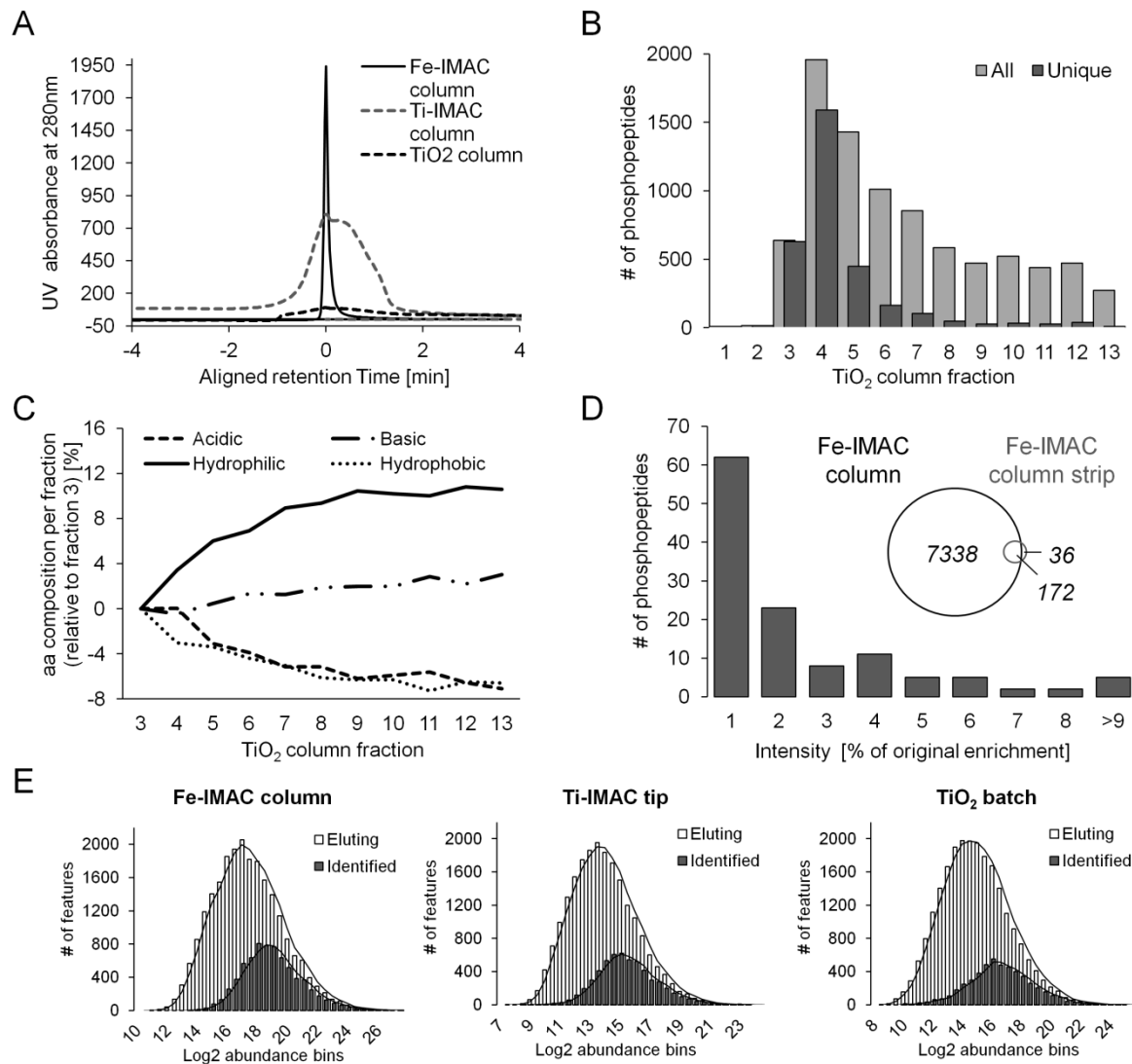


Figure 5. Elution efficiency. (A) UV absorbance comparison of the Ti-IMAC, Fe-IMAC and TiO₂ column phosphopeptide containing peak shows a broad elution profile for Ti-IMAC and TiO₂ columns which suggests a more inefficient elution compared to the Fe-IMAC column. (B) Number of unique and redundant phosphopeptides identified across fractions from a TiO₂ HPLC column enrichment of 1 mg A431 cell digest. The broad elution profile and the high degree of redundant identifications indicates inefficient elution. (C) Analysis of the absolute frequency of basic, acidic, hydrophilic and hydrophobic amino acids of phosphopeptides in the different TiO₂ column fractions (relative to the earliest fraction containing phosphopeptides, fraction 3). It is apparent that hydrophilic phosphopeptides are more difficult to elute from TiO₂ columns than other phosphopeptides. (D) When stripping the iron from Fe-IMAC columns, only very few phosphopeptides are identified in the stripped fraction. The majority of these are identified with merely 1 % of the intensity with which they were identified in a regular Fe-IMAC eluate indicating that the elution of Fe-IMAC columns is virtually complete. (E) Log₂ abundance bins of eluting and identified features across TiO₂ batch, Ti-IMAC tip and Fe-IMAC column direct measurement triplicates.

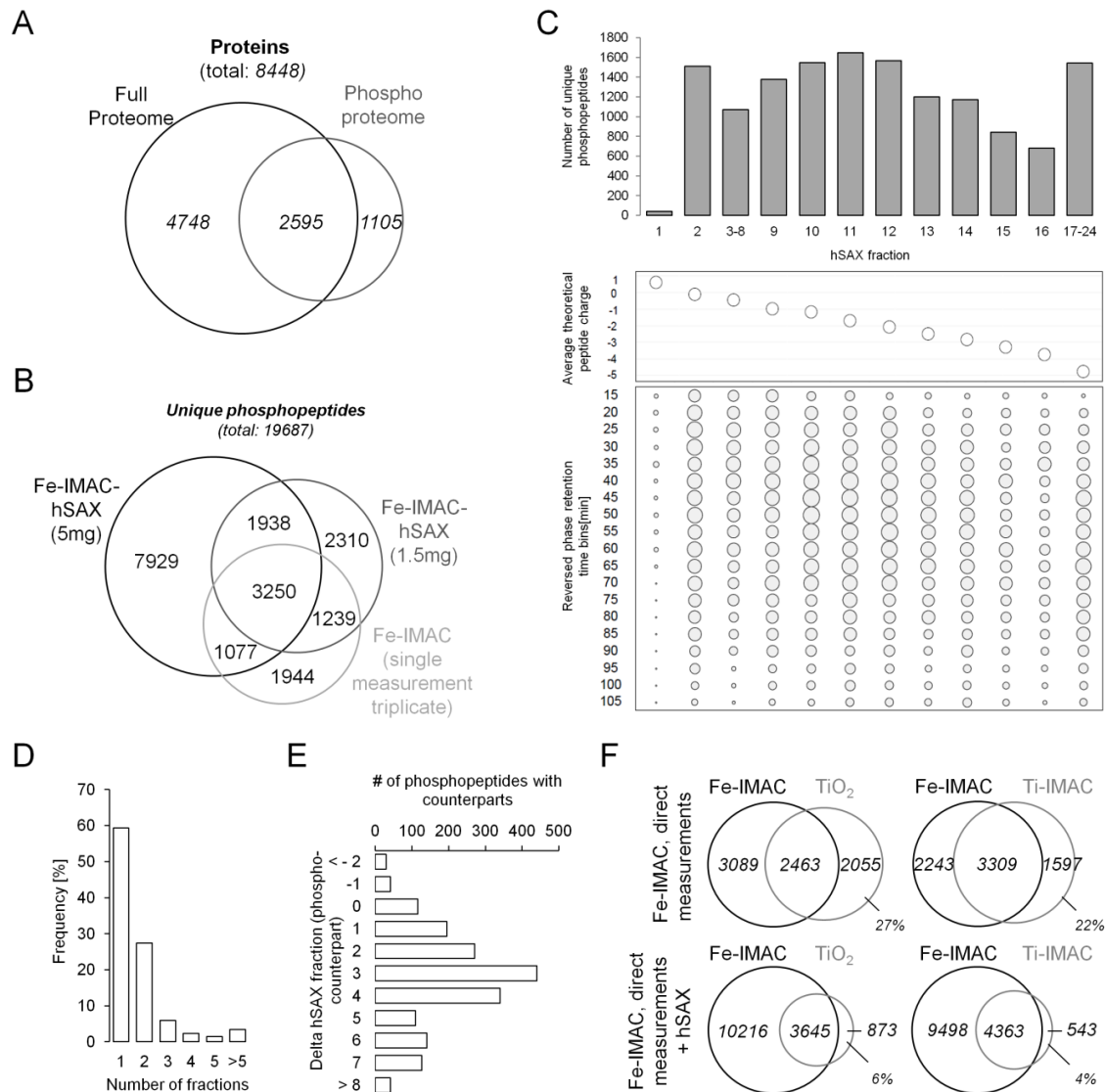


Figure 6. Hydrophilic SAX (hSAX) separation of Fe-IMAC eluates increases depth of phosphoproteome coverage. (A) Overlap of identified phosphoproteins from the 5 mg Fe-IMAC-hSAX dataset and the proteins identified with hSAX separation of 300 μ g A431 full proteome. (B) Comparison of phosphopeptides identified in Fe-IMAC column-based workflows with or without hSAX separation of Fe-IMAC eluates. (C) Top panel: number of phosphopeptide identifications across a hSAX separation of Fe-IMAC eluates. Middle panel: as one would expect, the average theoretical peptide charge decreases along the hSAX separation. Bottom panel: identification of phosphopeptides from hSAX fractions by reversed phase LC-MS/MS. The size of the dots indicates with the number of identified phosphopeptides. The data clearly demonstrates the strong orthogonality of the hSAX and reversed phase separations. (D) Separation efficiency of the hSAX fractionation shown as the percentage of phosphopeptides found in one or more fractions. (E) Difference in hSAX elution fraction for pairs of identified phosphopeptides and corresponding non-phosphorylated counterpart peptides. (F) Comparison of Fe-IMAC column (direct measurement, triplicate) to TiO_2 batch (direct measurement, triplicate) and Ti-IMAC tip (direct measurement, triplicate) phosphopeptide enrichments with or without Fe-IMAC hSAX separation. It is apparent that the number of phosphopeptides exclusively detected in TiO_2 batch or Ti-IMAC tip experiments is strongly reduced when the analytical depth of Fe-IMAC column enrichment is increased by a prior hSAX fractionation. This shows that often quoted orthogonality of the enrichment materials is much smaller than anticipated.

a major factor in determining which precursor is picked for fragmentation. Once the precursor ion complexity exceeds the fragmentation capacity of the mass spectrometer, the process becomes rather stochastic (35). In the conducted experiments, only 31 % of the eluting peptide like features (across triplicates) are actually identified (26 % for Ti-IMAC and 22 % for TiO₂) and those that are, predominantly represent highly abundant precursors (Figure 5E). As a consequence, the inherently biased nature of data dependent data acquisition is a major factor in the observed apparent complementarity between the different enrichment methods (12).

Large scale phosphoproteomics by coupling Fe-IMAC column enrichment to hSAX separation

As shown above, Fe-IMAC columns enable unbiased and comprehensive phosphopeptide enrichment and generally outperformed the other two methods/formats. Still, the restricted analytical capacity of the mass spectrometer used prevents very deep phosphoproteome analysis in direct LC-MS/MS measurements (Figure 5E). To increase analytical depth, fractionation of the isolated phosphopeptide pool can be performed prior to LC-MS/MS analysis. It was also hypothesized that the differences between the three phosphopeptide enrichment methods investigated in this study would diminish in such a setting as the aforementioned limitations of DDA would become less pronounced. The Trost laboratory in Dundee has recently shown good orthogonality of hydrophilic strong anion exchange and reversed phase chromatography for ordinary peptides (52). I therefore evaluated if hSAX would show the same desirable characteristics for phosphopeptides (52). In a first step, I reproduced hSAX performance for the fractionation of full proteome samples by separating 300 µg of A431 digest into 24 fractions and each of these was measured by LC-MS/MS using a 110 min reversed phase gradient. The identification of 48,250 unique peptides and 7,343 proteins in two days of instrument time corroborates the published results (Figure 6A). Next, 1.5 mg of A431 digest were subjected to Fe-IMAC column enrichment and separated into 24 hSAX fractions. This increased the number of identifications to 8,517 unique phosphopeptides (+13 % compared to direct measurements of 1 mg triplicates), but came at the expense of a three-fold increase in measurement time (Figure 6B). To test the merits of the combination of Fe-IMAC and hSAX further, the experiment was scaled up to 5 mg digest leading to the identification of 14,194 unique phosphopeptides (~3,700 proteins, of which 70 % were also identified in the full proteome in Figure 6A). As expected, hSAX separation of phosphopeptides closely followed their computed theoretical charge states (calculated for pH=8.5) and good orthogonality and separation efficiency compared to reversed phase separation was observed (Figure 6C and D).

Also in line with expectation, comparison of hSAX retention time differences between phosphopeptides and their respective non-phosphorylated counterparts (identified in the full proteome dataset) showed that the addition of a phosphate group considerably increased the binding to the hSAX material (Figure 6E), rendering the application of hSAX particularly suited for phosphopeptide fractionation. With a working 3D phosphopeptide separation scheme in hand, I revisited the apparent complementarity of phosphopeptide enrichment materials yet again. The Fe-IMAC data with and without hSAX separation was compared to the data obtained by triplicate direct LC-MS/MS measurement of TiO₂ and Ti-IMAC enriched samples. Figure 6F shows that the overlap of phosphopeptide identifications is substantially increased if the analytical depth is increased. Specifically, the fraction of exclusively identified phosphopeptides for TiO₂ enrichment goes down from 27 % to 6 % and from 22 % to 4 % for Ti-IMAC enrichment showing that these are actually purified by Fe-IMAC and are therefore not specific to purification on TiO₂ or Ti-IMAC resins. The remaining small discrepancy can likely still be attributed to insufficient sampling speed of the mass spectrometer.

Conclusion

In this study I have shown that the often quoted complementarity of different materials for phosphopeptide enrichment can actually be attributed to limited binding capacity, biased or incomplete elution, shortcomings in the physical formats (batch, tip) and limited analytical capacity of the mass spectrometer. All these can be largely overcome by using an Fe-IMAC HPLC column for bulk phosphopeptide enrichment from complex digests and an optional further hSAX separation step prior to LC-MS/MS analysis. These two chromatographic elements can be integrated into a flexible workflow (Figure 1) that allows the identification of 4-5,000 phosphopeptides within 4 hours of measurement time or 10-15,000 phosphopeptides within 48 hours of measurement time on a modestly performant mass spectrometer. Although the data shows that the Fe-IMAC column methodology scales over a wide range of digest quantities, it should be noted that downscaling samples would also require downscaling of the column. Alternatively, the Ti-IMAC material in tip format also gave good overall results at sample quantities below 100 μg digest. Owing to the fact that the Fe-IMAC column used here is commercially available, the method should be readily implemented. It can also be anticipated that the overall simplicity of the approach would also lead to improved reproducibility of results within one laboratory as well as better comparability of data acquired in different laboratories.

Acknowledgement

The author is indebted to Heiner Koch who performed initial experiments using the Fe-IMAC column during his M.Sc. thesis under my supervision and to Max Mundt and Guillaume Médard for their support with the synthesis of Ti-IMAC beads.

Abbreviations

CV	Coefficient of variation
DDA	Data dependent acquisition
ERLIC	Electrostatic repulsion hydrophilic interaction chromatography
FDR	False discovery rate
FLR	False localization rate
FT	Flow through
HCD	Higher energy collision induced dissociation
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High performance liquid chromatography
hSAX	Hydrophilic strong anion exchange
I.D.	Inner diameter
IDA	Iminodiacetate
IMAC	Immobilized metal ion affinity chromatography
LC	Liquid chromatography
PTM	Post translational modification
PTM	Post translational modification
SCX	Strong cation exchange
Ti-IMAC	Immobilized titanium (IV) ion affinity chromatography
TiO ₂	Titanium dioxide

References

1. Blume-Jensen, P., and Hunter, T. (2001) Oncogenic kinase signalling. *Nature* 411, 355-365.
2. Cohen, P. (2000) The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends in biochemical sciences* 25, 596-601.
3. Wilhelm, M., Schlegl, J., Hahne, H., Moghaddas Gholami, A., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J. H., Bantscheff, M., Gerstmair, A., Faerber, F., and Kuster, B. (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509, 582-587.
4. Lemeer, S., and Heck, A. J. (2009) The phosphoproteomics data explosion. *Curr Opin Chem Biol* 13, 414-420.
5. Andersson, L., and Porath, J. (1986) Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal Biochem* 154, 250-254.
6. Posewitz, M. C., and Tempst, P. (1999) Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal Chem* 71, 2883-2892.
7. Zhou, H., Xu, S., Ye, M., Feng, S., Pan, C., Jiang, X., Li, X., Han, G., Fu, Y., and Zou, H. (2006) Zirconium Phosphonate-Modified Porous Silicon for Highly Specific Capture of Phosphopeptides and MALDI-TOF MS Analysis. *Journal of Proteome Research* 5, 2431-2437.
8. Pinkse, M. W., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. (2004) Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal Chem* 76, 3935-3943.
9. Kweon, H. K., and Håkansson, K. (2006) Selective Zirconium Dioxide-Based Enrichment of Phosphorylated Peptides for Mass Spectrometric Analysis. *Analytical Chemistry* 78, 1743-1749.
10. Zhou, H., Low, T. Y., Hennrich, M. L., Toorn, H. v. d., Schwend, T., Zou, H., Mohammed, S., and Heck, A. J. R. (2011) Enhancing the Identification of Phosphopeptides from Putative Basophilic Kinase Substrates Using Ti (IV) Based IMAC Enrichment. *Molecular & Cellular Proteomics* 10, M110.006452.
11. Lai, A. C.-Y., Tsai, C.-F., Hsu, C.-C., Sun, Y.-N., and Chen, Y.-J. (2012) Complementary Fe³⁺- and Ti⁴⁺-immobilized metal ion affinity chromatography for purification of acidic and basic phosphopeptides. *Rapid Communications in Mass Spectrometry* 26, 2186-2194.
12. Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., and Aebersold, R. (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* 4, 231-237.
13. Tsai, C.-F., Hsu, C.-C., Hung, J.-N., Wang, Y.-T., Choong, W.-K., Zeng, M.-Y., Lin, P.-Y., Hong, R.-W., Sung, T.-Y., and Chen, Y.-J. (2013) Sequential Phosphoproteomic Enrichment through Complementary Metal-Directed Immobilized Metal Ion Affinity Chromatography. *Analytical Chemistry*.
14. Thingholm, T. E., Jensen, O. N., Robinson, P. J., and Larsen, M. R. (2008) SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Mol Cell Proteomics* 7, 661-671.
15. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jorgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4, 873-886.
16. Kettenbach, A. N., and Gerber, S. A. (2011) Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments. *Anal Chem* 83, 7635-7644.
17. Montoya, A., Beltran, L., Casado, P., Rodríguez-Prados, J.-C., and Cutillas, P. R. (2011) Characterization of a TiO₂ enrichment method for label-free quantitative phosphoproteomics. *Methods* 54, 370-378.
18. Li, Q. R., Ning, Z. B., Tang, J. S., Nie, S., and Zeng, R. (2009) Effect of peptide-to-TiO₂ beads ratio on phosphopeptide enrichment selectivity. *J Proteome Res* 8, 5375-5381.
19. Yue, X.-S., and Hummon, A. B. (2013) Combination of Multistep IMAC Enrichment with High-pH Reverse Phase Separation for In-Depth Phosphoproteomic Profiling. *Journal of Proteome Research* 12, 4176-4186.

20. Zhou, H., Di Palma, S., Preisinger, C., Peng, M., Polat, A. N., Heck, A. J. R., and Mohammed, S. (2013) Toward a Comprehensive Characterization of a Human Cancer Cell Phosphoproteome. *Journal of Proteome Research* 12, 260-271.
21. Ficarro, S. B., Salomon, A. R., Brill, L. M., Mason, D. E., Stettler-Gill, M., Brock, A., and Peters, E. C. (2005) Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites. *Rapid Commun Mass Spectrom* 19, 57-71.
22. Wang, J., Zhang, Y., Jiang, H., Cai, Y., and Qian, X. (2006) Phosphopeptide detection using automated online IMAC-capillary LC-ESI-MS/MS. *PROTEOMICS* 6, 404-411.
23. Pinkse, M. W. H., Mohammed, S., Gouw, J. W., van Breukelen, B., Vos, H. R., and Heck, A. J. R. (2008) Highly Robust, Automated, and Sensitive Online TiO₂-Based Phosphoproteomics Applied To Study Endogenous Phosphorylation in *Drosophila melanogaster*. *Journal of Proteome Research* 7, 687-697.
24. Ndassa, Y. M., Orsi, C., Marto, J. A., Chen, S., and Ross, M. M. (2006) Improved Immobilized Metal Affinity Chromatography for Large-Scale Phosphoproteomics Applications. *Journal of Proteome Research* 5, 2789-2799.
25. Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J., Rush, J., Lauffenburger, D. A., and White, F. M. (2005) Time-resolved Mass Spectrometry of Tyrosine Phosphorylation Sites in the Epidermal Growth Factor Receptor Signaling Network Reveals Dynamic Modules. *Molecular & Cellular Proteomics* 4, 1240-1250.
26. Zarling, A. L., Ficarro, S. B., White, F. M., Shabanowitz, J., Hunt, D. F., and Engelhard, V. H. (2000) Phosphorylated Peptides Are Naturally Processed and Presented by Major Histocompatibility Complex Class I Molecules in Vivo. *The Journal of Experimental Medicine* 192, 1755-1762.
27. Zarei, M., Sprenger, A., Gretzmeier, C., and Dengjel, J. (2012) Combinatorial Use of Electrostatic Repulsion-Hydrophilic Interaction Chromatography (ERLIC) and Strong Cation Exchange (SCX) Chromatography for In-Depth Phosphoproteome Analysis. *Journal of Proteome Research* 11, 4269-4276.
28. Zarei, M., Sprenger, A., Metzger, F., Gretzmeier, C., and Dengjel, J. (2011) Comparison of ERLIC-TiO₂, HILIC-TiO₂, and SCX-TiO₂ for Global Phosphoproteomics Approaches. *Journal of Proteome Research* 10, 3474-3483.
29. Dai, J., Wang, L.-S., Wu, Y.-B., Sheng, Q.-H., Wu, J.-R., Shieh, C.-H., and Zeng, R. (2009) Fully Automatic Separation and Identification of Phosphopeptides by Continuous pH-Gradient Anion Exchange Online Coupled with Reversed-Phase Liquid Chromatography Mass Spectrometry. *Journal of Proteome Research* 8, 133-141.
30. Villén, J., and Gygi, S. P. (2008) The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nature Protocols* 3, 1630-1638.
31. Engholm-Keller, K., Birck, P., Størling, J., Pociot, F., Mandrup-Poulsen, T., and Larsen, M. R. (2012) TiSH — a robust and sensitive global phosphoproteomics strategy employing a combination of TiO₂, SIMAC, and HILIC. *Journal of Proteomics* 75, 5749-5761.
32. Thakur, S. S., Geiger, T., Chatterjee, B., Bandilla, P., Fröhlich, F., Cox, J., and Mann, M. (2011) Deep and Highly Sensitive Proteome Coverage by LC-MS/MS Without Prefractionation. *Molecular & Cellular Proteomics* 10, M110.003699.
33. Pirmoradian, M., Budamgunta, H., Chingjin, K., Zhang, B., Astorga-Wells, J., and Zubarev, R. A. (2013) Rapid and deep human proteome analysis by single-dimension shotgun proteomics. *Molecular & cellular proteomics: MCP* 12, 3330-3338.
34. Köcher, T., Pichler, P., Swart, R., and Mechtler, K. (2012) Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. *Nature Protocols* 7, 882-890.
35. Michalski, A., Cox, J., and Mann, M. (2011) More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC-MS/MS. *Journal of Proteome Research* 10, 1785-1793.
36. Hebert, A. S., Richards, A. L., Bailey, D. J., Ulbrich, A., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2014) The One Hour Yeast Proteome. *Molecular & Cellular Proteomics* 13, 339-347.

37. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* 2, 1896-1906.
38. Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A. J., Zou, H., and Mohammed, S. (2013) Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat Protoc* 8, 461-480.
39. Seidler, J., Zinn, N., Haaf, E., Boehm, M. E., Winter, D., Schlosser, A., and Lehmann, W. D. (2011) Metal ion-mobilizing additives for comprehensive detection of femtomole amounts of phosphopeptides by reversed phase LC-MS. *Amino Acids* 41, 311-320.
40. Hahne, H., Pachi, F., Ruprecht, B., Maier, S. K., Klaeger, S., Helm, D., Médard, G., Wilm, M., Lemeer, S., and Kuster, B. (2013) DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nature Methods* 10, 989-991.
41. van den Toorn, H. W. P., Muñoz, J., Mohammed, S., Rajmakers, R., Heck, A. J. R., and van Breukelen, B. (2011) RockerBox: analysis and filtering of massive proteomics search results. *Journal of proteome research* 10, 1420-1424.
42. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nature methods* 4, 923-925.
43. Marx, H., Lemeer, S., Schliep, J. E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A. J. R., and Kuster, B. (2013) A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nature Biotechnology* 31, 557-564.
44. Wu, Z. X., Doondeea, J. B., Gholami, A. M., Janning, M. C., Lemeer, S., Kramer, K., Eccles, S. A., Gollin, S. M., Grenman, R., Walch, A., Feller, S. M., and Kuster, B. (2011) Quantitative Chemical Proteomics Reveals New Potential Drug Targets in Head and Neck Cancer. *Molecular & Cellular Proteomics* 10.
45. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 41, D808-815.
46. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-2504.
47. Mi, H., and Thomas, P. (2009) PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol Biol* 563, 123-140.
48. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic acids research* 40, D261-270.
49. Ye, J., Zhang, X., Young, C., Zhao, X., Hao, Q., Cheng, L., and Jensen, O. N. (2010) Optimized IMAC-IMAC Protocol for Phosphopeptide Recovery from Complex Biological Samples. *Journal of Proteome Research* 9, 3561-3573.
50. Mertins, P., Qiao, J. W., Patel, J., Udeshi, N. D., Clauser, K. R., Mani, D. R., Burgess, M. W., Gillette, M. A., Jaffe, J. D., and Carr, S. A. (2013) Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nature Methods* 10, 634-637.
51. Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A. J. R., Zou, H., and Mohammed, S. (2013) Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nature Protocols* 8, 461-480.
52. Ritoro, M. S., Cook, K., Tyagi, K., Pedrioli, P. G. A., and Trost, M. (2013) Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *Journal of Proteome Research* 12, 2449-2457.

Chapter 3

MALDI-TOF- and nESI-Orbitrap-MS/MS identify orthogonal parts of the phosphoproteome

Abstract

Phosphorylation is a reversible posttranslational protein modification which plays a pivotal role in intracellular signaling. Despite extensive efforts, phosphorylation site mapping of proteomes is still incomplete motivating the exploration of alternative methods that complement existing workflows. In this study, I compared tandem mass spectrometry (MS/MS) on matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) and nano-electrospray ionization (nESI) Orbitrap instruments with respect to their ability to identify phosphopeptides from complex proteome digests. Phosphopeptides were enriched from tryptic digests of cell lines using Fe-IMAC column chromatography and subjected to LC-MS/MS analysis. It was found that the two analytical workflows exhibited considerable orthogonality. For instance, MALDI-TOF-MS/MS favored the identification of phosphopeptides encompassing clear motif signatures for acidic residue directed kinases. The extent of orthogonality of the two LC-MS/MS systems was comparable to that of using alternative proteases such as Asp-N, Arg-C, chymotrypsin, Glu-C and Lys-C on just one LC-MS/MS instrument. Notably, MALDI-TOF-MS/MS identified an unexpectedly high number and percentage of phosphotyrosine sites (~20 % of all sites), possibly as a direct consequence of more efficient ionization. The data clearly show that LC-MALDI-MS/MS can be an useful complement to LC-nESI-MS/MS for phosphoproteome mapping and particularly so for acidic and phosphotyrosine containing peptides.

Introduction

Protein phosphorylation is a reversible posttranslational protein modification regulating intracellular signaling and many other fundamental biological processes (1). More than half of the (human) proteome is already known to be phosphorylated at least once (2) and predictions extrapolate this to be the case for as much as 90% of all mammalian proteins (3). In recent years, mass spectrometry-based proteomics has emerged as the prime technology for the qualitative and quantitative study of dynamic protein phosphorylation. Still, due to the often substoichiometric nature of the modification (4) as well as limitations in analytical technology, phosphopeptide enrichment prior to analysis is still generally required. Methods such as TiO₂ (5), Ti-IMAC (6), and Fe-IMAC (7,8) that all exploit the affinity of phosphate groups towards metal ions, enable the global enrichment of the phosphoproteome. Immuno-affinity methods based on immobilized antibodies directed against phosphorylated epitopes such as phosphotyrosine (9, 10, 11) or specific phosphorylation motifs (12), are also widely applied to study a specific part of the phosphoproteome.

Despite the impressive technological improvements in the past, enabling the identification and quantification of thousands of phosphorylation sites from a single biological system, the coverage of all known, let alone potential sites is far from complete (1). Two confounding factors in many current approaches are the fact that bottom up proteomics heavily relies on trypsin as the sole protease for protein digestion and on electrospray for peptide ionization. Albeit powerful, it is well known that both limit the part of the proteome that can be detected. Proteases with alternative and orthogonal cleavage specificities such as Asp-N, Arg-C, chymotrypsin, Glu-C and Lys-C are efficient means to increase coverage of the proteome (13, 14) and phosphoproteome (15, 16). Similarly, it has been shown that proteomic sequence coverage can be extended when using an alternative ionization technique because nanoelectrospray ionization (nESI) (17, 18) and matrix assisted laser desorption ionization (MALDI) (19, 20, 21), favor different physicochemical properties of peptides for ionization (22, 23, 24). Much of the early work comparing the merits of the two ionization methods used samples of low complexity (i. e. numbers of peptides analyzed), thus limiting the insight that could be obtained to a few dominating effects. Recent work from Hessling *et al.* (24), extended the comparison of the two ionization methods to a more global proteomic scale and the data generally confirmed the previously observed complementarity of nESI and MALDI. Parker *et al.* (25) as well as Aryal *et al.* (26) also showed orthogonality of MALDI and nESI for phosphopeptide analysis albeit for a limited sample set only. Therefore, in this study, I revisited

and extended the previous work to more complex phosphopeptide mixtures and more comprehensive experimental workflows. The data clearly shows that the two ionization methods are indeed partially orthogonal. In particular, MALDI identified a substantially higher portion of acidic phosphopeptides than nESI. Interestingly, the gain of employing MALDI analysis on the same sample analysed by nESI was comparable to using orthogonal proteases. MALDI also identified an unusually high number and percentage of phosphotyrosine sites including many that could not be detected by nESI even following extensive chromatographic fractionation or immunoprecipitation.

Materials and methods

Cell culture

Human epidermoid A431 cells were grown in Iscove's Modified Dulbecco's Medium (IMDM, Biochrom AG) supplemented with 10 % (v/v) Fetal Bovine Serum (Biochrom AG) and 1 % (v/v) Antibiotic/Antimycotic solution (Sigma Aldrich Corp.). Prior to lysis cells were treated with 1 mM pervanadate for 5 min. After harvesting, cells were washed two times with PBS and lysed in 8 M Urea, 40 mM Tris/HCl (pH 7.6) containing 1 x complete mini EDTA-free protease inhibitor (Roche Diagnostics) and 1 x phosphatase inhibitor cocktail (Sigma-Aldrich). The lysate was centrifuged at 20,000 x g for 1 h at 4 °C. Protein concentration was determined using the Bradford method (Coomassie (Bradford) Protein Assay Kit, Thermo Scientific). The supernatant was reduced with 10 mM DTT at 56 °C for 1 h and subsequently alkylated with 25 mM iodoacetamide for 45 min at room temperature in the dark.

Digestion and desalting

One mg of starting material from one A431 lysate batch was digested using six different enzymes: Trypsin (sequencing grade modified, Promega Corp., Madison), Asp-N (sequencing grade, Promega Corp., Madison), Glu-C (sequencing grade, Promega Corp., Madison), chymotrypsin (sequencing grade, Promega Corp., Madison), lysyl-endopeptidase (Lys-C, WAKO Chemicals GmbH, Neuss) and Arg-C (sequencing grade, Promega Corp., Madison). For trypsin digestion an enzyme-to-substrate ratio of 1 : 50 (20 µg trypsin) was used and the lysate was diluted 1 : 5 with 40 mM Tris/HCl (pH 7.6). For Asp-N digestion an enzyme-to-substrate ratio of 1 : 250 (4 µg Asp-N) was used and the lysate was diluted 1 : 3 with 40 mM Tris/HCl (pH 7.5). For Glu-C digestion an enzyme-to-substrate ratio of 1 : 100 (20 µg Glu-C) was used and the lysate was diluted 1 : 9 with 40 mM Tris/HCl (pH 7.5). For chymotrypsin digestion an enzyme-to-substrate ratio of 1 : 100 (20 µg chymotrypsin) was used and the lysate was diluted 1 : 7.6 with 40 mM Tris/HCl (pH 8.0) and 10mM CaCl₂. For Lys-C digestion an enzyme-to-substrate ratio of 1 : 100 (20 µg Lys-C) was used and the lysate was diluted 1 : 2.7 with 40 mM Tris/HCl (pH 8.0). For Arg-C digestion an enzyme-to-substrate ratio of 1 : 100 (20 µg Arg-C) was used and the lysate was diluted 1 : 7.6 with 40 mM Tris/HCl (pH 7.8), 5 mM CaCl₂, 2 mM EDTA and 10 mM DTT. All digestion reactions, except for chymotrypsin which was performed at 25°C, were carried out at 37°C and over night. The next day, digests were acidified with TFA to a pH of ~2 in order to stop enzymatic activity. SepPack columns (C18

cartridges, Sep-Pak Vac 1cc (50mg), Waters Corp.) were used for peptide desalting according to manufacturer's instructions. After elution with 0.3 ml of solvent B (0.07 % TFA, 50 % ACN), the volume was adjusted to 0.5 ml using solvent A (0.07 % TFA) in order to reduce the ACN concentration from 50 % to 30 %.

Phosphopeptide enrichment using a Fe-IMAC column

A Fe-IMAC column (4mm I.D. x 50 mm, ProPac IMAC-10, Thermo Fisher Scientific) was connected to an HPLC system (ÄKTA explorer FPLC system, Amersham Pharmacia) with a 1 ml sample loop and charged with iron according to manufacturer's instructions. Briefly, the column was rinsed with three column volumes of 20 mM formic acid and charged using three column volumes of 25 mM FeCl₃, 100 mM acetic acid. To wash off unbound iron-ions the column was flushed with 20 column volumes of 20 mM formic acid and subsequently removed from the HPLC system. The HPLC lines were first flushed with 20 ml ddH₂O water, followed by 20 ml 50 mM EDTA and 10 ml ddH₂O to remove remaining iron-ions. After re-connection, the column was equilibrated to baseline with Fe-IMAC solvent A (0.07 % TFA, 30 % ACN). The desalted digest (0.5 ml in 0.07 % TFA, 30 % ACN) was directly loaded (0.1 ml/min, 10 min) and unbound peptides were washed out with Fe-IMAC solvent A (0.3 ml/min, 16 min). Subsequent phosphopeptide elution was achieved by a linear gradient from 0 % to 45 % Fe-IMAC solvent B (0.5 % NH₄OH) (0.2 ml/min, 60min). After an increase to 100 % solvent B and a holding step for 5 min the column was re-equilibrated by switching to Fe-IMAC solvent A (30 min, 0.5 ml/min). 1 ml of the flow through and 1 ml of the phosphopeptide containing eluate were collected, dried down and stored at -80 °C. The column was recharged after a maximum of six enrichments.

hSAX separation

The dataset of hydrophilic strong anion (27) separated phosphopeptides (enrichment and separation of 5 mg pervanadate treated A431 digest) was taken from a previous study carried out in our laboratory (8). The dataset has previously been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD001060.

Phosphotyrosine immunoprecipitation

Immunoaffinity purification of phosphotyrosine containing peptides using P-Tyr-1000 antibody resin (PTMScan® Phospho-Tyrosine Rabbit mAb (P-Tyr-1000), Cell Signaling Technology), was essentially carried out as described by the manufacturer. Briefly, 80 µl of antibody beads were

washed four times with 1 ml of PBS and three times with 1 ml of IP buffer. 5 mg of digested and dried peptides were resuspended in IP buffer and combined with the beads. The sample was left slowly shaking at 4 °C over night. The next day beads were washed with 2 x 1 ml of IP buffer and 3 x 1 ml ddH₂O to remove unbound peptides. Bound peptides were eluted with 55 µL of 0.15 % TFA for 10 min. The elution step was repeated with 45 µL of 0.15 % TFA. Eluates were combined, dried down and desalted using StageTips (28).

LC- nESI-MS/MS

Nanoflow LC-MS/MS was performed by coupling an Eksigent nanoLC-Ultra 1D+ (Eksigent, Dublin, CA) to an Oribtrap Velos (Thermo Scientific, Bremen, Germany). Peptides were delivered to a trap column (100 µm I.D. x 2 cm, packed with 5 µm C18 resin, Reprosil PUR AQ, Dr. Maisch, Ammerbuch, Germany) at a flow rate of 2 µl/min for 12 min and 5 µl/min for 13 min in 100 % loading solvent A (0.1 % FA in ddH₂O). After 25 minutes of loading and washing, peptides were transferred to an analytical column (75 µm x 40 cm C18 column Reprosil PUR AQ, 3 µm, Dr. Maisch, Ammerbuch, Germany) using solvent A (0.1 % FA, 5 % DMSO in ddH₂O) and separated with a linear gradient from 2 % to 27 % solvent B (0.1 % FA, 5 % DMSO in ACN) at a flow rate of 300 nL/min (29). A 210 min gradient (with a total turnaround time of 235 min) was applied to measure Fe-IMAC column enrichments, whereas phosphotyrosine enrichments were separated with a linear 110 min gradient (135 min turnaround). To prevent potential iron adsorption to the reversed phase stationary phase, 5 µl of 1 µM deferoxamine dissolved in ddH₂O was injected in between phosphopeptide measurements (30).

Peptides were ionized using 2.2 kV spray voltage and a capillary temperature of 275 °C. The mass spectrometer was operated in data dependent acquisition mode, automatically switching between MS and MS/MS. Full scan MS spectra (m/z 360 – 1300) were acquired in the Orbitrap at 30,000 (m/z 400) resolution, an AGC target value of 1e6 and a maximum injection time of 100 ms. For internal calibration the signal at m/z 401.922718 was used as a lock mass (29). High resolution HCD-MS/MS spectra were generated for up to 10 precursors with a normalized collision energy of 35 %. The precursor ion count for triggering an MS/MS event was set to 500 with a repeat count of 1, a repeat duration of 10 s and a dynamic exclusion of 20 s. Fragment ions were read out in the Orbitrap mass analyzer at a resolution of 7,500 (isolation window 2 Th). The MS/MS AGC target value was set to 4e4 with a maximum ion injection time of 250 ms.

LC-MALDI-MS/MS

Samples were dissolved in 14 μl of a 5 % ACN, 0.1 % FA solution (5 min sonification). 7.9 μl of each sample was injected (NanoAcquity, full loop injection, overfill factor of 1.2; loading time of 60 min with a flow of 300 nl/min), separated with a 110 min gradient (5 - 40 %, ACN with 0.1 % TFA; flow rate of 350 nl/min) and spotted into 1200 spots (Sunchrom Spotter, 8 sec per spot, 1 $\mu\text{l}/\text{min}$ DHB matrix (10 mg/ml; in 50 % ACN; 0.1 % TFA; 1 % phosphoric acid; standard peptides, 10 min delay). A mass range from 750 to 4000 Da was selected with a focus mass of 2000 Da. For MS spectra, a total of 8 x 250-shot sub-spectrum was accumulated (i.e., 2000 shots per spectrum). The sample plate was moved with continuous stage motion with a stage velocity of 1000 $\mu\text{m}/\text{sec}$. The laser pulse rate was set to 400 Hz. The method contained peak smoothing with FFT and Poisson Denoise. Minimal S/N for peak detection was specified as 5. Cluster area S/N optimization was performed with a threshold of 20. Internal calibration was used with a minimal S/N of 10 and a mass tolerance of ± 0.3 m/z. At least 1 peak needed to match and the maximal outlier error was 10 ppm. Only monoisotopic peaks were used. The m/z of the four peptides from the internal peptide spike-in were used as reference masses. The interpretation method selected monoisotopic precursors within a mass range of 750 to 4000 Da (minimum signal to noise set to 50) over the full retention time range for MS/MS. A maximum of 5 precursors per fraction were selected and MS/MS spectra of the strongest precursors (based on MS1 intensity) were acquired first. Exclusion lists contained masses with decimal place that could be excluded for peptides. Salt adduct masses 21.982 and 37.956 were excluded via adduct exclusion lists and adduct tolerance m/z was set to ± 0.03 . In addition, precursors were excluded within a resolution of 200. Minimum chromatogram peak width was required to be 1 fraction and fraction to fraction precursor mass tolerance was set to 200 ppm. 1 kV positive mode with collision induced dissociation and automatic acquisition control. A relative precursor mass window of 200 resolution (FWHM) was defined. Metastable suppressor was activated. For MS/MS spectra, a maximum of 12 x 250-shot sub-spectrum was accumulated. Acceptance of MS/MS spectra was passed after 12 sub-spectra or earlier if the final spectrum reached the desired high quality (maximal 3000 shots per spectrum). The sample plate was moved with continuous stage motion with a stage velocity of 800 $\mu\text{m}/\text{sec}$. A laser pulse rate of 1000 Hz was selected. The processing method specified usage of Savitsky-Golay for peak smoothing with 5 points across peaks and a 4th polynomial order. Minimal S/N for peak detection was set to 15 and the minimal peak width at full width half maximum (FWHM) was 1.5 bins. Cluster area S/N optimization was performed with a threshold of 15. Calibration was the default.

Peptide and protein identification

For peptide and protein identification, peak lists were extracted from raw files using Mascot Distiller v2.4.3.0 (Matrix Science, UK) and subsequently searched against the human Uniprot database (26.10.2010, containing 88,381 entries) using Mascot (v2.4.1) with the following parameters: carbamidomethyl cysteine as a fixed modification; phosphorylation of serine, threonine and tyrosine, oxidation of methionine and N-terminal protein acetylation as variable modifications. For spectra acquired with nESI, precursor tolerance was set to 10 ppm (50 ppm for MALDI) and fragment ion tolerance to 0.05 Da (0.2 Da for MALDI). For each proteolytic enzyme up to 2 missed cleavage sites were allowed. Identified phosphopeptides were filtered using two criteria: (i) 1 % PSM FDR using Rockerbox (31) in combination with percolator (v1.0) (32) (ii) remaining peptides corresponding to accepted PSMs were filtered according to a 5 % global peptide FDR cutoff based on the formula published by Marx et al. (33) (HCD, global FDR). The peptide table created after Rockerbox processing was used for further analysis. To filter for confidently localized phosphopeptides and phosphorylation sites an additional 5 % FLR cutoff based on the Mascot delta score was applied (33). Statistically significant motif positions were extracted using pLOGO (34). KinomeXplorer was used for the analysis of kinase substrate relationships which were subsequently filtered for NetworkKin scores > 6 (35).

Data availability

The mass spectrometry proteomics data and associated supplementary information have been deposited to the ProteomeXchange Consortium (36) via the PRIDE partner repository with the dataset identifier PXD003114.

Results and discussion

ESI and MALDI ionization lead to the identification of orthogonal sets of phosphorylation sites

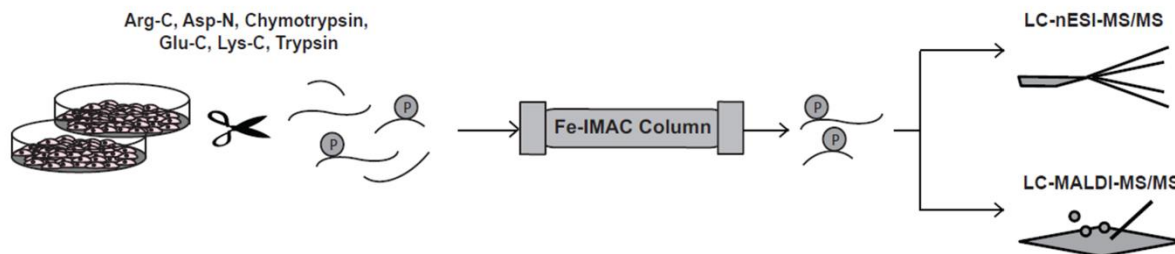


Figure 1. Schematic workflow depicting a Fe-IMAC HPLC column setup for the purification of phosphopeptides from complex lysates, which were digested using either Arg-C, Asp-N, chymotrypsin, Glu-C, Lys-C or trypsin. One half of each sample was measured by nESI-MS/MS (Thermo Orbitrap Velos) and the other half by MALDI-MS/MS (SCIEX 5800 TOF/TOF).

As complementarity of MALDI and nESI for peptide identification in bottom up proteomics has been demonstrated for unmodified peptides already (22, 24), I sought to investigate to which extent this phenomenon extends to phosphopeptides. To address this, I first treated human A431 cells with sodium pervanadate in order to boost tyrosine phosphorylation, then digested the proteome and performed phosphopeptide enrichment using IMAC column chromatography (8) and subjected identical samples to LC-MS/MS measurements on a SCIEX 5800 MALDI TOF/TOF and a nESI Orbitrap Velos (Figure 1). From single analysis runs, each representing half of one mg enriched protein digest, I identified a total of 4,353 unique phosphorylation sites using nESI and MALDI (6,773 non-redundant phosphopeptides representing 2,095 proteins) and using Mascot at a false localization rate (FLR) of $< 5\%$ (33) (Figure 2A). Although much fewer phosphorylation sites were identified by LC-MALDI compared to LC-nESI-MS/MS, 50 % of the MALDI sites (852) were not detected by nESI (Figure 2A). To ascertain that this was not simply owing to under sampling by the nESI mass spectrometer, the nESI analysis was repeated using five times the amount of starting material (five mg protein digest from the same sample) and separating the Fe-IMAC phosphopeptide pool into 24 fractions by hydrophilic strong anion chromatography (hSAX) (27) (the dataset has previously been acquired and used (8)). This analysis resulted in the identification of $> 11,000$ phosphopeptides corresponding to 9,171 phosphorylation sites at a FLR $< 5\%$. Despite the much deeper coverage, 36 % of the phosphorylation sites identified by LC-MALDI (619) were still not contained in the much broader

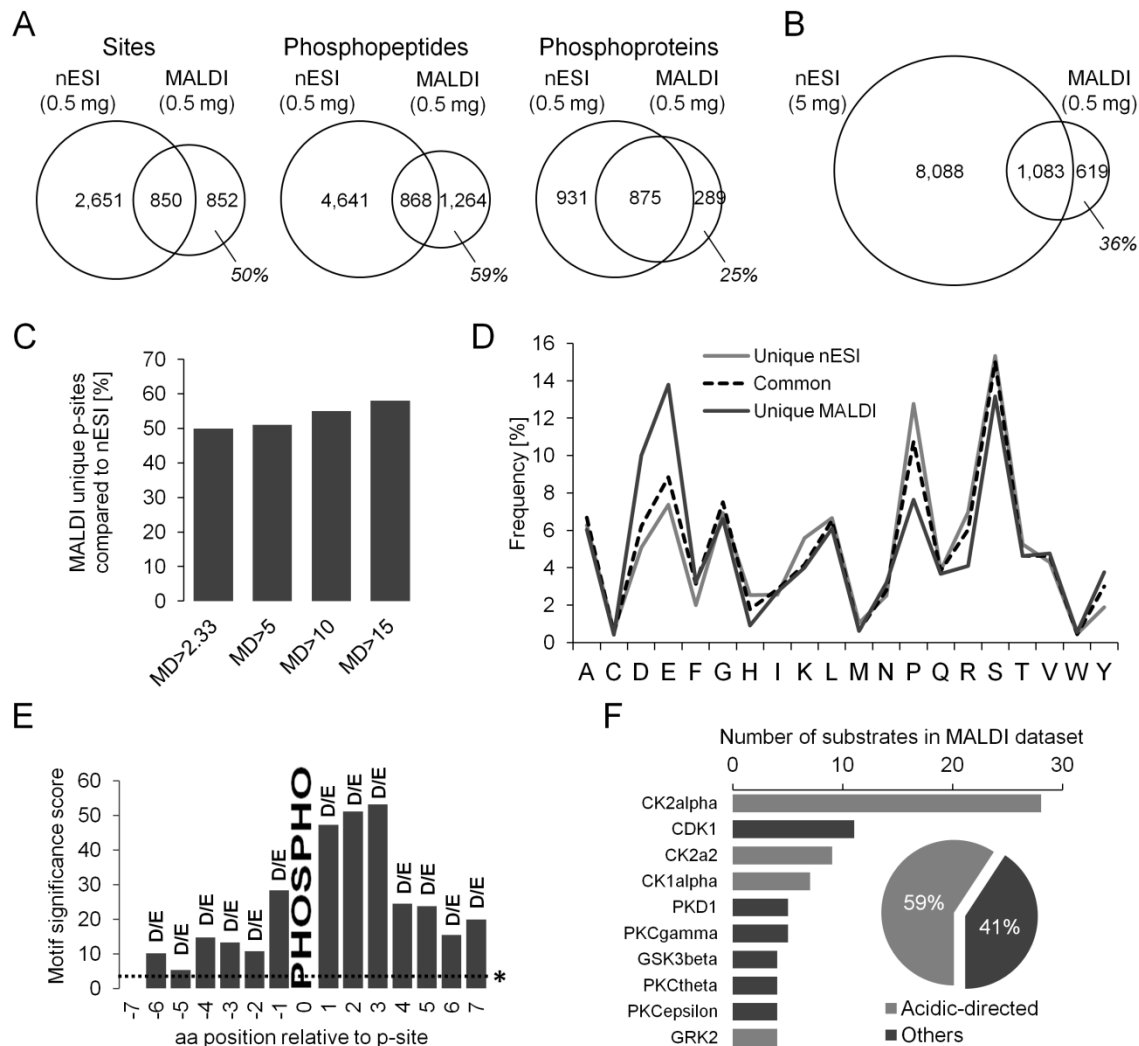


Figure 2. Comparison of nESI and MALDI single measurements of the same phosphopeptide enrichment. (A) Number of unique phosphorylation sites (FLR < 5 %), phosphopeptides and phosphoproteins for single measurements using nESI (Orbitrap Velos) and MALDI (SCIEX 5800 TOF/TOF) identification from the same A431 digest (1 mg total). (B) Even if nESI phosphorylation site coverage is increased due to extensive phosphopeptide separation (hSAX, 24 fractions), 36 % of the sites identified by MALDI single measurements remain unique. (C) Unique phosphorylation sites of MALDI compared to nESI single measurements (in percent of the MALDI dataset) are plotted dependent on the Mascot Delta score. The increasing percentage of uniquely identified sites suggests that orthogonality is not caused by false phosphorylation site assignment. (D) Global amino acid frequency distribution of unique nESI phosphopeptides, unique MALDI phosphopeptides and those overlapping shows a clear enrichment of acidic amino acids in the MALDI dataset. (E) Compared to a whole proteome background, acidic residues which are preferentially ionized by MALDI, are highly overrepresented at positions surrounding the phosphorylation site. Motif significance scores were calculated using pLOGO; the dotted line marked with an asterisk indicates statistical significance (p -value < 0.05). (F) Kinase substrate enrichment using the phosphorylation sites uniquely identified with MALDI reveals that four of the top ten kinases identified are acidic directed. The pie chart depicts that the associated, acidic kinase substrates constitute almost two thirds of all identified substrates.

LC- nESI data set (Figure 2B) clearly indicating that the detection of these sites was strongly favored by MALDI. Using more stringent phosphorylation site localization criteria to $< 1\%$ FLR, by increasing the required Mascot Delta Score to > 15 for both MALDI and nESI (Figure 2C) (33, 37), did not change this observation excluding the possibility that the apparent orthogonality of LC-MALDI and LC-nESI may have resulted from false phosphorylation site assignments.

For the phosphorylation sites detected exclusively by LC-MALDI, it was found that the frequency of acidic residues in the respective tryptic peptide was twice as high as in those detected by LC-nESI (increase from 5% to 10% for aspartate and from 7% to 14% for glutamate; Figure 2D). Compared to a whole proteome background, these acidic residues are significantly overrepresented in positions surrounding the phosphorylation site which indicates the presence of kinase motifs (analyzed using pLOGO (34); Figure 2E). Indeed, kinase substrate enrichment analysis using KinomeXplorer (35) revealed that the majority of the substrates are likely phosphorylated by acidic residue-directed kinases (Figure 2F). Consistent with previous work (23, 24), I did not detect significant differences in phosphopeptide length and hydrophobicity indicating that neither chromatographic factors nor the efficiency of peptide fragmentation/detection by MS can explain the apparent orthogonality of LC-MALDI and LC-nESI. However, I noticed that the frequency of tyrosine and phenylalanine residues was considerably higher in the LC-MALDI data compared to the LC-nESI data (4 % vs. 2 % for Y and 3 % vs. 2 % for F). This is consistent with very early work on the MALDI process (19) and possibly indicating that the presence of aromatic amino acids aids in the desorption and ionization of these phosphopeptides. Conversely, proline residues appeared substantially underrepresented in LC-MALDI (8 % vs. 13 % in LC-nESI), an observation for which there is no plausible explanation at present.

In contrast to observation made in earlier studies, which mainly analyzed samples of low molecular complexity, the LC-MALDI data did neither indicate an increase in basic amino acid frequency in general (24), nor a preference towards the occurrence of C-terminal arginine in particular (23, 24, 38). I did also not observe a preferred detection of multiply phosphorylated or large phosphopeptides which has previously been reported (26). Given that the samples and measurement details used in this study were not entirely comparable to those employed in the cited earlier work, it should be emphasized that the above discrepancies may have multiple causes not investigated further here.

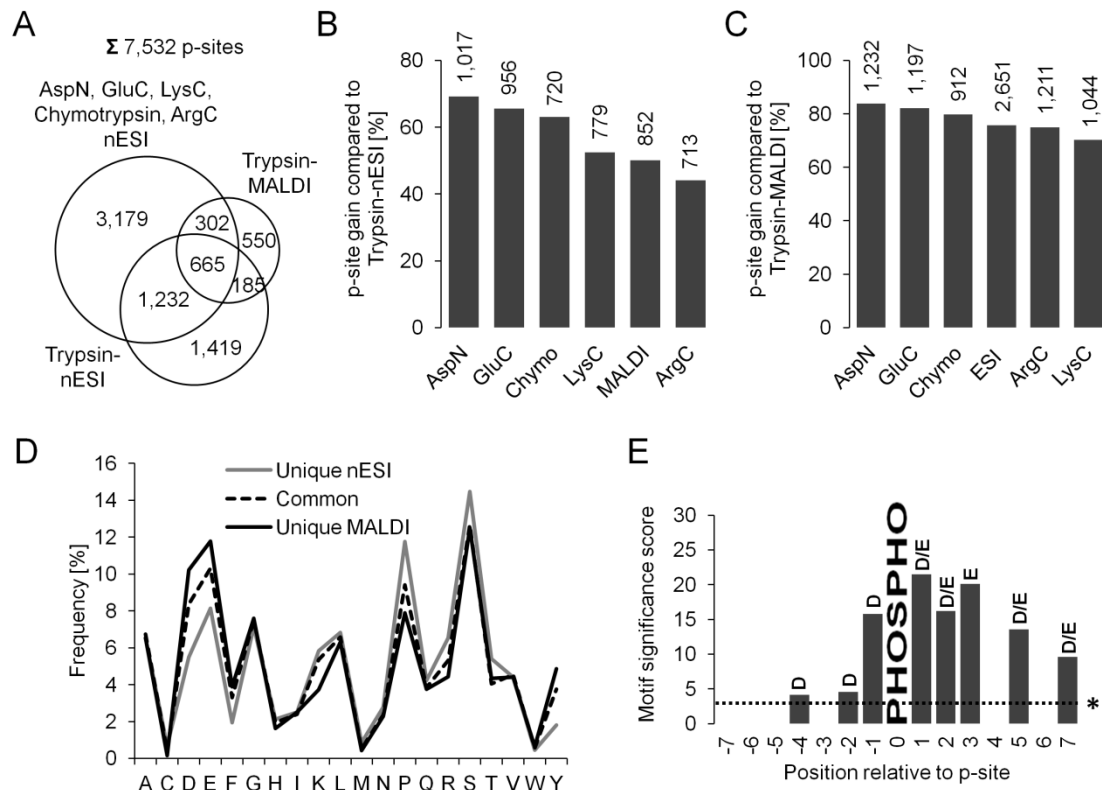


Figure 3. Comparison of MALDI against the phosphorylation site orthogonality of different digestion enzymes. (A) Venn diagram depicting the overlap of unique phosphorylation sites (FLR < 5 %) identified by single measurements with different digestion enzymes (Asp-N, Arg-C, chymotrypsin, Glu-C and Lys-C), trypsin-nESI and trypsin-MALDI. Trypsin-MALDI identifies 550 unique phosphorylation sites compared to the 6,982 phosphorylation sites identified with nESI. (B) The phosphorylation site gain (FLR < 5 %) of different digestion enzymes (Asp-N, Arg-C, chymotrypsin, Glu-C and Lys-C) compared to a standard trypsin-nESI workflow (in percent of the respective dataset) clearly shows that orthogonal proteases unlock different parts of the phosphoproteome. Absolute numbers of uniquely identified phosphorylation sites are shown above bars. (C) The orthogonality of trypsin and Asp-N, Arg-C, chymotrypsin, Glu-C and Lys-C recorded with nESI ionization is higher if trypsin digested peptides are ionized by MALDI compared to nESI. This indicates that the combination of these approaches increases phosphoproteome coverage even further. Absolute numbers of uniquely identified phosphorylation sites are shown above bars. (D) Global amino acid frequency distribution of combined phosphopeptides digested with different proteases and identified by nESI, identified by MALDI and phosphopeptides identified with both ionization methods reveals an overrepresentation of acidic residues in the MALDI dataset. (E) Overrepresented aspartate and glutamate residues in the MALDI dataset are located at statistically significant positions surrounding the phosphorylation site. Motif significance scores were calculated using pLOGO; the dotted line marked with an asterisk indicates statistical significance (p -value < 0.05).

Phosphorylation site identification orthogonality of LC-MALDI and LC-nESI is comparable to using different proteases in conjunction with LC-nESI analysis

In order to further substantiate that LC-MALDI-MS/MS detects a distinct set of phosphorylation sites compared to LC-nESI, I compared phosphorylation site identification results obtained by using different digestion enzymes. For this purpose, one milligram of A431 lysate each was

digested with Asp-N, Arg-C, chymotrypsin, Glu-C and Lys-C, respectively, followed by phosphopeptide enrichment using Fe-IMAC chromatography. Phosphopeptides were successfully and selectively (> 95 %) purified for all employed proteases (Table 1) and each phosphopeptide pool was analyzed by LC-nESI-MS/MS using a four hour LC gradient on an Orbitrap Velos. Cumulatively, 5,378 phosphorylation sites (FLR < 5 %) were identified across the alternative digestion enzymes of which only 35 % overlapped with trypsin LC-nESI (Figure 3A) which is similar to what has been recently reported by Swaney et al. (14) based on whole proteome digests and Giansanti et al. (39) who analyzed phosphopeptide enrichments. The number of acquired tandem mass spectra was comparable to that obtained for trypsin indicating that overall MS performance was comparable (Table 1). However, the identification rate of tandem mass spectra was considerably lower possibly as a result of e.g. more complex peptide fragment spectra of non-tryptic peptides and the fact that LC and MS parameters were not specifically optimized for non-tryptic peptides in this study.

Table 1. Summary of phosphopeptide identifications from different digests and a nESI and MALDI measurement using trypsin.

	Acquired Spectra	PSM ID rate [%]	Unique phosphopeptides	p-sites (FLR < 5 %)	Selectivity [%]	Multiply phosphorylated [%]
AspN-nESI	30,450	20	2,471	1,470	97	8
LysC-nESI	26,358	29	2,166	1,484	96	8
Chymotrypsin-nESI	22,793	26	1,607	1,142	98	6
GluC-nESI	28,876	19	2,301	1,456	98	8
ArgC-nESI	24,984	29	2,143	1,615	99	11
Trypsin-nESI	31,212	58	5,509	3,501	91	8
Trypsin-MALDI	7,171	32	2,132	1,702	95	3

The gain in phosphorylation site identification to that of using the single measurement trypsin dataset alone ranged from 44 % for Arg-C (713 sites) to 69 % for Asp-N (1,017 sites; Figure 3B). Interestingly, the gain in phosphorylation site identification provided by the LC-MALDI dataset (trypsin) was comparable, both in relative and absolute terms, to that of the use of alternative proteases and LC-nESI analysis again indicating that LC-MALDI-MS/MS provides access to phosphorylation sites not covered by LC-nESI-MS/MS. This absolute and relative gain was even more pronounced when phosphorylation sites obtained with different digestion enzymes and nESI ionization are compared to trypsin-MALDI sites, ranging from 912 (80 %) additional sites for chymotrypsin to 1,232 (84 %) for AspN (Figure 3C). When combining the data from all protease digests, about one third (550) of all phosphorylation sites detected by LC-

MALDI-MS/MS were not found in any other experiment (Figure 3A). I next analyzed the alternative protease digests by LC-MALDI-MS/MS and found the same trends observed for trypsin, notably the overrepresentation of D, E, Y and F residues, the underrepresentation of P residues and the strong preference for acidic amino acids surrounding the phosphorylation site (Figures 3D, E)

MALDI-MS/MS preferentially desorps/ionizes/detects phosphotyrosine containing peptides

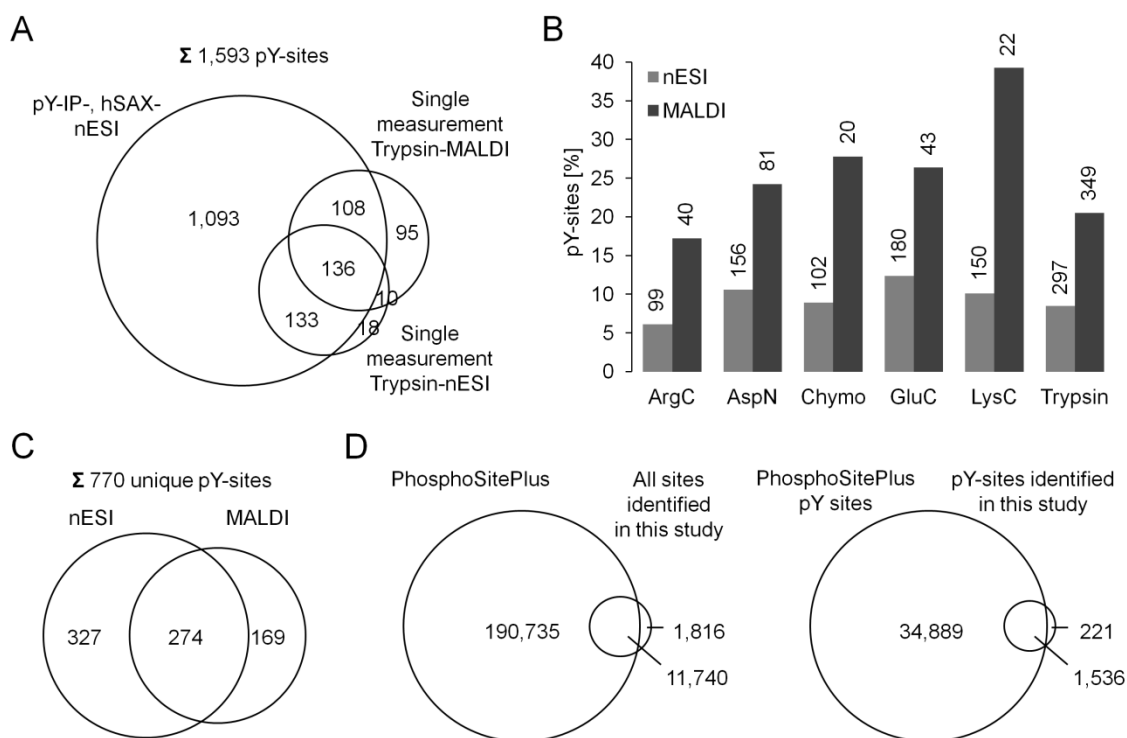


Figure 4. MALDI preferentially ionizes phosphotyrosine containing peptides. (A) Venn diagram depicting the overlap of phosphotyrosine sites (FLR < 5 %) identified in trypsin-MALDI single measurement (1 mg A431 digest), trypsin-nESI single measurement (1 mg A431 digest) and by the combination of hSAX (24 fractions, 5 mg A431 digest) and phosphotyrosine-immunoprecipitation (pY-1000, 5 mg digest). Whereas nearly all sites of trypsin-nESI single measurement are also covered by the 1,470 sites identified in the deep phosphotyrosine-proteome, 95 sites (27 % of the dataset) are only identified in the trypsin-MALDI single measurement. This indicates preferential ionization of certain phosphotyrosine sites by MALDI. (B) Comparison of the percentage (absolute numbers are shown above bars) of phosphotyrosine sites in relation to all sites identified using the indicated protease in combination with nESI or MALDI (FLR < 5 %). It is apparent that MALDI has a preference for the ionization of phosphotyrosine containing peptides. (C) Overlap of phosphotyrosine sites identified using different protease digests (Asp-N, Arg-C, chymotrypsin, Glu-C, Lys-C and trypsin), followed by nESI and MALDI. (D) Comparison of phosphorylation sites identified in this study against those recorded in the PhosphoSitePlus database. 1,816 out of 13,556 confidently localized phosphorylation sites are not yet recorded in PhosphoSitePlus. This clearly indicates non complete phosphoproteome coverage. The present work increases the number of phosphotyrosine sites recorded by 221.

In addition to the phosphorylation site identification complementarity between LC-MALDI and LC-nESI tandem mass spectrometry observed above, I made another unexpected observation. Tyrosine phosphorylation (pY) accounted for a surprisingly high absolute (349) and relative (20.5 %) incidence of all identified phosphorylation sites in LC-MALDI-MS/MS data (trypsin). The respective figures for LC-nESI-MS/MS were 297 sites (8.5 %; Figure 4A). Here again, 55 % of the phosphotyrosine sites found by single LC-MS/MS experiments were exclusively identified by LC-MALDI-MS/MS. In order to ascertain that this apparent complementarity can be attributed to the employed ionization technique rather than undersampling by the mass spectrometer, pY immunoprecipitation was used to enrich phosphotyrosine containing peptides from a tryptic digest representing five mg protein from pervanadate treated A431 cells. 1,900 phosphotyrosine containing peptides corresponding to 1,219 unique phosphotyrosine sites were identified (FLR < 5 %). I combined this data with the hSAX fractionation data presented above (containing 814 phosphotyrosine sites, FLR < 5 %) in order to compensate for biases potentially introduced by either method (40). Figure 4A shows that despite extending the number of pY-sites by more than five fold, 27% of the pY-sites (95) identified by LC-MALDI-MS/MS (trypsin) are not found in any other data set. In contrast and in line with expectations, nearly all of the sites previously identified in single LC-nESI-MS/MS measurements were also covered in the combined hSAX and phosphotyrosine immunoprecipitation dataset (Figure 4A). The preference for the detection of pY-sites by LC-MALDI-MS/MS was confirmed when analyzing the LC-MALDI and LC-nESI-MS/MS data of the alternative digestion enzymes. Even though the absolute numbers are quite low, the percentage of pY-sites was on average three fold higher in LC-MALDI than LC-nESI-MS/MS data for all proteases investigated (Figure 4B). When confining the comparison of pY-sites identified from the same sample but using different ionization methods the results are consistent in that 38 % of all pY-sites identified by LC-MALDI-MS/MS are not found in the LC-nESI data (Figure 4C). It has previously been observed that the aromatic amino acids phenylalanine and tyrosine can enhance the ionization of peptides and the data supports this notion (see also above) (41). In light of the data presented above, it appears tempting to speculate that the addition of a phosphate group to tyrosine possibly also promotes the desorption/ionization process, which is supported by the fact that the removal of a phosphor group from tyrosine residues alters the ultraviolet absorbance of corresponding peptides (42). However, this will remain a matter of further investigations in the future.

Complementary approaches are valuable for the increase of phospho(tyrosine)site coverage

Over the course this study, evidence for a total of 13,556 confidently localized phosphorylation sites of which 1,757 were found on tyrosine residues has been collected (Figure 4D). In light of the observed complementarity for LC-MALDI and LC-nESI-MS/MS, it was to be found out if any of the sites observed here are potentially novel by comparing the data to the PhosphoSitePlus database (43). Owing to the massive increase in available data, this resource has grown from ~100,000 phosphorylation sites in March 2014 to ~200,000 sites in March 2015. At the time of writing, 1,816 (13 %) of the 13,556 phosphorylation sites identified in this work were not yet recorded in PhosphoSitePlus (Figure 4D). Despite the fact that phosphotyrosine has been extensively studied and is of proportionally much lower occurrence than serine or threonine phosphorylation, 13 % (221) of the 1,757 identified phosphotyrosine sites found in this study have not yet been observed or recorded in PhosphoSitePlus (Figure 4D). Some such examples include i) phosphorylation at serine 835 on STAT2, a transcription factor downstream of several receptor associated kinases, ii) phosphorylation at serine 83 on PTPN12, a phosphatase dephosphorylating several tyrosine kinases and iii) phosphorylation at tyrosine 131 on the human tyrosine kinase SYK, which has previously only been detected in mice and rats where it regulates kinase activity (44). I do not mean to over emphasize the discovery of novel sites in this particular study but find it noteworthy because the data clearly indicates that there is still a need to develop or employ novel techniques or combinations thereof for comprehensive phosphorylation site mapping of proteomes.

Conclusion

In this study, LC-MALDI and LC-nESI-MS/MS was used as a means to extend phosphorylation site coverage and it was found that either approach exhibits distinct preferences with respect to the detected phosphopeptides. Importantly, LC-MALDI-MS/MS showed a bias towards the identification of acidic phosphopeptides as well phosphotyrosine sites. The present work clearly shows that, despite a decade of large-scale efforts, the human phosphoproteome is still incompletely mapped. The complementarity of LC-MALDI and LC-nESI-MS/MS highlighted in this study may thus have a role to play in closing the gaps. There are indeed numerous examples in the literature, where the application of MALDI tandem MS was required to identify phosphorylation sites of biological significance (45). The combination of phosphotyrosine immuno-affinity enrichment and MALDI-MS/MS may be particularly useful in the future given the outstanding importance of tyrosine phosphorylation in physiological and pathological processes which may thus lead to a renaissance of the approach.

Acknowledgement

The author wishes to thank Dr. Christoph Roesli for conducting the LC-MALDI-MS/MS measurements and Heiner Koch, Tobias Baumann and Max Mundt for excellent experimental assistance.

Abbreviations

FLR	False localization rate
hSAX	Hydrophilic strong anion exchange
I.D.	Inner diameter
IMAC	Immobilized metal ion affinity chromatography
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS/MS	Tandem mass spectrometry
nESI	(Nano-)Electrospray ionization
pY	Phosphotyrosine
TiO ₂	Titaniumdioxide
TOF	Time-of-flight

References

1. Ruprecht, B., and Lemeer, S. (2014) Proteomic analysis of phosphorylation in cancer. *Expert Rev. Proteomics* 11, 259–267
2. Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J.-H., Bantscheff, M., Gerstmair, A., Faerber, F., and Kuster, B. (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509, 582–587
3. Sharma, K., D'Souza, R. C. J., Tyanova, S., Schaab, C., Wiśniewski, J. R., Cox, J., and Mann, M. (2014) Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.* 8, 1583–1594
4. Olsen, J. V., and Mann, M. (2013) Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry. *Mol. Cell. Proteomics* 12, 3444–3452
5. Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. R. (2004) Selective Isolation at the Femtomole Level of Phosphopeptides from Proteolytic Digests Using 2D-NanoLC-ESI-MS/MS and Titanium Oxide Precolumns. *Anal. Chem.* 76, 3935–3943
6. Zhou, H., Low, T. Y., Hennrich, M. L., Toorn, H. van der, Schwend, T., Zou, H., Mohammed, S., and Heck, A. J. R. (2011) Enhancing the Identification of Phosphopeptides from Putative Basophilic Kinase Substrates Using Ti (IV) Based IMAC Enrichment. *Mol. Cell. Proteomics* 10, M110.006452
7. Andersson, L., and Porath, J. (1986) Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal. Biochem.* 154, 250–254
8. Ruprecht, B., Koch, H., Medard, G., Mundt, M., Kuster, B., and Lemeer, S. (2015) Comprehensive and Reproducible Phosphopeptide Enrichment Using Iron Immobilized Metal Ion Affinity Chromatography (Fe-IMAC) Columns. *Mol. Cell. Proteomics* 14, 205–215
9. Steen, H., Kuster, B., Fernandez, M., Pandey, A., and Mann, M. (2002) Tyrosine Phosphorylation Mapping of the Epidermal Growth Factor Receptor Signaling Pathway. *J. Biol. Chem.* 277, 1031–1039
10. Blagoev, B., Ong, S.-E., Kratchmarova, I., and Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* 22, 1139–1145
11. Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J., Rush, J., Lauffenburger, D. A., and White, F. M. (2005) Time-resolved Mass Spectrometry of Tyrosine Phosphorylation Sites in the Epidermal Growth Factor Receptor Signaling Network Reveals Dynamic Modules. *Mol. Cell. Proteomics* 4, 1240–1250
12. Giansanti, P., Stokes, M. P., Silva, J. C., Scholten, A., and Heck, A. J. R. (2013) Interrogating cAMP-dependent Kinase Signaling in Jurkat T Cells via a Protein Kinase A Targeted Immune-precipitation Phosphoproteomics Approach. *Mol. Cell. Proteomics MCP* 12, 3350–3359
13. Choudhary, G., Wu, S.-L., Shieh, P., and Hancock, W. S. (2003) Multiple Enzymatic Digestion for Enhanced Sequence Coverage of Proteins in Complex Proteomic Mixtures Using Capillary LC with Ion Trap MS/MS. *J. Proteome Res.* 2, 59–67
14. Swaney, D. L., Wenger, C. D., and Coon, J. J. (2010) Value of Using Multiple Proteases for Large-Scale Mass Spectrometry-Based Proteomics. *J. Proteome Res.* 9, 1323–1329
15. Wiśniewski, J. R., and Mann, M. (2012) Consecutive Proteolytic Digestion in an Enzyme Reactor Increases Depth of Proteomic and Phosphoproteomic Analysis. *Anal. Chem.* 84, 2631–2637
16. Gilmore, J. M., Kettenbach, A. N., and Gerber, S. A. (2012) Increasing phosphoproteomic coverage through sequential digestion by complementary proteases. *Anal. Bioanal. Chem.* 402, 711–720
17. Yamashita, M., and Fenn, J. B. (1984) Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* 88, 4451–4459
18. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64–71
19. Karas, M., Bachmann, D., and Hillenkamp, F. (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem.* 57, 2935–2939
20. Karas, M., Bachmann, D., Bahr, U., and Hillenkamp, F. (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int. J. Mass Spectrom. Ion Process.* 78, 53–68

21. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 2, 151–153
22. Bodnar, W. M., Blackburn, R. K., Krise, J. M., and Moseley, M. A. (2003) Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J. Am. Soc. Mass Spectrom.* 14, 971–979
23. Stapels, M. D., and Barofsky, D. F. (2004) Complementary Use of MALDI and ESI for the HPLC-MS/MS Analysis of DNA-Binding Proteins. *Anal. Chem.* 76, 5423–5430
24. Hessling, B., Büttner, K., Hecker, M., and Becher, D. (2013) Global relative quantification with liquid chromatography-matrix-assisted laser desorption ionization time-of-flight (LC-MALDI-TOF)--cross-validation with LTQ-Orbitrap proves reliability and reveals complementary ionization preferences. *Mol. Cell. Proteomics MCP* 12, 2911–2920
25. Parker, L., Engel-Hall, A., Drew, K., Steinhardt, G., Helseth, D. L., Jabon, D., McMurry, T., Angulo, D. S., and Kron, S. J. (2008) Investigating quantitation of phosphorylation using MALDI-TOF mass spectrometry. *J. Mass Spectrom. JMS* 43, 518–527
26. Aryal, U. K., and Ross, A. R. S. (2010) Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. *Rapid Commun. Mass Spectrom.* 24, 219–231
27. Ritorto, M. S., Cook, K., Tyagi, K., Pedrioli, P. G. A., and Trost, M. (2013) Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *J. Proteome Res.* 12, 2449–2457
28. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896–1906
29. Hahne, H., Pachi, F., Ruprecht, B., Maier, S. K., Klaeger, S., Helm, D., Médard, G., Wilm, M., Lemeer, S., and Kuster, B. (2013) DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat. Methods* 10, 989–991
30. Seidler, J., Zinn, N., Haaf, E., Boehm, M. E., Winter, D., Schlosser, A., and Lehmann, W. D. (2011) Metal ion-mobilizing additives for comprehensive detection of femtomole amounts of phosphopeptides by reversed phase LC-MS. *Amino Acids* 41, 311–320
31. van den Toorn, H. W. P., Muñoz, J., Mohammed, S., Rajmakers, R., Heck, A. J. R., and van Breukelen, B. (2011) RockerBox: analysis and filtering of massive proteomics search results. *J. Proteome Res.* 10, 1420–1424
32. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* 4, 923–925
33. Marx, H., Lemeer, S., Schliep, J. E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A. J. R., and Kuster, B. (2013) A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nat. Biotechnol.* 31, 557–564
34. O’Shea, J. P., Chou, M. F., Quader, S. A., Ryan, J. K., Church, G. M., and Schwartz, D. (2013) pLogo: a probabilistic approach to visualizing sequence motifs. *Nat. Methods* 10, 1211–1212
35. Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., Palma, A., Cesareni, G., Jensen, L. J., and Linding, R. (2014) KinomeXplorer: an integrated platform for kinome biology studies. *Nat. Methods* 11, 603–604
36. Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., Dienes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P.-A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H.-J., Albar, J. P., Martinez-Bartolomé, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* 32, 223–226
37. Lemeer, S., Kunold, E., Klaeger, S., Raabe, M., Towers, M. W., Claudes, E., Arrey, T. N., Strupat, K., Urlaub, H., and Kuster, B. (2012) Phosphorylation site localization in peptides by MALDI MS/MS and the Mascot Delta Score. *Anal. Bioanal. Chem.* 402, 249–260
38. Krause, E., Wenschuh, H., and Jungblut, P. R. (1999) The Dominance of Arginine-Containing Peptides in MALDI-Derived Tryptic Mass Fingerprints of Proteins. *Anal. Chem.* 71, 4160–4165
39. Giansanti, P., Aye, T. T., van den Toorn, H., Peng, M., van Breukelen, B., and Heck, A. J. R. (2015) An Augmented Multiple-Protease-Based Human Phosphopeptide Atlas. *Cell Rep.* 11, 1834–1843
40. Di Palma, S., Zoumaro-Djajoon, A., Peng, M., Post, H., Preisinger, C., Munoz, J., and Heck, A. J. R. (2013) Finding the same needles in the haystack? A comparison of phosphotyrosine peptides

- enriched by immuno-affinity precipitation and metal-based affinity chromatography. *J. Proteomics* 91, 331–337
41. Nishikaze, T., and Takayama, M. (2006) Cooperative effect of factors governing molecular ion yields in desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom. RCM* 20, 376–382
 42. Zhang, Z. Y., Maclean, D., Thiemeseffler, A. M., Roeske, R. W., and Dixon, J. E. (1993) A Continuous Spectrophotometric and Fluorometric Assay for Protein Tyrosine Phosphatase Using Phosphotyrosine-Containing Peptides. *Anal. Biochem.* 211, 7–15
 43. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270
 44. Castro, R. O. de, Zhang, J., Jamur, M. C., Oliver, C., and Siraganian, R. P. (2010) Tyrosines in the Carboxyl Terminus Regulate Syk Kinase Activity and Function. *J. Biol. Chem.* 285, 26674–26684
 45. Fernández-Sáiz, V., Targosz, B.-S., Lemeer, S., Eichner, R., Langer, C., Bullinger, L., Reiter, C., Slotta-Huspenina, J., Schroeder, S., Knorn, A.-M., Kurutz, J., Peschel, C., Pagano, M., Kuster, B., and Bassermann, F. (2013) SCFFbxo9 and CK2 direct the cellular response to growth factor withdrawal via Tel2/Tti1 degradation and promote survival in multiple myeloma. *Nat. Cell Biol.* 15, 72–81

Chapter 4

Evaluation of kinase activity profiling using chemical proteomics

Abstract

Protein kinases are important mediators of intracellular signaling and are reversibly activated by phosphorylation. Immobilized kinase inhibitors can be used to enrich these often low abundance proteins, to identify targets of kinase inhibitors or to probe their selectivity. It has been suggested that binding of kinases to affinity beads reflects a kinase's activation status, a concept that is under considerable debate. To assess the merits of the idea, a series of experiments including quantitative phosphoproteomics and purification of kinases by single or mixed affinity matrices from signaling activated or resting cancer cells was performed. The data show that mixed affinity beads generally bind kinases independent of their activation status and experiments using individual immobilized kinase inhibitors show mixed results in terms of preference for binding of the active or inactive conformation. Taken together, activity or conformation dependent binding to such affinity resins depends i) on the kinase, ii) on the affinity probe and iii) on the activation status of the lysate or cell. As a result, great caution should be exercised when inferring kinase activity from such binding data. The results also suggest that assaying kinase activity using binding data is restricted to a limited number of well chosen cases.

Introduction

Protein kinases are important mediators of intracellular signaling and exert their function via the posttranslational addition of a phosphate group to serine, threonine or tyrosine residues. In concert with phosphatases, the stoichiometry of phosphorylation at a given site is reversibly regulated in response to a wide variety of factors (1). Aberrant activation of kinases impairs this tightly controlled balance and can lead to diseases such as cancer, making kinases prominent targets for the development of small molecule inhibitors (2). Since kinases share a structurally conserved ATP binding pocket, achieving drug selectivity remains a challenge (3, 4). Amongst different biochemical tools (5, 6), chemoproteomic methods (7, 8, 9, 10, 11) including Kinobeads (12) can be used to assess compounds for their selectivity. Such broad spectrum kinase inhibitors immobilized on beads enable the purification of native kinases directly from cell lysates. When configured as a competition binding assay, an inhibitor of interest can be dosed into lysates or into cells and thus compete with binding to the affinity matrix. In conjunction with quantitative mass spectrometry, Kinobeads enable the dose-dependent determination of inhibitor selectivity against hundreds of endogenous kinases and ATP binding proteins.(12, 13) Similar kinome affinity enrichments have been applied to the analysis of kinase expression across cell lines and tissues (14, 15), the elucidation of resistance mechanisms (16), the discovery of kinome reprogramming (17) and the global determination of nucleotide cofactor affinities across the kinome (18).

How much of a given kinase is captured on Kinobeads or similar multiplexed inhibitor beads (MIBs (17)) mainly depends on the level of kinase expression in the cell, the affinity of the kinase to the immobilized compound and, at least for some kinases, on their conformation/activity (12, 19). Unfortunately, determining the latter is not straightforward. Several recent articles from the same laboratory have claimed that MIBs are capable to “quantitatively measure activation changes in the majority of kinases”, can detect “altered kinome activity profiles in response to stimuli or kinase inhibitors” and that “global changes in kinase activation can be determined after drug treatment” (20, 21). These claims stem from experiments utilizing MIBs after global activation of signaling pathways by phosphatase inhibition using pervanadate (17). Most of the affinity matrices used in these studies are immobilized type I inhibitors, which are generally thought to enrich kinases both in the “DFG-in” and the “DFG-out” conformation (22). Surprisingly, the aforementioned papers conclude that kinase activation increases the affinity of tyrosine kinases to the immobilized inhibitors on a global scale. In a very recent study, MIB columns are even accredited to “capture active

kinases”, further extending and generalizing the overall notion (23). Despite being a very appealing concept, our laboratory and others have doubts that the idea can be generalized to such an extent.

Therefore, in this study, the influence of kinase activity on the binding to immobilized inhibitor mixtures was revisited and the experimental results arrive at more refined and somewhat opposing conclusions. Briefly, it was found that kinase binding to immobilized kinase inhibitors is largely independent of kinase activity. Based on the quantitative comparison of kinase affinity binding to cellular activity changes, as determined by quantitative phosphoproteomics, it can be concluded that the influence of activity on kinase-bead binding has to be assessed on a case by case basis as the composition of the bead mixture itself as well as the activation state of the cell also strongly influences the obtained results. Therefore, chemoproteomic experiments are of limited use as a general read-out for kinase activity.

Materials and methods

Cell culture

The human neuroblastoma cell line SK-N-BE(2) was grown in DMEM/Ham's F-12 medium (Biochrom AG) supplemented with 10 % (v/v) Fetal Bovine Serum (Biochrom AG) and 1 % antibiotic/antimycotic solution (Sigma Aldrich). BT-474 breast cancer cells were grown in DMEM/Ham's F-12 medium supplemented with 15 % (v/v) Fetal Bovine Serum and 1 % antibiotic/antimycotic solution. Prior to lysis, BT-474 cells were treated with increasing concentrations (1 nM to 10 μ M) of lapatinib (LC Laboratories) for 1 h and SK-N-BE(2) cells were treated with 100 μ M pervanadate for 15 min. Pervanadate was always prepared freshly and activated by mixing equimolar amounts of H₂O₂ and Na₃VO₄. Immediately before harvesting of cells at roughly 90 % confluency, they were washed twice with ice cold PBS. Cells were lysed in 1 x CP buffer (50 mM Tris-HCl, pH 7.5, 5 % glycerol, 1.5 mM MgCl₂, 150 mM NaCl) supplemented with 0.8 % NP-40, 1 mM DTT, 25 mM NaF and freshly added protease and phosphatase inhibitors (5 x phosphatase inhibitor cocktail 1, Sigma-Aldrich; 5 x phosphatase inhibitor cocktail 2, Sigma-Aldrich; 1 mM Na₃VO₄ and 20 nM calyculin A). Protein extracts were clarified by ultracentrifugation for 1 h at 150.000 g and 4 °C. Protein concentration was determined by the Bradford method (Coomassie (Bradford) Protein Assay Kit, Thermo Scientific) and lysates were stored at -80 °C until further use.

Compound synthesis and coupling

All compounds included in KBy were synthesized according to the procedures described by Médard *et al.* (13). Linkable Dasatinib was synthesized according to the procedures described in the patent WO2013055780A1. SB203580 was synthesized in 4 steps according to Munoz modification (24) of Biftu procedure (25). VI16743 was generously provided by Dr. Josef Wissing from the Helmholtz Centre for Infection Research, Braunschweig, Germany. Bisindolylmaleimide X (BisX) was commercially sourced. Compounds were immobilized on sepharose beads through covalent linkage using their primary amino groups (12). NHS-activated sepharose (GE Healthcare, Freiburg, Germany) and the compounds (1, 2, and 4 μ mol per ml beads) were equilibrated in DMSO. 15 μ l of triethylamine was added to start the coupling reaction and the mixture was incubated on an end-over-end shaker for 16 - 20 h in the dark. Free NHS-groups on beads were blocked by adding 50 μ l amino ethanol and incubation on an

end-over-end shaker for 16 - 20 h in the dark. Coupled beads were washed and stored in ethanol at 4 °C in the dark. The coupling reaction was monitored by LC-MS.

Western blot

Prior to SDS-PAGE, samples were heated in 2 x LDS buffer containing 50 mM DTT for 10 min at 90 °C. 30 µg protein per lysate was loaded onto NuPAGE 4 - 12 % Bis/Tris gels, which were run at 200 V for 50 min in 1 x MOPS buffer. After gel electrophoresis, a wet chamber transfer to PVDF membranes (0.45 µm, BioRad) was performed using the XCell II™ Blot Module (Invitrogen) according to the manufactures instructions. The primary antibodies for AKT, pS473-AKT, MAPK1/3 and pT202/pY204-MAPK1/3 were purchased from Cell Signaling Technology. The anti-phosphotyrosine antibody pY-99 was obtained from Santa Cruz Biotechnology.

Protein digestion and dimethyl labeling

For the phosphoproteomics experiments, urea was added to the cleared lysates to a final concentration of 8 M. Afterwards protein extracts were reduced with 10 mM DTT at 56 °C for 1 h and alkylated with 25 mM iodoacetamide for 30 min at room temperature in the dark. The protein mixture was diluted with 40 mM Tris/HCl to a final urea concentration of 1.6 M. Protein digestion was performed by adding sequencing grade trypsin (Promega, 1 : 100 enzyme to substrate ratio) and incubation at 37 °C for 4 h. Subsequently, another portion of trypsin (1 : 100) was added and digestion was continued overnight. The following day, samples were acidified to a pH of 2 by addition of TFA and desalted using SepPack columns (C18 cartridges Sep-Pak Vac 1cc (200 mg), Waters Corp., desalting solvent A: 0.07 % TFA; desalting solvent B: 0.07 % TFA, 50 % ACN). Dimethyl labelling was performed on column as described (26). Briefly, after peptides were bound to the C18 material and washed with 8 ml of desalting solvent A, 5 ml light (10 % of 600 mM NaBH₃CN, 10 % of 4 % formaldehyde, 80 % of 50 mM sodium phosphate; all percentages as v/v) or intermediate (10 % of 600 mM NaBH₃CN, 10 % of 4 % D-formaldehyde, 80 % of 5 mM sodium phosphate; all percentages as v/v) labelling solution was slowly passed through the cartridges. Subsequently, 8 ml of desalting solvent A was used to wash out remaining labelling reagents. Labelled peptides were eluted using 1 ml desalting solvent B, dried *in vacuo* and stored at -80 °C.

Fe-IMAC enrichment and high-pH reversed-phase C18 tip fractionation

Phosphopeptide enrichment using a Fe-IMAC column (ProPac IMAC-10 column, 4 mm I.D. x 50 mm, Thermo Fisher Scientific) connected to an Aekta HPLC system was essentially performed as described previously (27). Briefly, the column was charged with FeCl_3 and equilibrated with IMAC solvent A (30 % ACN, 0.07 % TFA). Dried peptides were reconstituted in IMAC solvent A, loaded onto the column (10 min, 0.1 ml/min) and washed (16 min, 0.3 ml/min). Subsequently, phosphopeptides were eluted by a linear gradient from 0 – 45 % solvent B (0.5 % NH_4OH , 0.2 ml/min) over 60 min. The phosphopeptide elution peak was collected, dried down and stored at $-80\text{ }^\circ\text{C}$.

High-ph reversed-phase tip fractionation was performed in 200 μl pipette tips packed with three C18 extraction disks (\O 1.5 mm, 3M Empore) and fixed in 1.5 ml eppendorf tubes. All solvents were passed through the tip by centrifugation (2000 rpm, room temperature). First, tips were primed using 50 μl of 100 % ACN, followed by 50 μl of 60 % ACN in 25 mM NH_4COOH and 2 x 50 μl of 25 mM NH_4COOH , pH 10. Next, dried phosphopeptides were re-dissolved in 100 μl of 25 mM NH_4COOH , pH 10 and loaded onto the C18 material. After re-application of the flow through, phosphopeptides were eluted using increasing concentrations of ACN (2.5 %, 7.5 %, 12.5 %, 60 % ACN in 25 mM NH_4COOH). The previously stored flow through was combined with the 60 % ACN eluate, resulting in a total of four fractions, which were dried down and stored at $-20\text{ }^\circ\text{C}$ until LC-MS/MS measurement. The procedure for full proteome fractionation was identical except for the use of elution solvents containing 2.5 %, 5 %, 7.5 %, 12.5 %, 60 % ACN in 25 mM NH_4COOH . Here, the flow through fraction was combined with fraction 17.5 % ACN and the 5 % ACN fraction was combined with 50 % ACN fraction, resulting in a total of six fractions.

Kinase affinity pull downs

Kinobead pull downs were conducted in a 96-well plate format as described (13) using 2 mg protein and 35 μl of settled KBy, the MIB-like mix of immobilized compounds or single immobilized compounds. Chemoproteomic enrichments were conducted in technical triplicates. Beads were washed twice with 1 ml 1 x CP and then equilibrated with 1 ml 1 x CP/0.4 % NP-40. Residual liquid was removed via centrifugation (1200 rpm, 2 min, $4\text{ }^\circ\text{C}$). 1 ml lysate diluted to a concentration of 2 mg/ml protein and 0.4 % NP-40 was then added to the beads (for drug competition assays lysates were pre-incubated with the specified concentration of drug or an equivalent volume of DMSO for 45 min). After incubation for 30 min ($4\text{ }^\circ\text{C}$, on end-over-end

shaker), beads were washed three times with 1 ml of 1x CP/0.4% NP-40 and twice with 1 ml 1 x CP/0.2 % NP-40. Wash solvents were removed via centrifugation (1200 rpm, 2 min, 4 °C). For elution, samples were incubated with 60 µl of 2 x NuPAGE LDS Sample Buffer containing 50 mM DTT in a thermoshaker (30 min, 50 °C, 700 rpm). The eluate was harvested via centrifugation (1200 rpm, 2 min, 4 °C) and 30 µl thereof was alkylated with 55 mM chloroacetamide (CAA) for 30 min at room temperature in the dark. Prior to MS analysis, samples were desalted and concentrated by SDS-PAGE and digested as previously described (28).

LC-MS/MS analysis

Table 1. Detailed LC-MS/MS parameters for each type of experiment.

	Quantitative Phosphoproteomics	Quantitative Phosphoproteomics	Chemoproteomics	Chemoproteomics	Full Proteome
	Orbitrap Velos	Q-Exactive	Orbitrap Elite	Velos	Q-Exactive Plus
source voltage	2.2 kV	1.7 kV	2.2 kV	2.2 kV	2.2 kV
MS1 AGC	1E06	3E06	1E06	1E06	3E06
MS1 max IT	100 ms	250 ms	100 ms	100 ms	100 ms
MS2 AGC	4E04	5E04	2E04	4E04	1E05
MS2 max IT	250 ms	120 ms	100 ms	100 ms	50 ms
MS1 resolution	30000	35000	60000	30000	70000
MS2 resolution	7500	17500	15000	7500	17500
MS1 m/z range	360-1300 m/z	350-1500 m/z	360-1300 m/z	300-1300 m/z	360-1300 m/z
TopN	10	20	15	10	20
Isolation Window	2.0 Th	1.5 Th	2.0 Th	2.0 Th	1.7 m/z
Fragmentation	HCD	HCD	HCD	HCD	HCD
NCE	35 %	25 %	30 %	40 %	25 %
Dynamic exclusion	20 s	40 s	20 s	30 s	35 s
Internal calibration m/z	401.92272	445.120025	401.92272	445.120025	401.92272
LC system	Eksigent	Eksigent	Eksigent	Eksigent	Eksigent
HPLC solvent A	0.1 % FA, 5 % DMSO in ddH ₂ O	0.1 % FA in ddH ₂ O	0.1 % FA, 5 % DMSO in ddH ₂ O	0.1 % FA in ddH ₂ O	0.1 % FA, 5 % DMSO in ddH ₂ O
HPLC solvent B	0.1 % FA, 5 % DMSO in ACN	0.1 % FA in ACN	0.1 % FA, 5 % DMSO in ACN	0.1 % FA in ACN	0.1 % FA, 5 % DMSO in ACN
Gradient length	225 min	180 min	100 min	225 min	110 min
Gradient	2 %-28 %	2 %-28 %	3 %-32 %	7 %-35 %	4 %-32 %

Detailed LC-MS/MS parameters can be found in table 1. Briefly, dissolved peptides (0.1 % FA in ddH₂O) were first delivered to a trap column and after 10 min of loading and washing, peptides were transferred to an analytical column. For mass spectrometric measurements of KBy pull downs performed with BT-474 lysates (lapatinib experiments) an Orbitrap Velos was used. Half of the phosphopeptide fractions were measured on an Orbitrap Velos and an Orbitrap Q-Exactive, respectively. An Orbitrap Elite was used for all other kinase affinity purifications in SK-N-BE(2) lysate (bead mixture and single inhibitor pull downs). Depending on the LC solvent composition, internal calibration was performed using a dimethyl sulfoxide cluster (m/z 401.922720) (29) or $(\text{Si}(\text{CH}_3)_2\text{O})_6\text{H}^+$ (m/z 445.120025) which is present in ambient air.

Data analysis

Data analysis was performed using MaxQuant v1.4.0.5 (30) with the integrated search engine Andromeda (31). For peptide and protein identification, raw files were searched against the UniProtKB database (v22.07.13) with carbamidomethylated cysteine as fixed modification and phosphorylation of serine, threonine and tyrosine, oxidation of methionine and N-terminal protein acetylation as variable modifications. For dimethyl labeling experiments dimethyl Lys0 and dimethyl Nter0 was set as the light and dimethyl Lys4 and dimethyl Nter4 as the heavy label. Trypsin/P was specified as the proteolytic enzyme with up to two missed cleavage sites allowed. Precursor tolerance was set to 10 ppm and fragment ion tolerance to 0.05 Da. Peptide identifications required a minimal length of 6 amino acids and all datasets were adjusted to 1 % PSM and 1 % protein FDR. For kinase affinity enrichments, protein identifications are only reported, if they are based on more than three unique and/or razor peptides. A two-sided t-test was used to assess statistical significance. For *in silico* combination of identified kinases from single inhibitor enrichments, an in-house programmed tool called “kinase-blender” was used. Inhibitor curves were fitted using a four parameter inhibitor response model with the initial residual binding fixed at 1 and a hill slope constraint of -1. Phosphopeptides were filtered for a localization probability > 0.75 (class I sites) (32). Quantification was performed label-free for kinase affinity enrichments, and by dimethyl labeling for enriched phosphopeptide fractions (using the MS1 intensities of the peptide ions in both cases). To determine protein intensities in label-free experiments, the intensities of razor and unique peptides were summed up. Resulting protein intensities were then normalized to obtain LFQ intensities. The match-between-runs option, which aligns features of different mass spectrometric runs based on accurate retention time and mass, was enabled for label-free and dimethyl labeled experiments.

Upregulated phosphorylation sites were further analyzed for enriched kinase substrate relationships using KinomeXplorer (NetworkKIN score > 2, number of substrates > 2, top two kinases per phosphorylation site) (33). Phosphoprotein interactions and enriched pathways were extracted using STRING (version 9.1; combined score > 0.9) (34) and KEGG (FDR corrected p-value < 0.01) (35). Resulting networks were exported and visualized with Cytoscape (version 3.1) (36). Phosphorylation site annotations were derived from PhosphoSitePlus™ (37).

Data availability

All raw mass spectrometry files and MaxQuant results files are available via PRIDE and ProteomeXchange with the accession code PXD002635 (38).

Results and discussion

Phosphoproteomic characterization of kinase activation upon pervanadate treatment

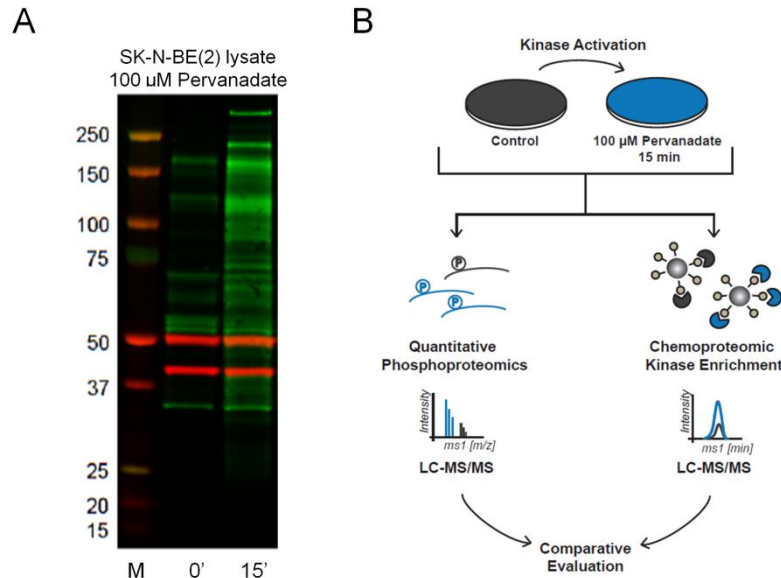


Figure 1. Experimental strategy for the characterization of activity dependent kinase binding to immobilized inhibitors. (A) SK-N-BE(2) cells were treated with 100 μM pervanadate for the indicated time points. Immunoblot analysis using an anti-phosphotyrosine antibody (pY-99) showed that tyrosine phosphorylation was increased by tyrosine phosphatase inhibition as expected. (B) Kinases in the SK-N-BE(2) neuroblastoma cell line were activated using 100 μM pervanadate for 15 min. Phosphorylation changes were quantified by dimethyl labelling-based phosphoproteomics and subsequently compared against differential kinase binding in a chemoproteomics setup, utilizing two different bead mixtures and the corresponding nine single probes.

Pervanadate addition to intact cells is known to indirectly activate tyrosine kinases by inhibition of tyrosine phosphatases (39, 40) and immunoblot analysis indeed showed a global boost in tyrosine phosphorylation in the model cell line SK-N-BE(2) used in this study upon 15 min treatment with 100 μM pervanadate (Figure 1A). To characterize and quantify the accompanying changes in protein phosphorylation, dimethyl labeled phosphopeptides were purified by Fe-IMAC chromatography (27) and the pool of phosphopeptides was analyzed by LC-MS/MS (Figure 1B). Moreover, the Fe-IMAC flow through which contains unphosphorylated peptides was measured in order to be able to determine changes in absolute phosphorylation site stoichiometries (also referred to as occupancies, see Olsen *et al.* (41)). From a total of 5,842 quantified phosphorylation sites, 210 sites were classified as down- and 886 sites as upregulated upon pervanadate treatment using an arbitrary fold change cut-off of $\log_2 < -1$ or > 1 (Figure 2A). As expected, 82 % of the 400 quantified phosphotyrosine sites were regulated

(Figure 2B) and a strong increase in phosphorylation sites of kinases that are known to increase their enzymatic activity was observed (Figure 2C) (e. g. MAPK1/3 (39, 42), EGFR (43), IGF1R (44), MAPK14 (42) and RAF1 (39), including several activation loop sites such as MAPK1-pT185/pY187, MAPK14-pY182, EGFR-pY869 and IGF1R-pY1161/pY1165). The data suggests that this is true for additional kinases such as RPS6KB1, RPS6KB4 and MAP2K4 (Figure 2C). Notably, I also found increased abundance of phosphorylation sites that are known to negatively regulate enzymatic activity leading to inactivation of these kinases (e. g. ROCK2, CHECK1). Phosphorylation increase was not only evident in relative but also in absolute terms indicating that a large molar fraction of several key kinases is activated by the treatment (Figure 2D). Notable examples for an increase in phosphorylation site occupancy of kinase activity determining sites include MAPK1-pY187 (from 12 % to 77 %), MAPK3-pY204 (from 10 % to 71 %) and AKT3-pS472 (from 29 % to 76 %) (see Figure 2E). In addition to the kinases themselves, I also analyzed phosphorylation levels of known kinase substrates as a second line of evidence for modulated kinase activity upon pervanadate treatment. To infer kinase activity from this data, all upregulated sites (\log_2 fold changes > 1 following pervanadate treatment) were analyzed with KinomeXplorer (33), a tool that scores kinase-substrate relationships based on linear motifs and network associations. Exemplified for AKT and MAPK1/3 substrates, the number and \log_2 fold changes of regulated substrates clearly indicated increased activation of the upstream kinase upon pervanadate treatment of cells (Figure 2F). Immunoblot analysis of the activity regulating sites pT202/pY204-MAPK1/3 and pS473-AKT (Figure 2G), confirmed the proteomic data and showed that such substrate footprints can serve as useful predictors of kinase activation. By the same token, activation of many tyrosine kinases such as PDGFRB, ERBB2, EGFR, FYN and ABL can be inferred because their substrates show increased phosphorylation (Figure 2H). In addition, pathway enrichment analysis of regulated protein phosphorylation identified e.g. the ERBB, VEGF and Insulin pathways underscoring the global impact of pervanadate treatment on cellular signaling (data not shown). Although it can neither be generalized that all kinases are affected by the treatment nor that abundance regulation of activity determining sites is always directly connected to kinase activation (45, 46), it can be concluded at this point that many key kinases show clear evidence for relative and absolute changes in activity.

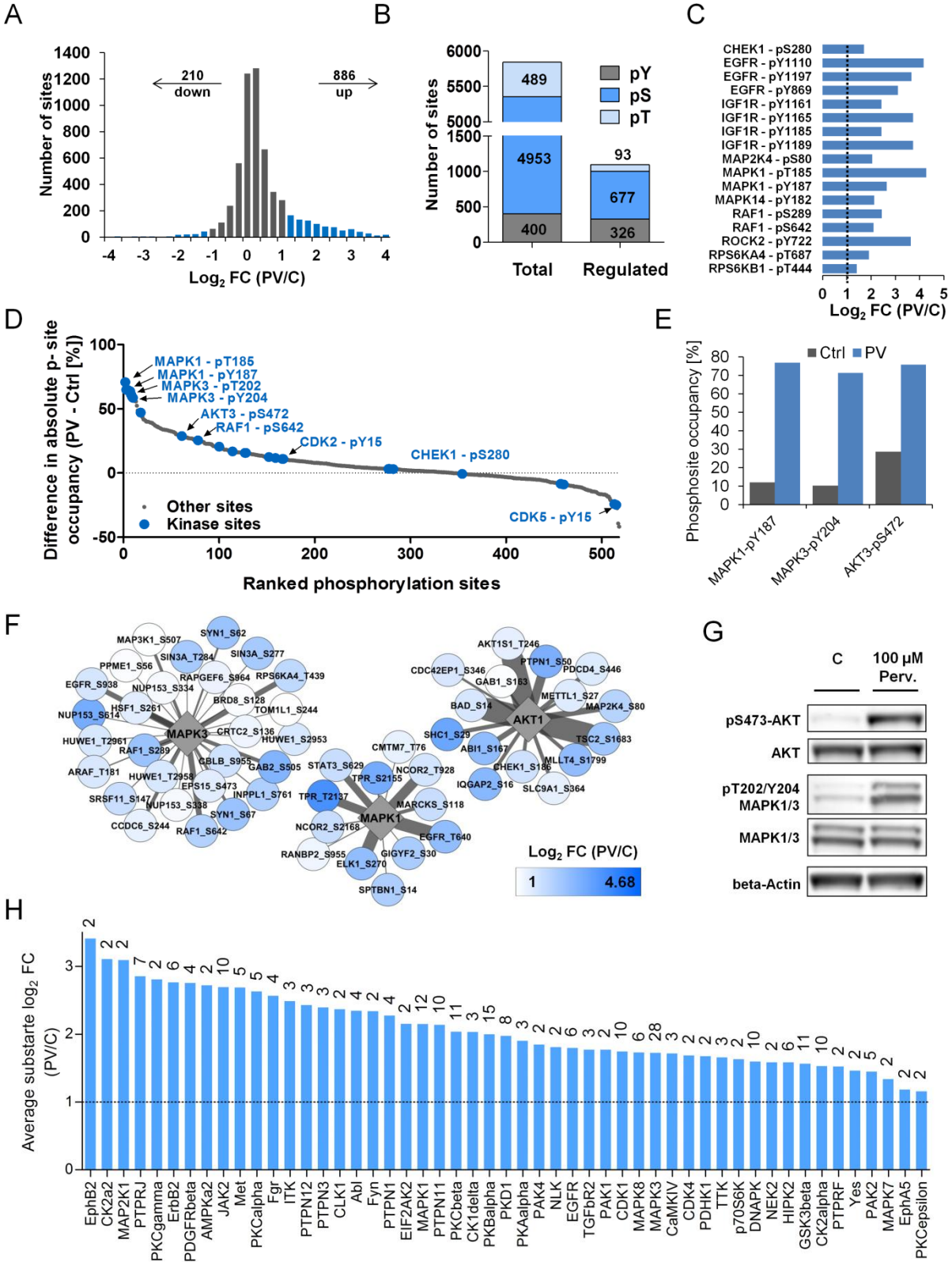


Figure 2. Quantitative phosphoproteomic characterization of kinase activation after pervanadate (PV) treatment. (A) Histogram of 5,842 confidently localized and quantifiable phosphorylation sites showing that 15 % of all sites were upregulated (\log_2 fold change > 1) indicating a global increase of phosphorylation upon pervanadate treatment (B) Due to tyrosine phosphatase inhibition, phosphotyrosine sites (7 % of all sites) were overrepresented compared to typical phosphoproteomic studies and the majority of phosphotyrosine sites (82 %) was found regulated by the treatment. (C) Activity regulating sites of key kinase receptors (MAPK1- pT185/pY187, MAPK14-pY182, EGFR-pY869 and IGF1R-pY1161/pY1165 are part of the kinase activation loop) and signaling nodes were found to be highly upregulated providing direct evidence for alteration of the state of kinase activity. (D) Strong absolute differences in phosphorylation site occupancy upon pervanadate treatment indicate that a large fraction of several key kinases is activated. Kinase sites are colored in blue and annotated activity influencing kinase sites are highlighted. (E) Phosphorylation site occupancy increase for MAPK1-pY187, MAPK3-pY204 and AKT3-pS472 upon treatment. (F) After assignment of upregulated phosphorylation sites to their respective kinases, AKT1, MAPK3 and MAPK1 substrates were found to be highly enriched, providing a footprint of kinase activity. Node colors indicate the \log_2 phosphorylation change upon pervanadate treatment and the edge thickness represents the NetworkKIN score for each kinase substrate interaction. (G) Immunoblot analysis using antibodies against p-T202/p-Y204-MAPK1/3 and p-S473-AKT indicated activation of AKT, MAPK1 and MAPK3. (H) Average \log_2 fold change for all substrates matching to a kinase or phosphatase using KinomeXplorer. Numbers on top of the bars indicate how many substrates per kinase/phosphatase were found in the dataset of upregulated phosphorylation sites.

Kinase binding to immobilized inhibitor beads is largely independent of kinase activity

With clear evidence for pervanadate induced activation of several kinases in hand, I sought to examine if and to what extent differences in kinase activation status translates into differential binding to immobilized, broadly selective kinase inhibitors. To effect this, two different inhibitor bead mixtures were composed: Kinobeads version gamma (KBy) (13) and a MIB-like mixture consisting of compounds that either exactly match or closely resemble those described by Duncan *et al.* (17) (Figure 3). Strikingly, the vast majority of kinases for which an activation change was observed in the pervanadate treatment (either by the direct identification of an increased abundance of an activity inducing site or indirectly by the identification of an abundance increase in a kinase substrate) was not significantly differentially bound by MIB-like or KBy beads (Figure 4A). In other words, binding of these kinases was either only marginally or not at all influenced by their activation status, irrespective of which bead mixture was used. However, a small number of kinases (IGF1R, MAPK1, MAPK3, RPS6KB1 and RAF1) showed a reproducible affinity increase upon activation. But except for RPS6KB1 (for KBy) and MAPK3 (for MIB-like), none of the affinity changes fully reflected the quantitative extent of regulation observed at the phosphorylation level (Figure 4A). This suggests that even if some few kinases can be preferentially bound to beads in one conformation, the extent to which this preference can be detected will depend on the relative proportion of the respective kinase conformation in the cell or lysate. Interestingly, and despite strong activation of EGFR by pervanadate (deduced from the increased abundance of enzymatic activity inducing sites: pY1110, pY1197, pY869 and

several substrate sites; Figure 4A), the protein is actually lost from KBy upon pervanadate treatment. Kinase substrate footprints reveal the same trend for CLK1 (MIB-like) and for PDGFRB (KBy and MIB-like) (Figure 4A). This observation implies that these kinases are preferentially captured in their inactive conformation.

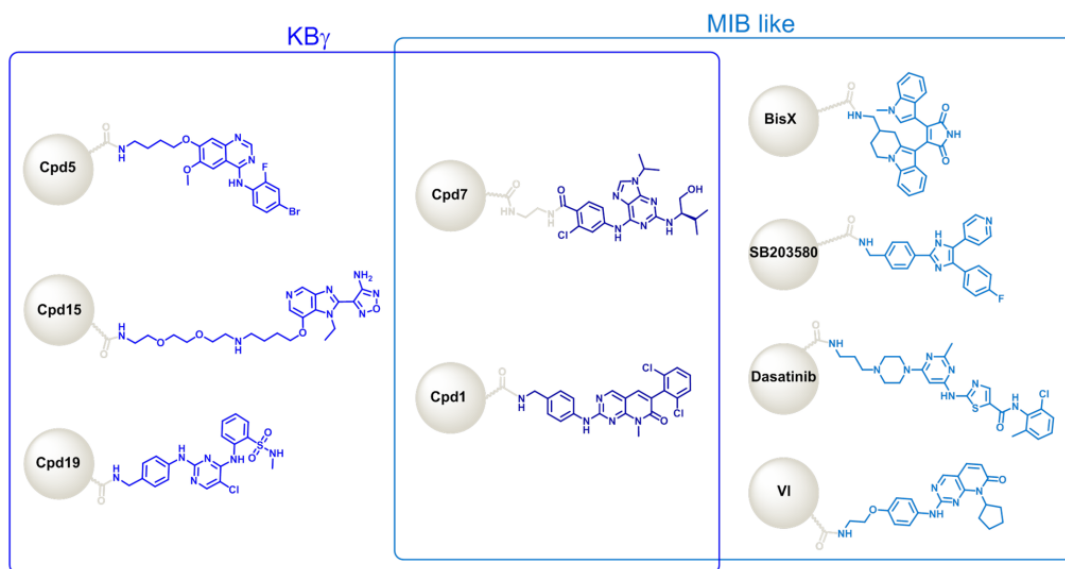


Figure 3. Bead mix compositions and structures of the immobilized inhibitors used in this study.

Since phosphorylation site abundance information is not available for all bead enriched kinases, differential kinase binding was compared globally using KBy and MIB-like beads. A volcano plot summarizing the results of replicate enrichments show that only 25 out of the 165 kinases quantified in KBy pull downs changed significantly in abundance (14 up and 11 down; $p < 0.01$; Figure 4B). Similar results were obtained for MIB-like beads: out of 122 quantified kinases, 16 showed increased and 16 showed reduced abundance upon pervanadate treatment ($p < 0.01$; Figure 4B). In addition, no significant differences in binding behavior were detected for the different kinase families (Figure 4C). These findings are in clear contrast to what was reported by Duncan *et al.* (17), where a comparable experimental setup was used and in which the authors concluded that a global increase of tyrosine kinase binding to MIBs was observed upon pervanadate treatment. 113 kinases were commonly detected using KBy and MIB-like beads. Whereas RPS6KA3, RAF1, ARAF, PDGFRB or DDR1/2 showed similar binding behavior on both beads, others including MAPK1/3 and RPS6KA1 were differentially detected in only one of the two datasets (Figure 4D). From this data it can be concluded that in addition to the kinase itself, the inhibitor composition of the bead matrix also affects conformation dependent binding.

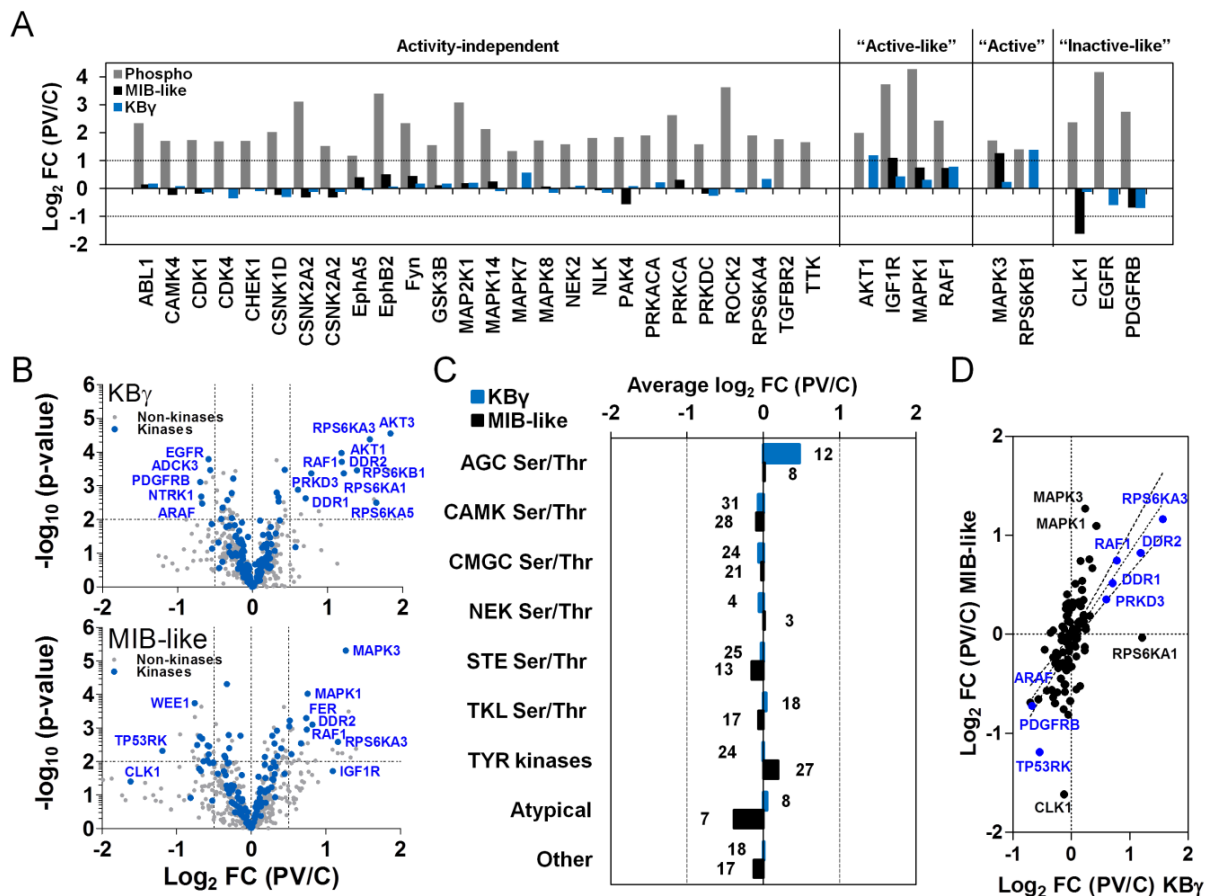


Figure 4. Enrichment of kinases by immobilized inhibitor beads with and without pervanadate (PV) treatment. (A) Kinases binding to MIB-like and KBy beads upon pervanadate treatment was compared to the phosphorylation intensity changes (\log_2 scale) of kinase activity regulating sites or kinase substrates (in cases where both information was available, the activity regulating site is depicted). Whereas the majority of kinases bound to both affinity matrices in an activity independent fashion, some few kinases were captured in an “active-like” or “inactive-like” conformation. Strikingly only two kinases showed a phosphorylation change comparable to that of the binding increase to affinity matrices following pervanadate treatment (RPS6KB1 for KBy and MAPK3 for the MIB-like mix). (B) Volcano plots showing pervanadate induced \log_2 fold changes of quantified kinases (165 for KBy and 122 for the MIB-like mix with ≥ 4 peptides) plotted against their p-values (technical triplicates). For both the MIB-like bead mix and KBy, the number of kinases showing increased binding (KBy: 14, MIB-like: 16; $p < 0.01$) was comparable to that showing decreased binding upon signaling activation (KBy: 11, MIB-like: 16; $p < 0.01$). Hence, the majority of kinases is not bound in an activity dependent fashion and no global correlation between chemoproteomic affinity increase and signaling activation is evident. (C) Average pervanadate induced \log_2 fold changes of enriched kinase families for KBy and the MIB-like beads suggesting that kinase capture is independent of kinase family (numbers indicate family members). (D) A comparison of pervanadate induced \log_2 fold changes for overlapping kinases in KBy and MIB-like experiments (113 kinases) reveals that activity-based kinase binding is dependent on the composition of the affinity matrix (dashed lines indicates the mean including the 95 % confidence interval).

Pharmacological kinase inactivation does not change affinity to immobilized inhibitor beads

Next, I sought to confirm the above observation that kinase binding to immobilized inhibitor beads is generally not dependent on kinase activity, by reversing the above pervanadate activation experiments. Specifically, I used the highly selective EGFR/ERBB2 inhibitor lapatinib to treat the lapatinib sensitive and ERBB2 overexpressing breast cancer cell line BT-474 in order to inactivate EGFR/ERBB2 as well as further downstream kinases in the same pathway (in biological duplicates; as a control, I performed KBy selectivity profiling in lysates of BT-474 cells to confirm the exquisite selectivity of this inhibitor). Upon treatment of cells with increasing concentrations of lapatinib (1 nM – 10 μ M), KBy pull downs were performed from each lysate. Would kinase affinity to immobilized inhibitor beads depend on the activation status of the respective kinase, one would expect to observe decreased binding to KBy in this experiment. Strikingly, apart from the direct lapatinib targets ERBB2 and EGFR, only one out of 178 kinases quantified in this experiment (RPS6KB1; a known downstream kinase in the EGFR/ERBB2 pathway), was dose-dependently lost from the beads (IC_{50} = 38 nM; Figure 5A, Figure 5B). Competition experiments carried out in lysates show that RPS6KB1 is not a direct target of lapatinib (Figure 5B). RPS6KB1 was also the only kinase for which the binding increase after pervanadate treatment nearly fully mirrored the observed phosphorylation site changes in the above experiments using SK-N-BE(2) cells (Figure 4A), which is something one would expect if only the active kinase was bound in the first place. In addition, I was able to determine several site occupancies. While increasing doses of lapatinib led to the complete loss of phosphorylation of AKT1-pS472 (from 26% to 0%, see Figure 5C), MAPK1-pY187 (from 46 % to 0 %, see Figure 5C) and GSK3A-pY279 (from 86 % to 0 % (data not shown), the total amount of the respective kinases captured on KBy remained the same. This clearly shows that the beads bind both, the active and inactive form of the kinases. Surprisingly, ARAF, IRAK1, SPK1, ULK1 and ERBB3 showed increased binding to KBy upon lapatinib treatment (Figure 5A). Unfortunately, these proteins were not detected in the above pervanadate treatment and phosphorylation analysis. Hence, no conclusions can be drawn at this stage as to which conformation of these kinases may be responsible for the observed effect.

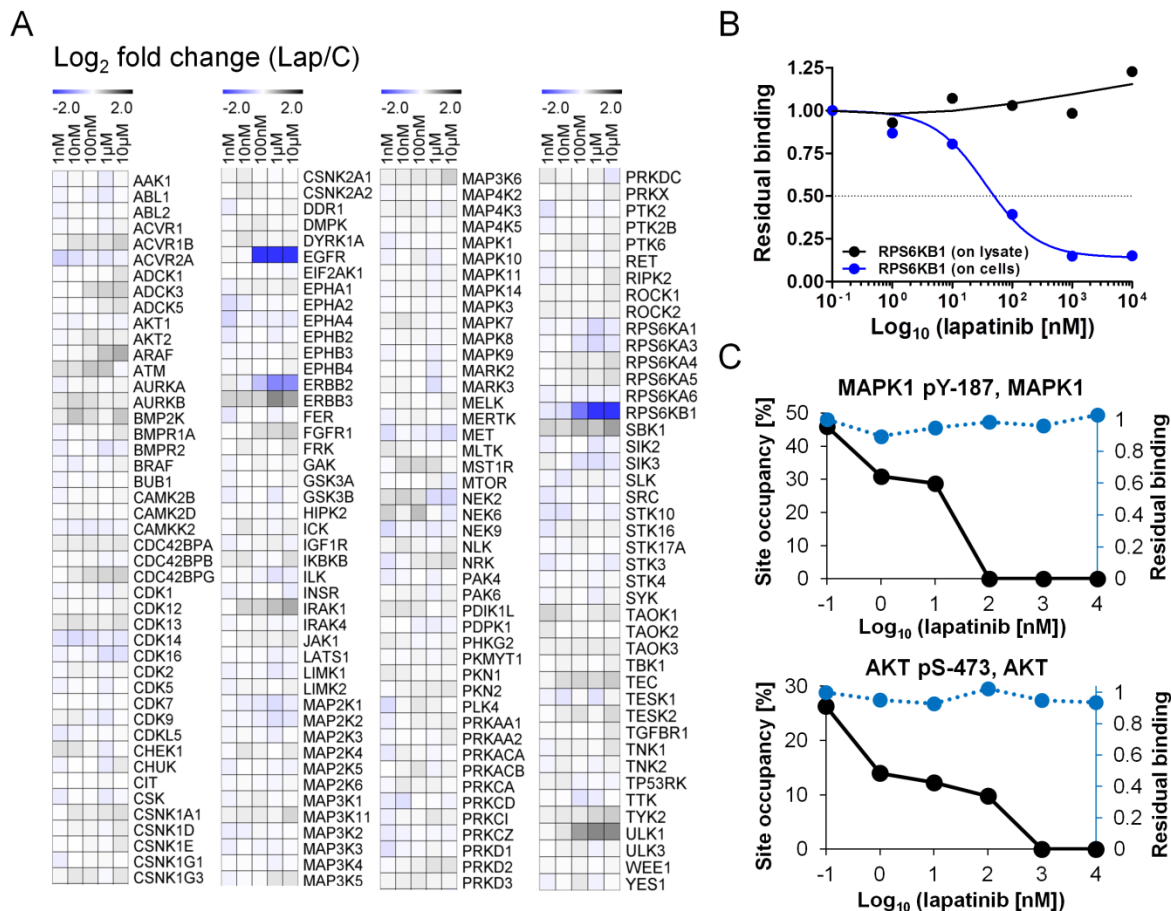


Figure 5. Enrichment of kinases by immobilized inhibitor beads with and without EGFR inhibitor treatment (A) Heat map with color coded log_2 fold change of the 178 quantified kinases in lapatinib (lap) treated (1 nM - 10 μM for 60 min) and control cells (C). Apart from the direct lapatinib targets EGFR and ERBB2, only the downstream kinase RPS6KB1 is dose-dependently lost from KBy. (B) Apart from the primary targets of lapatinib, only one known downstream pathway member, RPS6KB1, showed a dose-dependent inactivation in cells indicating that lapatinib actually shuts down the pathway in cells. The control Kinobead experiments using lapatinib treatment in lysates showed no effect confirming that RPS6KB1 is not a direct target of lapatinib. (C) MAPK1 and AKT kinase binding profiles to KBy were not altered, while the occupancy of their activity inducing phosphorylation sites show a clear dose-dependent decrease of binding to KBy implying that even though MAPK1/3 and AKT activity was reduced by the drug, their binding to Kinobeads was unaffected. This provides direct evidence for the fact that both conformations of the respective kinases are bound to the beads.

Use of immobilized inhibitor bead mixtures occlude activity related binding affinity effects

The above experiments have shown that binding of kinases to immobilized inhibitor beads cannot generally be correlated with a kinase's activation status, but they also indicated that for some kinases, this does actually occur. To investigate the role the immobilized inhibitors have to

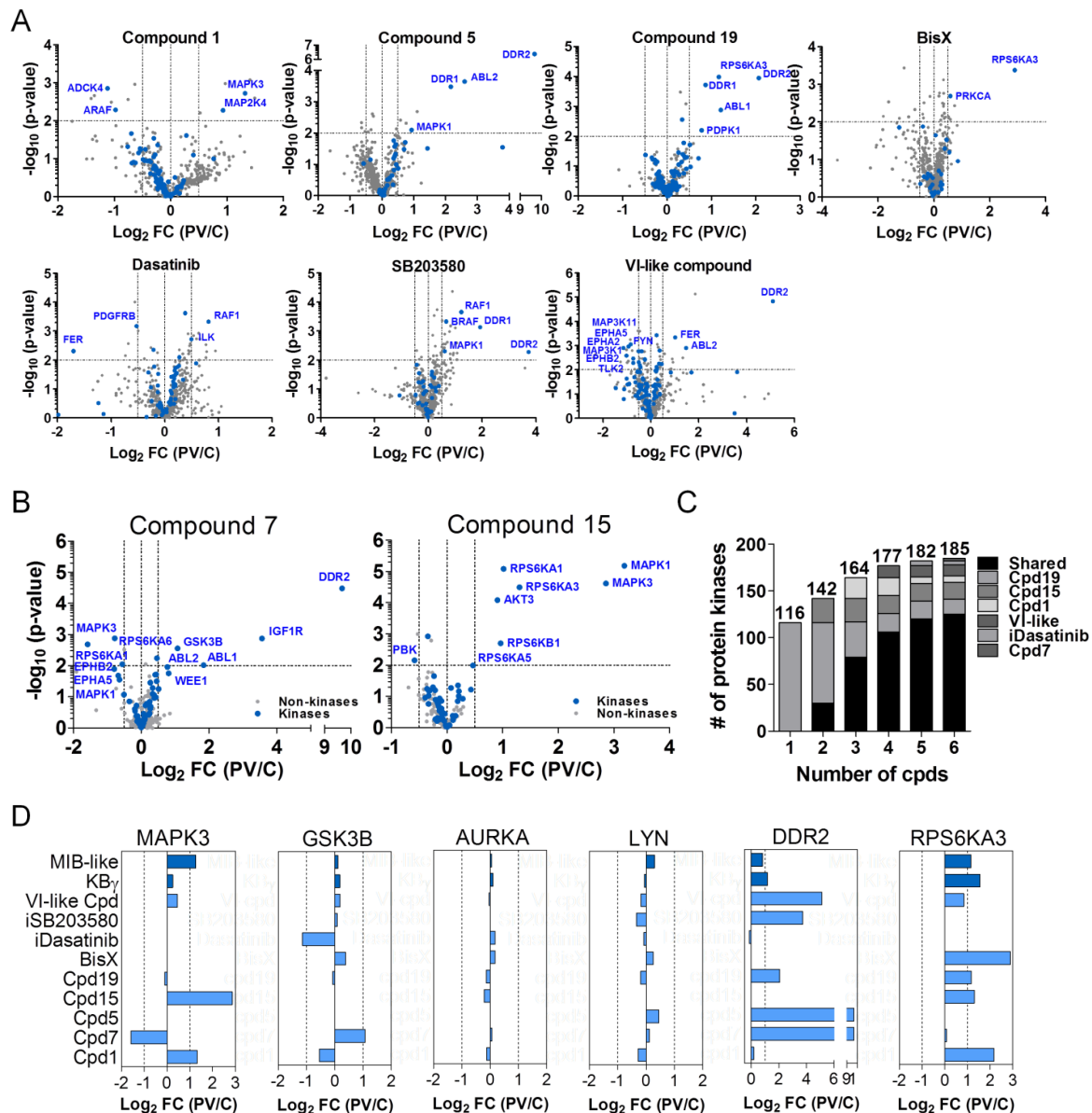


Figure 6. Decomposition of immobilized inhibitor mixtures into single probes (A) Volcano Plots showing the kinase \log_2 fold changes between pervanadate (PV) treated and control cells (C) for seven individual inhibitor resins plotted against their p-value (technical triplicates). (B) Volcano plots summarizing triplicate pull down experiments using compound 15 and compound 7 showing that kinases like MAPK1, MAPK3 and RPS6KA1 exhibit conformation dependent but opposing binding behavior towards the two compounds. Bead specific conformational selectivity might therefore be diluted upon mixing several inhibitor beads. In conclusion, both kinase activity and the immobilized inhibitor itself determine whether or not a kinase is captured in a conformation dependent fashion. (C) Results from experiments using single inhibitor beads representing the MIB-like and the KBy mixtures, were combined in silico. The inherent redundancy of kinase capture by the beads is illustrated by the number of shared kinases between the inhibitor beads which is increasing with the number of compounds. (D) Examples of pervanadate induced \log_2 fold changes of kinases across the nine different single compounds and the bead mixtures (if a kinase was not identified, no column is shown). Whereas some kinases are differentially bound by compounds (e.g. MAPK3, GSK3B) or show conformational enrichment (e.g. DDR2, RPS6KA3), the majority of kinases does not show such differential binding (illustrated here by LYN and AURKA).

play in this context, the KBy and MIB-like bead mixtures were decomposed and pull downs with and without pervanadate treatment of SK-N-BE(2) cells using the nine individual inhibitor resins were conducted (Figure 6A and 6B). It turned out that there is a lot of redundancy in the kinases that are enriched by any of these affinity beads and more than half of all kinases are enriched by two or more affinity matrices (Figure 6C). It is therefore not unreasonable to expect that activity dependent capture of kinases by one immobilized compound could be masked by another if used in a mixture as noted in Médard *et al.* (13). This notion is illustrated by a comparison of the range of kinases enriched by compound 7 and compound 15 (both included in KBy) shown in Figure 6B. While MAPK1, MAPK3 and RPS6KA1 bind significantly stronger to compound 15 in response to pervanadate treatment (\log_2 fold changes of 3.2, 2.9, 1.0 and p -values of 6.6E-6, 2.4E-5, 8.3E-6), they are lost from compound 7 upon pervanadate treatment (\log_2 fold changes of -0.6, -1.6, -0.6 and p -values of 2.8E-2, 2.1E-3, 9.0E-3). This indicates that the two compounds preferentially bind to different conformations of those kinases. Mixing both resins can therefore mask such opposing trends and complicate the delineation of activity dependent binding events. Figure 6D shows examples of a number of kinases with diverse binding behavior upon pervanadate treatment for all nine inhibitor resins as well as KBy and MIB-like beads. LYN and AURKA are captured on any of the beads without apparent preference for an active or inactive conformation (which is true for the vast majority of all kinases). For MAPK3 and GSK3B, opposing effects are observed for individual resins but much of these are equalized in the bead mixtures. DDR2 and RPS6KA3 are notable cases for which a marked increase in affinity was observed for several compounds upon pervanadate treatment and this effect was also visible (albeit much less strongly) in mixed bead experiments. Notably, for both kinases, an extensive structural rearrangement upon activation is supported by crystallographic data (47, 48). Hence, for certain combinations of kinase and immobilized compound, it may be possible to infer kinase activity from the binding data (compound 5 and ABL1/ABL2 is another example; Figure 6A). This may be useful in cases where specific phospho-antibodies indicating kinase activity are not available (such as for DDR2). More generally, the data set and analysis provided in this study highlight a number of kinases for which immobilized inhibitors could potentially provide an experimental assay platform for the discovery of conformation selective molecules, a concept that should be tested by further experiments in the future.

Conclusion

The present work refines the characterization of conformation dependent binding of kinases to immobilized inhibitors. In contrast to previous reports, only very few kinases are preferentially enriched in an activity indicating conformation, whereas binding of the majority of kinases is unaffected by activating or inactivating stimuli. This calls previous claims of global activity dependent kinase binding into question and rather suggests that neither conformational selectivity nor concomitant loss or gain of kinase binding is predictable but must be carefully characterized on a case by case basis.

Acknowledgements

The author wants to thank Jana Zecha for conduction of the chemo- and phosphoproteomic experiments during her M.Sc. thesis under my supervision and for the fruitful discussion throughout the project. The author also gratefully acknowledges Stephanie Heinzlmeir and Guillaume Médard for their help with compound synthesis, Susan Klaeger for mass spectrometric measurements and Andreas Klaus for excellent technical assistance.

Abbreviations

ATP	Adenosine triphosphate
C/Ctrl	Controle
FDR	False discovery rate
I.D.	Inner diameter
IMAC	Immobilized metal ion affinity chromatography
KBy	Kinobeads gamma
LC	Liquid chromatography
MIB	Multiplexed inhibitor beads
MS	Mass spectrometry
ppm	Parts per million
PSM	Peptide spectrum match
PV	Pervanadate
p-Y/S/T	Phosphotyrosine, -serine, -threonine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

References

1. Hunter, T. (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80, 225–236
2. Blume-Jensen, P., and Hunter, T. (2001) Oncogenic kinase signalling. *Nature* 411, 355–365
3. Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* 9, 28–39
4. Schirle, M., Bantscheff, M., and Kuster, B. (2012) Mass spectrometry-based proteomics in preclinical drug discovery. *Chem. Biol.* 19, 72–84
5. Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H., and Peterson, J. R. (2011) Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1039–1045
6. Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., Hocker, M., Treiber, D. K., and Zarrinkar, P. P. (2011) Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1046–1051
7. Wissing, J., Jansch, L., Nitz, M., Dieterich, G., Hornberger, R., Kéri, G., Wehland, J., and Daub, H. (2007) Proteomics analysis of protein kinases by target class-selective prefractionation and tandem mass spectrometry. *Mol. Cell. Proteomics MCP* 6, 537–547
8. Midland, A. A., Whittle, M. C., Duncan, J. S., Abell, A. N., Nakamura, K., Zawistowski, J. S., Carey, L. A., Earp Iii, H. S., Graves, L. M., Gomez, S. M., and Johnson, G. L. (2012) Defining the expressed breast cancer kinome. *Cell Res.* 22, 620–623
9. Patricelli, M. P., Nomanbhoy, T. K., Wu, J., Brown, H., Zhou, D., Zhang, J., Jagannathan, S., Aban, A., Okerberg, E., Herring, C., Nordin, B., Weissig, H., Yang, Q., Lee, J.-D., Gray, N. S., and Kozarich, J. W. (2011) In Situ Kinase Profiling Reveals Functionally Relevant Properties of Native Kinases. *Chem. Biol.* 18, 699–710
10. Ku, X., Heinzlmeir, S., Helm, D., Médard, G., and Kuster, B. (2014) New Affinity Probe Targeting VEGF Receptors for Kinase Inhibitor Selectivity Profiling by Chemical Proteomics. *J. Proteome Res.* 13, 2445–2452
11. Pachel, F., Plattner, P., Ruprecht, B., Médard, G., Sewald, N., and Kuster, B. (2013) Characterization of a Chemical Affinity Probe Targeting Akt Kinases. *J. Proteome Res.* 12, 3792–3800
12. Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* 25, 1035–1044
13. Médard, G., Pachel, F., Ruprecht, B., Klaeger, S., Heinzlmeir, S., Helm, D., Qiao, H., Ku, X., Wilhelm, M., Kuehne, T., Wu, Z., Dittmann, A., Hopf, C., Kramer, K., and Kuster, B. (2015) Optimized Chemical Proteomics Assay for Kinase Inhibitor Profiling. *J. Proteome Res.*
14. Gholami, A. M., Hahne, H., Wu, Z., Auer, F. J., Meng, C., Wilhelm, M., and Kuster, B. (2013) Global Proteome Analysis of the NCI-60 Cell Line Panel. *Cell Rep.* 4, 609–620
15. Wu, Z., Doondeea, J. B., Gholami, A. M., Janning, M. C., Lemeer, S., Kramer, K., Eccles, S. A., Gollin, S. M., Grenman, R., Walch, A., Feller, S. M., and Kuster, B. (2011) Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer. *Mol. Cell. Proteomics MCP* 10, M111.011635
16. Cooper, M. J., Cox, N. J., Zimmerman, E. I., Dewar, B. J., Duncan, J. S., Whittle, M. C., Nguyen, T. A., Jones, L. S., Ghose Roy, S., Smalley, D. M., Kuan, P. F., Richards, K. L., Christopherson, R. I., Jin, J., Frye, S. V., Johnson, G. L., Baldwin, A. S., and Graves, L. M. (2013) Application of multiplexed kinase inhibitor beads to study kinome adaptations in drug-resistant leukemia. *PLoS One* 8, e66755
17. Duncan, J. S., Whittle, M. C., Nakamura, K., Abell, A. N., Midland, A. A., Zawistowski, J. S., Johnson, N. L., Granger, D. A., Jordan, N. V., Darr, D. B., Usary, J., Kuan, P.-F., Smalley, D. M., Major, B., He, X., Hoadley, K. A., Zhou, B., Sharpless, N. E., Perou, C. M., Kim, W. Y., Gomez, S. M., Chen, X., Jin, J., Frye, S. V., Earp, H. S., Graves, L. M., and Johnson, G. L. (2012) Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* 149, 307–321

18. Becher, I., Savitski, M. M., Savitski, M. F., Hopf, C., Bantscheff, M., and Drewes, G. (2013) Affinity Profiling of the Cellular Kinome for the Nucleotide Cofactors ATP, ADP, and GTP. *ACS Chem. Biol.* 8, 599–607
19. Brehmer, D., Godl, K., Zech, B., Wissing, J., and Daub, H. (2004) Proteome-wide identification of cellular targets affected by bisindolylmaleimide-type protein kinase C inhibitors. *Mol. Cell. Proteomics MCP* 3, 490–500
20. Graves, L. M., Duncan, J. S., Whittle, M. C., and Johnson, G. L. (2013) The dynamic nature of the kinome. *Biochem. J.* 450, 1–8
21. Stuhlmiller, T. J., Earp, H. S., and Johnson, G. L. (2014) Adaptive reprogramming of the breast cancer kinome. *Clin. Pharmacol. Ther.* 95, 413–415
22. Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* 2, 358–364
23. Stuhlmiller, T. J., Miller, S. M., Zawistowski, J. S., Nakamura, K., Beltran, A. S., Duncan, J. S., Angus, S. P., Collins, K. A. L., Granger, D. A., Reuther, R. A., Graves, L. M., Gomez, S. M., Kuan, P.-F., Parker, J. S., Chen, X., Sciaky, N., Carey, L. A., Earp, H. S., Jin, J., and Johnson, G. L. (2015) Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep.* 11, 390–404
24. Munoz, L., Selig, R., Yeung, Y. T., Peifer, C., Hauser, D., and Laufer, S. (2010) Fluorescence polarization binding assay to develop inhibitors of inactive p38alpha mitogen-activated protein kinase. *Anal. Biochem.* 401, 125–133
25. Biftu, T., Feng, D., Ponpipom, M., Girotra, N., Liang, G.-B., Qian, X., Bugianesi, R., Simeone, J., Chang, L., Gurnett, A., Liberator, P., Dulski, P., Leavitt, P. S., Crumley, T., Misura, A., Murphy, T., Rattray, S., Samaras, S., Tamas, T., Mathew, J., Brown, C., Thompson, D., Schmatz, D., Fisher, M., and Wyvratt, M. (2005) Synthesis and SAR of 2,3-diarylpyrrole inhibitors of parasite cGMP-dependent protein kinase as novel anticoccidial agents. *Bioorg. Med. Chem. Lett.* 15, 3296–3301
26. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. R. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* 4, 484–494
27. Ruprecht, B., Koch, H., Medard, G., Mundt, M., Kuster, B., and Lemeer, S. (2015) Comprehensive and Reproducible Phosphopeptide Enrichment Using Iron Immobilized Metal Ion Affinity Chromatography (Fe-IMAC) Columns. *Mol. Cell. Proteomics* 14, 205–215
28. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* 68, 850–858
29. Hahne, H., Pachi, F., Ruprecht, B., Maier, S. K., Klaeger, S., Helm, D., Médard, G., Wilm, M., Lemeer, S., and Kuster, B. (2013) DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat. Methods* 10, 989–991
30. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372
31. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* 10, 1794–1805
32. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S., and Mann, M. (2010) Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* 3, ra3
33. Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., Palma, A., Cesareni, G., Jensen, L. J., and Linding, R. (2014) KinomeXplorer: an integrated platform for kinome biology studies. *Nat. Methods* 11, 603–604
34. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., Mering, C. von, and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808–D815
35. Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30
36. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504

37. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270
38. Vizcaíno, J. A., Côté, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., Griss, J., Alpi, E., Birim, M., Contell, J., O’Kelly, G., Schoenegger, A., Ovelleiro, D., Pérez-Riverol, Y., Reisinger, F., Ríos, D., Wang, R., and Hermjakob, H. (2013) The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 41, D1063–D1069
39. Zhao, Z., Tan, Z., Diltz, C. D., You, M., and Fischer, E. H. (1996) Activation of mitogen-activated protein (MAP) kinase pathway by pervanadate, a potent inhibitor of tyrosine phosphatases. *J. Biol. Chem.* 271, 22251–22255
40. Boersema, P. J., Foong, L. Y., Ding, V. M. Y., Lemeer, S., Breukelen, B. van, Philp, R., Boekhorst, J., Snel, B., Hertog, J. den, Choo, A. B. H., and Heck, A. J. R. (2010) In-depth Qualitative and Quantitative Profiling of Tyrosine Phosphorylation Using a Combination of Phosphopeptide Immunoaffinity Purification and Stable Isotope Dimethyl Labeling. *Mol. Cell. Proteomics* 9, 84–99
41. Olsen, J. V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127, 635–648
42. Park, J., and Liu, A. Y. (2000) Pervanadate induces the hyperphosphorylation but not the activation of human heat shock factor 1. *J. Cell. Physiol.* 185, 348–357
43. Reynolds, A. R., Tischer, C., Verveer, P. J., Rocks, O., and Bastiaens, P. I. H. (2003) EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation. *Nat. Cell Biol.* 5, 447–453
44. Shisheva, A., and Shechter, Y. (1993) Mechanism of pervanadate stimulation and potentiation of insulin-activated glucose transport in rat adipocytes: dissociation from vanadate effect. *Endocrinology* 133, 1562–1568
45. Dalby, K. N., Morrice, N., Caudwell, F. B., Avruch, J., and Cohen, P. (1998) Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK. *J. Biol. Chem.* 273, 1496–1505
46. Lara, R., Seckl, M. J., and Pardo, O. E. (2013) The p90 RSK Family Members: Common Functions and Isoform Specificity. *Cancer Res.* 73, 5301–5308
47. Carafoli, F., Bihan, D., Stathopoulos, S., Konitsiotis, A. D., Kvangsakul, M., Farndale, R. W., Leitinger, B., and Hohenester, E. (2009) Crystallographic insight into collagen recognition by discoidin domain receptor 2. *Struct. Lond. Engl.* 1993 17, 1573–1581
48. Malakhova, M., Tereshko, V., Lee, S.-Y., Yao, K., Cho, Y.-Y., Bode, A., and Dong, Z. (2008) Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2. *Nat. Struct. Mol. Biol.* 15, 112–113

Chapter 5

Multi proteomic dissection of lapatinib mode of action and resistance in ERBB2 overexpressing breast cancer

Abstract

Despite initially high response rates of the small molecule kinase inhibitor lapatinib in ERBB2 overexpressing breast cancer, the acquisition of drug resistance frequently occurs. Here, I used an established cell line model and applied a combination of mass spectrometry-based proteomics, chemoproteomics and phosphoproteomics in an effort to globally assess the molecular consequences of lapatinib treatment and resistance. The resulting dataset, which collectively comprises quantitative values for > 7,800 proteins, > 300 protein kinases and > 15,000 phosphopeptides enabled deep insight into signaling recovery and molecular reprogramming upon resistance. Importantly, the approach readily confirms and extends a previously described mechanism of resistance (AXL overexpression, PIK3 reactivation) and reveals the occurrence of a wealth of other, novel, pharmacologically actionable targets (e.g. CDK1/2, the spliceosome or EEF2K). Although a comparison of the dataset to previously described mechanisms of ERBB2 resistance in breast cancer suggests a great heterogeneity and context specific functionality of molecular resistance drivers, it also confirms the addiction to anaerobic glycolysis. In contrast to previous, expression-based mechanisms, this study uncovers a phosphorylation mediated reprogramming of LDHA and PDHA1 activity which increases the sensitivity of the resistant cells to glycolysis inhibition. As glucose addiction can occur via multiple different routes, this phenotype might potentially represent a common and targetable convergence point and, as such, a universal “Achilles heel” of resistance to ERBB2 targeted therapies in breast cancer.

Introduction

The receptor tyrosine kinase ERBB2 (Her2) is overexpressed in 20 % - 30 % of all breast tumors and leads to an increase in the proliferative and invasive potential which is associated with poor patient survival (1, 2). Mechanistically, high levels of ERBB2 cause homodimerization, autophosphorylation and activation of downstream signaling pathways also in the absence of a stimulating ligand (3). The aberrant signal is mainly transferred via the PIK3/AKT/mTOR and the RAF/MEK/ERK kinase cascades and ultimately results in uncontrolled cell growth. Pharmacological efforts directed towards ERBB2 resulted in the FDA approval of trastuzumab (4, 5, 6), a monoclonal antibody which prevents ERBB2 dimerization, and lapatinib (7, 8), a small molecule EGFR/ERBB2 inhibitor which blocks the kinases active site. Despite initially high response rates to targeted therapies, the acquisition of drug resistance frequently, if not inevitably, occurs. Owing to the early FDA approval of ERBB2 targeted therapies and their clinical prevalence in breast cancer, a wealth of different resistance mechanisms has been described to date. Examples include a signaling switch to other ERBB family members (EGFR (9), ERBB3 (10)), the compensatory upregulation of alternative receptor tyrosine kinases (EPHA2 (11), IGFR (12, 13), MET (14) or MERTK (RON) (15)), the activation of downstream kinases (e.g. PRKACA (16), SRC (17, 18, 19) or activating PIK3CA mutations (20, 21)). Non-kinase mediated mechanisms of resistance are for instance the overexpression of CCNE (22), the loss of signaling proteins (e.g. PTEN (23), CDKN1B (p27KIP1) (24), metabolic dependency on glucose metabolism (either by HSF1 mediated LDHA overexpression (25) or the overexpression of several metabolic enzymes (26)) and the activation of estrogen receptor signaling (27).

Liu and colleagues established the overexpression of the receptor tyrosine kinase AXL as a novel cause for lapatinib resistance in the ERBB2 overexpressing breast cancer cell line BT-474 (28). In their study they showed that AXL engaged PIK3 which in turn restored proliferation by recovery of the AKT/mTOR signaling branch. Here I used this cell line model of lapatinib resistance and employed explorative mass spectrometry to profile the proteome, kinome and phosphoproteome changes in an effort to robustly characterize resistance development on a global scale and to identify additional and potentially common molecular events inherent to ERBB2 inhibitor resistance. Collectively, this in depth, multi proteomic analysis offers an unprecedented perspective on the molecular mechanisms of resistance and shows that the acquisition is accompanied by many targetable alterations. The results concomitantly suggest the existence of common convergence nodes, such as the increased metabolization of glucose.

Material and methods

Cell culture and reagents

Parental BT-474 cells and its lapatinib resistant clone BT-474-J4 were grown in DMEM/HamsF12 medium (BioRad) supplemented with 15 % (v/v) Fetal Bovine Serum (FBS) and 1 % (v/v) Antibiotic/Antimycotic solution (Sigma). Resistant BT-474-J4 cells were cultured in the continuous presence of 1 μ M lapatinib. Biological replicates were prepared at different days using a different passage of both cell lines. Lapatinib, dasatinib and bosutinib were purchased from LC Laboratories, 2-deoxy-D-glucose was purchased from Sigma Aldrich and saracatinib, selumetinib linsitinib, BMS-387032 and SCH-727965 were purchased from Selleckchem. For viability/drug treatment assays, cells were seeded in 96-well plates at a concentration of 4×10^4 cells/well with complete culture medium. The next day, cells were exposed to increasing concentrations of the inhibitor or vehicle control for the indicated amount of time. Cell viability was measured using the AlamarBlue® Cell Viability Assay (ThermoFisher Scientific) according to manufacturer's instructions. Sigmoidal dose response curves were fitted using a nonlinear regression model in GraphPad Prism v.5.01.

Invasion and migration assay

The invasive and migrative potential of parental and resistant BT-474 cell lines was assessed in a transwell assay. Cells were starved in serum-free medium for 24 h, before they were placed in ThinCert 24-well cell culture inserts with 8 μ m pores (Greiner Bio-One). For invasion measurements, BD Matrigel (VWR) was diluted to a concentration of 50 μ g/ml with ice-cold coating buffer (0.01 M Tris, 0.7 % NaCl, pH 8.0). 100 μ l of diluted matrigel solution was pipetted into each insert and allowed to gelatinize for 2 h in the incubator. To determine migration, ThinCerts were left uncoated. Each chamber was filled with 200 μ l serum free medium containing 1×10^5 cells. The lower compartment was filled with basal medium supplemented with 20 % FBS which acted as a chemoattractant. After incubation for 48 hours at 37 °C, 5 % CO₂, the cells and residual liquid on the upper surface of the membrane was carefully removed with a cotton swab. The inserts were transferred to a new 24-well plate containing 600 μ l of pre-warmed PBS (with Ca²⁺/Mg²⁺) per well. Staining was performed by the addition of 450 μ l of 8 μ M Calcein-AM solution (in pre-warmed, serum-free medium) per well and incubation for 45 min. Subsequently, the inserts were washed with PBS (37 °C) and transferred to a plate containing 500 μ l accutase per well. Finally, fluorescence was measured (excitation at 485 nm, emission at 520 nm) using a FLUOstar Omega - Multi-mode microplate reader (BMG Labtech).

Cell lysis

Prior to harvest, cells were washed two times with PBS. For (phospho) proteome preparation, cells were lysed in 8 M Urea, 40 mM Tris/HCl (pH 7.6), 1 x EDTA-free protease inhibitor mixture (complete mini, Roche) and 1 x Phosphatase inhibitor cocktail (Sigma). The lysate was centrifuged at 20,000 g for 45 min at 4 °C. For Kinobead experiments, cells were lysed in 1 x CP buffer (50 mM Tris-HCl, pH 7.5, 5 % glycerol, 1.5 mM MgCl₂, 150 mM NaCl) supplemented with 0.8 % NP-40, 1 mM DTT, 25 mM NaF and freshly added protease and phosphatase inhibitors (5 x phosphatase inhibitor cocktail 1, Sigma–Aldrich; 5 x phosphatase inhibitor cocktail 2, Sigma–Aldrich; 1 mM Na₃VO₄ and 20 nM calyculin A). Protein extracts were clarified by ultracentrifugation at 150,000 g for 1 h at 4 °C. Protein concentration for phosphoproteome and kinome samples was determined by the Bradford method (Coomassie (Bradford) Protein Assay Kit, Thermo Scientific) and the cleared lysates were stored at -80 °C until further use.

Digestion and dimethyl labeling for phospho- and full proteome preparation

The urea containing lysate was reduced with 10 mM DTT at 56 °C for 30 min and alkylated with 55 mM chloroacetamide for 30 min at room temperature in the dark. The protein mixture was diluted with 40 mM Tris/HCl to a final urea concentration of 1.6 M. Digestion was performed by adding sequencing grade Trypsin (Promega) in an enzyme-to-substrate ratio of 1:100 and incubation for 4 h at 37 °C. Subsequently, another 1:100 trypsin was added for overnight digestion at 37 °C. The next day, samples were acidified with TFA to a pH of 2 in order to stop trypsin activity. SepPak columns (C18 cartridges Sep-Pak Vac 1 cc (50 mg), Waters Corp., solvent A: 0.07 % TFA, solvent B: 0.07 % TFA, 50 % ACN) were used for peptide desalting according to manufacturer's instructions. Dimethyl labeling was performed on column as described previously (29). Briefly, after peptides were bound to the C18 material and washed with 4 ml of desalting solvent A, 5 ml light (10 % of 600 mM NaBH₃CN, 10 % of 4 % formaldehyde, 80 % of 50 mM sodium phosphate; all percentages are v/v), intermediate (10 % of 600 mM NaBH₃CN, 10 % of 4 % D₂-formaldehyde, 80 % of 5 mM sodium phosphate; all percentages are v/v) or heavy (10 % of 600 mM NaBD₃CN, 10 % of 4 % ¹³C-D₂-formaldehyde, 80 % of 5 mM sodium phosphate; all percentages as v/v) labeling solution was slowly passed through the cartridges. Subsequently, 8 ml of desalting solvent A was used to wash out remaining labeling reagent. Labeled peptides were eluted using 1 ml desalting solvent B, dried down and stored at -80 °C.

Fe-IMAC column enrichment and (phospho) peptide fractionation

Phosphopeptide enrichment was essentially performed as previously described (30). A Fe-IMAC column (ProPac IMAC-10 column, 4 mm I.D. x 50 mm, Thermo Fisher Scientific) was connected to an Aekta HPLC system and charged with Fe³⁺ ions using 25 mM FeCl₃ solution and equilibrated with IMAC solvent A (30 % ACN, 0.07 % TFA). Dimethyl-labeled peptides were reconstituted in IMAC solvent A, combined and loaded onto the column (10 min, 0.1 ml/min). The flow through which contains non-phosphorylated peptides was collected in a volume of 1.5 ml. Subsequently, phosphopeptides were eluted by a step-wise gradient from 0 – 12 % solvent B (0.3 % NH₄OH) in 5.3 min (0.6 ml/min) and from 12 % B to 24.75 % solvent B in 17 min (0.2 ml/min). Upon collection of the phosphopeptide elution peak (in a total volume of 1 ml), the column was flushed with 50 % solvent B (2.5 min; 1 ml/min) and re-equilibrated with solvent A (14 min; 1 ml/min). Both, the flow through and the elution fraction were dried down and stored at -80 °C.

High pH reversed-phase micro-column fractionation was performed in 200 µl pipette tips which were fixed in 1.5 ml Eppendorf tubes and packed with five C18 extraction disks (Ø 1.5 mm, 3M Empore) (31). All solvents were passed through the tip by centrifugation (~500 g). First, tips were primed using 25 µl of 100 % ACN, followed by 50 µl of 60 % ACN in 25 mM NH₄COOH (pH 10) and 2 x 50 µl of 25 mM NH₄COOH (pH 10). Next, dried phosphopeptides were re-dissolved in 100 µl of 25 mM NH₄COOH (pH 10) and loaded onto the C18 material. After re-application of the flow through, phosphopeptides were sequentially eluted using increasing concentrations of ACN (2.5 %, 7.5 %, 12.5 %, 60 % ACN in 25 mM NH₄COOH; 40 µl each). The flow through was combined with the 60 % ACN eluate, resulting in a total of four fractions which were dried down and stored at -20 °C. hSAX fractionation of 300 µg Fe-IMAC column flow through into 24 fractions was performed as described previously (32, 30).

Kinase affinity pull downs

Kinobead pull downs were conducted in a 96 well plate format as described previously (33). 35 µl settled beads (Kinobeads gamma, KBy) were washed twice with 1 ml of 1 x CP and then equilibrated with 1 ml 1 x CP/0.4 % NP-40. Residual liquid was removed via centrifugation (1200 rpm, 2 min, 4 °C). 1 ml lysate was diluted 1:1 resulting in a final concentration of 2 mg/ml protein and a reduction of the NP-40 concentration from 0.8 % to 0.4 %. The equilibrated beads were combined with 1 ml of the diluted lysate and incubated for 30 min at 4 °C on an end-over-end shaker. After removal of the lysate by centrifugation (1200 rpm, 2 min, 4 °C), the beads

were washed three times with 1 ml of 1 x CP/0.4 % NP-40 and twice with 1 ml 1 x CP/0.2 % NP-40 (1200 rpm, 2 min, 4 °C). To elute and reduce captured proteins, beads were incubated with 60 µl of 2 x NuPAGE LDS Sample buffer supplemented with 50 mM DTT (thermoshaker, 30 min, 50 °C, 700 rpm). The eluate was harvested via centrifugation (1200 rpm, 2 min, 4 °C) and one half was alkylated using 55 mM chloroacetamide (CAA) for 30 min at room temperature in the dark. Prior to MS analysis, samples were concentrated by SDS-PAGE and digested as previously described (34).

LC-MS/MS measurements

For full- and phosphoproteome fractions, nanoflow LC-MS/MS was performed by coupling an Agilent 1290 (Agilent technologies, Middelburg, Netherlands) to an Orbitrap Q Exactive Plus (Thermo Scientific, Bremen, Germany). Peptides were delivered to a trap column (100 µm I.D. x 2 cm, packed with 3 µm C18 resin, Reprosil PUR AQ, Dr. Maisch, Ammerbuch, Germany) at a flow rate of 5 µl/minute in 100 % loading solvent A (0.1 % FA, in HPLC grade water). After 10 min of loading and washing, peptides were transferred to an analytical column (75 µm I.D. x 40 cm C18 column Reprosil PUR AQ, 3 µm, Dr. Maisch, Ammerbuch, Germany) and separated using a 75 min (105 min for phospho) gradient and from 0 % to 40 % (36 % for phospho) solvent B (0.1 % FA in 80 % ACN) at a flow rate of 200 nl/min (solvent A: 0.1 % FA in HPLC grade water). Peptides were ionized using 1.9 kV spray voltage and a capillary temperature of 320 °C. The mass spectrometer was operated in data dependent acquisition mode, automatically switching between MS1 and MS2. Full scan MS spectra (m/z 375 – 1600) were acquired in the Orbitrap for a maximum of 250 ms (10 ms for phospho) at 35,000 (70,000 for phospho) resolution and an AGC target value of $3e6$. High resolution HCD-MS2 spectra were generated for up to 10 precursors with a normalized collision energy of 25 %. The underfill ratio was set to 1 % with a dynamic exclusion of 12 s (18 s for phospho). Fragment ions were acquired in the Orbitrap mass analyzer at a resolution of 17,500 (isolation window of 1.5 Th) and an AGC target value of $5e4$ with a maximum ion injection time of 120 ms.

For Kinobead eluates, nanoflow LC-MS/MS was performed by coupling an UltiMate 3000 nano LC system (Thermo Scientific, Bremen, Germany) to a Q Exactive HF (Thermo Scientific, Bremen, Germany). Peptides were delivered to a trap column (100 µm I.D. x 2 cm, packed with 5 µm C18 resin, Reprosil PUR AQ, Dr. Maisch, Ammerbuch, Germany) at a flow rate of 5 µl/minute in 0.1 % FA in HPLC grade water. After 10 minutes of loading and washing, peptides were transferred to an analytical column (75 µm I.D. x 40 cm C18 column Reprosil PUR AQ, 3 µm, Dr. Maisch, Ammerbuch, Germany) and separated using a 110 min gradient

from 0 % to 33 % solvent B (0.1 % FA, 5 % DMSO in ACN) at a flow rate of 300 nl/min (solvent A: 0.1 % FA, 5 % DMSO in HPLC grade water). Peptides were ionized using 2.2 kV spray voltage and a capillary temperature of 275 °C. The mass spectrometer was operated in data dependent acquisition mode, automatically switching between MS1 and MS2. Full scan MS spectra (m/z 360 – 1300) were acquired in the Orbitrap for a maximum of 10 ms at 60,000 resolution and an AGC target value of $3e6$. High resolution HCD MS2 spectra were generated for up to 12 precursors with a normalized collision energy of 25 %. The underfill ratio was set to 1 % with a dynamic exclusion of 30 s. Fragment ions were acquired in the Orbitrap mass analyzer at a resolution of 15,000 (isolation window of 1.7 Th) and an AGC target value of $2e5$ with a maximum ion injection time of 75 ms.

Data analysis

Data analysis was performed using MaxQuant v1.4.0.5 (35) and the integrated search engine Andromeda (36). For peptide and protein identification, raw files were searched against the UniProtKB database (v22.07.13, containing 88,381 entries) with carbamidomethylated cysteine as fixed modification and phosphorylation of serine, threonine and tyrosine, oxidation of methionine and N-terminal protein acetylation as variable modifications. For dimethyl labeling experiments, dimethyl-Lys0 and dimethyl-Nter0 were specified as the light, dimethyl-Lys4 and dimethyl-Nter4 as the medium and dimethyl-Lys8 and dimethyl-Nter8 as the heavy label. Trypsin/P was set as the proteolytic enzyme for which up to two missed cleavage sites were allowed. Precursor tolerance was set to 10 ppm and fragment ion tolerance to 20 ppm. Peptide identifications required a minimal length of seven amino acids and all datasets were adjusted to 1 % PSM and 1 % protein FDR. For kinase affinity enrichments, protein identifications are only reported, if they are based on three or more unique and/or razor peptides. Quantification was performed label-free for kinase affinity enrichments and dimethyl-based for phosphoproteome and full proteome fractions. To determine protein intensities in label-free experiments, the intensities of razor and unique peptides were summed up. Resulting protein intensities were then normalized to obtain LFQ intensities. The match-between-runs option, which aligns features of different mass spectrometric runs based on accurate retention time and mass, was enabled for label-free and dimethyl experiments. To facilitate further data analysis, the results were either imported into the MaxQuant associated software suite Perseus or into Excel (Microsoft). A two-sided t-test was used to assess statistical significance. Phosphopeptide and protein p -values were corrected for multiple testing using the method of Benjamini-Hochberg and allowing an FDR of 1 %. Phosphopeptides were filtered for a localization probability > 0.75

(class 1 sites) (37). Upregulated phosphorylation sites were further analyzed for enriched kinase substrate relationships using KinomeXplorer (required NetworkKIN score > 4, number of substrates > 1, top two kinases per phosphorylation site) (38). MotifX (39) was used to extract phosphorylation site motifs from the acquired dataset (significance = 0.001, width = 13). Phosphoprotein interactions were extracted using STRING (version 9.1; combined score > 0.9) (40) and KEGG was used for pathways enrichment analysis (41). Resulting networks were exported and visualized within the Cytoscape environment (version 3.1) (42). Phosphorylation site annotations were derived from PhosphoSitePlus™ (43). The software tool DAVID was used for functional annotation enrichment (44).

Results and discussion

Mass spectrometry-based workflow for global, multi proteomic profiling of lapatinib mode of action and resistance

The aim of this study is a global phosphoproteomic and proteomic characterization of lapatinib mode of action and resistance in an established (28), ERBB2 positive breast cancer cell line model. First, it was confirmed that the resistant cell line remains insensitive to lapatinib treatment (IC_{50} of 4.4 μ M for the resistant cells and 56 nM for the parental cells; Figure 1A), responsive to AXL inhibition (IC_{50} of 82 nM for the multi kinase inhibitor BMS-777607 which also targets AXL; Figure 1A) and has a proliferation rate comparable to the one of parental cells (Figure 1B). Since AXL expression was found to be directly dependent on the selective pressure of lapatinib, the growth medium of the resistant cell line was supplemented with 1 μ M lapatinib throughout the study (Figure 1C). For the mass spectrometry-based (phospho)proteomics analysis of parental cells which were either left untreated or exposed to 1 μ M lapatinib for 30 min versus the resistant cell line, triple dimethyl labeling of proteome digests in combination with Fe-IMAC column-based phosphopeptide enrichment was used (Figure 1D). In addition to the analysis of altered protein abundance, this setup enables a direct quantitative comparison of phosphorylation changes induced by lapatinib treatment of parental cells (mode of action analysis) to the corresponding rewiring events arising in resistance. To facilitate robust statistical analysis, a total of four biological replicates were conducted. PCA analysis of the twelve experimental states (three dimethyl channels in four replicates) revealed that the proteome and phosphoproteome samples cluster according to biology rather than technical batch and clearly shows a separation of the parental and the resistance state along the first PCA axis (Figure 1E). The quantitative reproducibility and dataset quality is further underscored by excellent correlation between biological replicates (average Pearson R of > 0.97 for phosphopeptides and > 0.98 for proteins; data not shown). To gain insight into complementary protein kinase changes not observed in the full proteome dataset, triplicate kinase affinity enrichment by Kinobeads (KBy) followed by quantification in a label-free fashion was performed (Figure 1C). Collectively, the dataset, which required close to eight days of measurement time, comprises the quantification of > 7,800 proteins, > 300 protein kinases and > 15,000 unique phosphopeptides (> 9,800 unique phosphorylation sites) which makes this, to the best of my knowledge, the most comprehensive and robust characterization of molecular events accompanying kinase inhibitor resistance to date.

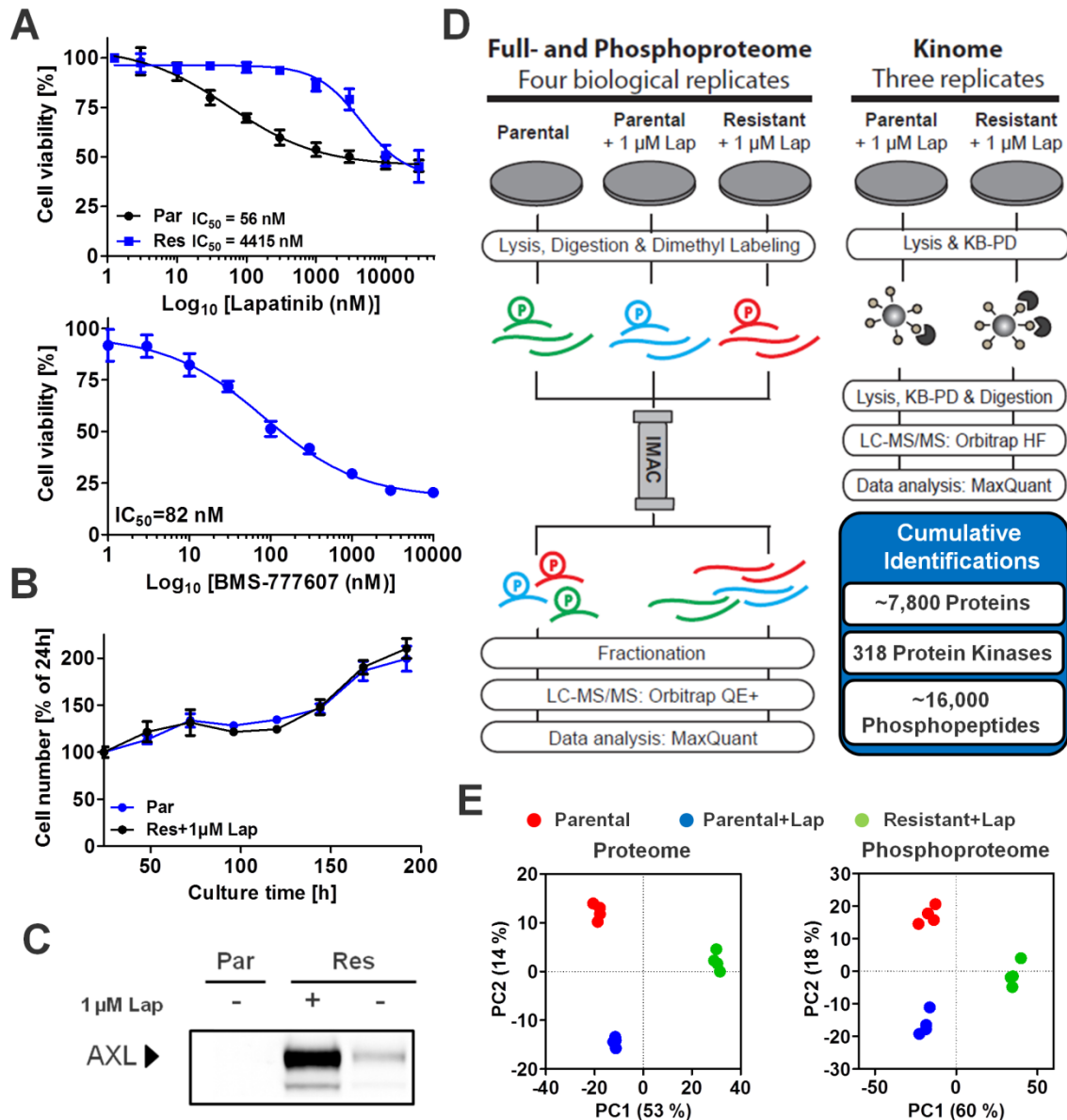


Figure 1. Workflow and dataset characterization. (A) Compared to parental BT-474 cells (Par), much higher doses of lapatinib are necessary to reduce viability of the resistant clone BT-474-J4 (Res). As shown by Liu et al. (28), the resistant cells also remain responsive to BMS-777607, a multi kinase inhibitor which targets AXL. (B) Proliferation of parental and resistant cell lines was monitored over the course of seven days. Despite the presence of 1 μM lapatinib, the growth rate of the resistant cell line is virtually identical to that of the parental line which is left without the inhibitor. (C) Western Blot analysis indicates that resistant cells which are cultured in the absence of lapatinib for several weeks lose the expression of AXL. (D) In the mass spectrometry-based (phospho) proteomic workflow used in this study, digested peptides from three different experimental conditions were dimethyl-labeled, combined and enriched for phosphopeptides using Fe-IMAC chromatography. To increase proteome and phosphoproteome coverage, the column flow through was separated using hydrophilic strong anion exchange chromatography and the phosphopeptide containing eluate was fractionated using high-pH reversed-phase micro-columns. The whole procedure was repeated in four independent biological replicates. In order to increase kinome coverage, Kinobead pull downs were conducted in three replicates. (E) PCA analysis of the three dimethyl encoded experimental states which were conducted in four biological replicates shows that samples cluster based on biology rather than dimethyl batch.

Phosphoproteomic analysis of lapatinib mode of action in parental cells

With this comprehensive dataset at hand, I first analyzed phosphorylation changes induced by treatment of the parental cell line with 1 μM lapatinib for 30 min. To assess statistical significance, a t-test with post hoc Benjamini-Hochberg correction (FDR < 1 %) was performed

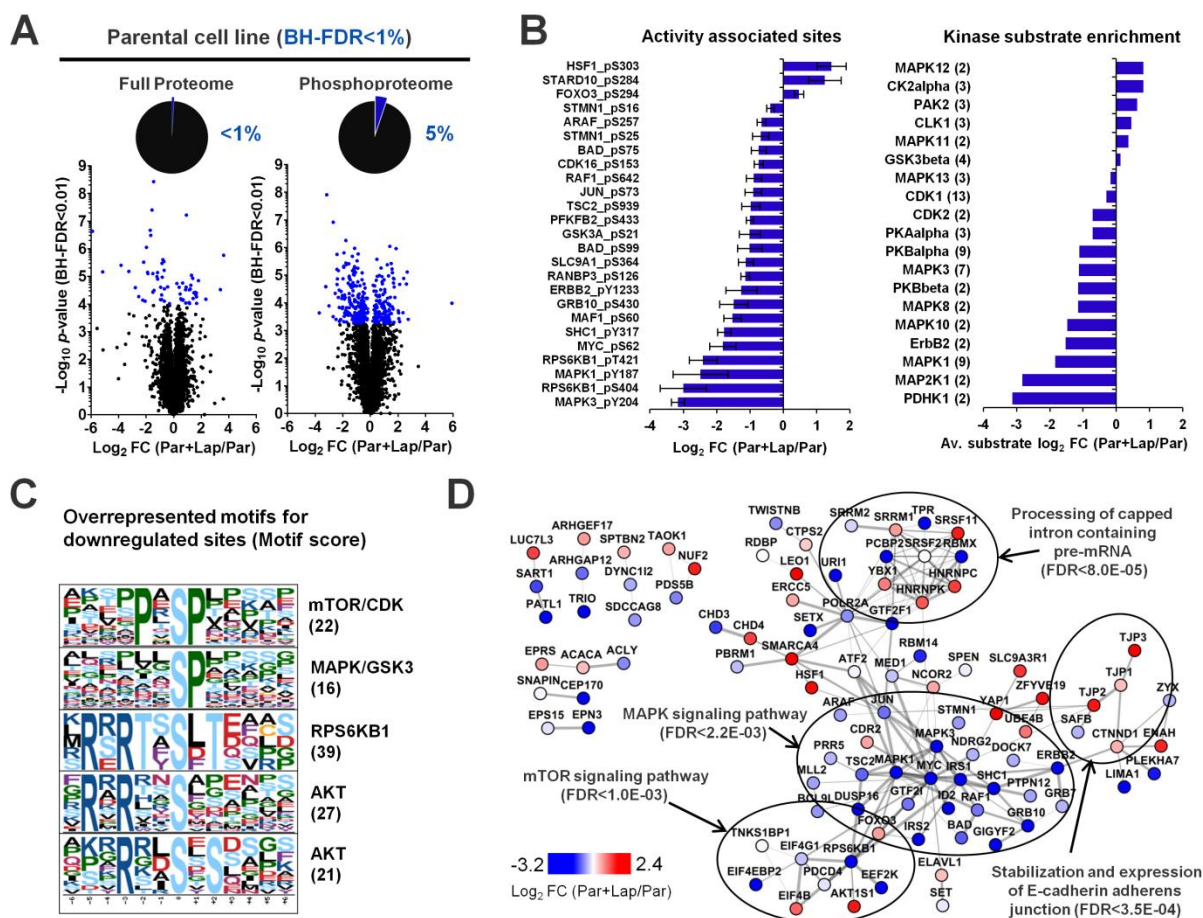


Figure 2. Phosphoproteomic analysis of lapatinib mode of action in parental cells. (A) Volcano plots showing significantly changing proteins and localized phosphosites (found in a minimum of three biological replicates; BH-FDR < 0.01, corrected for multiple testing) upon 30 min treatment with 1 μM lapatinib. Whereas less than 1 % of the proteome is significantly changing, 5 % of the phosphoproteome is perturbed. (B) Bar plots display the average \log_2 FC of sites which are known to have a functional impact on protein activity and the average \log_2 FC of kinase substrate sites. The number of associated substrates is indicated in brackets. (C) Phosphorylation motifs of many well-known kinases are enriched amongst the significantly inhibited phosphosites (p -value < 0.001). The motif-score is indicated in brackets. (D) Protein-protein interaction map of sites on phosphoproteins which are significantly changing upon short term lapatinib treatment. KEGG annotation (Cytoscape plugin) reveals significantly enriched pathways which confirms known (inactivation of MAPK, mTOR signaling) and uncovers new (mRNA processing and impact on adherens junctions) impact of lapatinib on cellular signaling.

on localized sites (localization probability > 0.75) which were at least seen in three out of four biological replicates. This analysis revealed that, despite the inhibitor's exquisite selectivity, a total of 349 sites (5 % of the dataset) are significantly changing upon lapatinib treatment (204 sites down and 145 sites up). As expected, less than one percent of the measured proteins are changing within the same period of time (Figure 2A). The 204 downregulated sites contain several known, ERBB2 pathway associated kinase phosphorylation sites such as ERBB2-pY-1233 (\log_2 FC of -2.9; p -value of 3.6E-4), MAPK3-pY-204 (\log_2 FC -3.2; p -value of 1.2E-08) or RPS6KB1-pS-404 (\log_2 FC of -3.0; p -value of 1.3E-4) which all directly reflect the kinase's activity (Figure 2B). Both, the analysis of enriched kinase substrates (Figure 2B) and the overrepresented motifs among the inhibited phosphorylation sites (Figure 2C) largely corroborate those kinase activity changes (e.g. for MAPK1, MAPK3, RPS6KB1 and ERBB2) and suggest the inhibition of additional kinases (e.g. AKT, mTOR). Global protein interaction analysis of changing phosphoproteins using the STRING database in combination with KEGG annotation of the extracted network confirms the known perturbation of the MAPK (e.g. < 2.2E-3) and mTOR signaling branch (FDR < 2.2E-3). Moreover, it reveals a previously unknown and, surprisingly even more significant, impact of the inhibitor on the stabilization and expression of E-cadherin adherens junctions (FDR < 3.5E-4) and on the processing of capped intron containing pre-mRNA (FDR < 8.0E-5) which represents an important part of mRNA splicing (Figure 2D). Interestingly, splicing is implicated in virtually all steps of tumor biology and frequently acquires oncogenic potential (45). Hence, the spliceosome is increasingly recognized as druggable tumor target (46) and splicostatins (or analogues thereof) inhibit parental BT-474 cells with low nanomolar affinity (46, 47). Impaired splicing might thus be a previously unappreciated mechanism of lapatinib action in breast cancer. In addition, phosphoproteomic analysis also points to the regulation of receptor tyrosine kinase adaptor proteins such as IRS1/2, SHC1 or GAB2 and the altered activity of central transcription factors JUN and MYC (Figure 2B).

Resistance acquisition is accompanied by extensive reprogramming of the proteome, kinome and phosphoproteome

It is currently largely unknown to which extent resistance acquisition alters the global molecular repertoire of a cell compared to its parental counterpart. Towards this end, I stringently analyzed the proportion of significant changes (t-test with post hoc BH-FDR adjustment to < 1 %) of the phosphoproteome between parental and resistant cells (Figure 3A) and found that 1816 localized phosphorylation sites change significantly (841 up and 976 down). This corresponds to

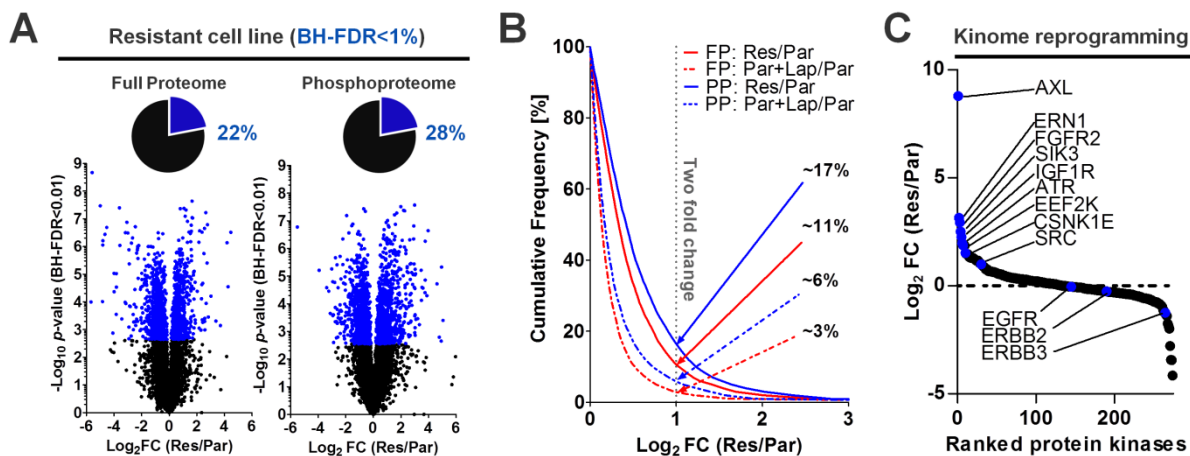


Figure 3. Global molecular reprogramming of a resistant cell line. (A) Volcano plots depict significantly changing proteins and localized phosphosites (found in a minimum of three biological replicates; BH-FDR < 1 %) in the resistant (Res) compared to the parental (Par) cell line. The magnitude of phosphoproteome (28 % of all sites) and proteome (22 % of all proteins) alterations suggest that resistance is accompanied by a fundamental impact on the proteomic composition of a tumor cell. (B) The cumulative frequency plot shows the quantitative magnitude of protein and phosphorylation changes upon resistance acquisition in relation to the changes observed after 30 min lapatinib treatment of the parental cell line. (C) Averaged quantitative kinase data from the full proteome and the affinity purification dataset were ranked according to their \log_2 FC between the parental and the resistant cell line.

over one quarter of all covered sites, which is surprisingly large in light of the relatively moderate change elicited by short term treatment with the very selective inhibitor lapatinib (5 % of all sites; Figure 2A). Also 22 % (1422) of all proteins change significantly (655 up and 768 down) (Figure 3A). From a strictly quantitative point of view, 17 % of all sites and 11 % of all proteins are altered at least two fold which further underscores the magnitude of molecular reprogramming upon resistance acquisition (Figure 3B). As kinases are frequently involved in the resistant phenotype and contain a high proportion of clinically actionable targets, kinome perturbations were of special interest. In addition to the known overexpression of AXL, this analysis indicates the reprogramming of several other kinases which have not yet been associated with lapatinib resistance (Figure 3C). Some of them have known or emerging roles in breast cancer biology (e.g. ERN1 (49), FGRF2 (50), ATR (51) and EEF2K (51, 52), with \log_2 fold changes of 3.1, 2.9, 2.0 and 1.9) whereas others are entirely undefined in this context (e.g. SIK3, \log_2 FC 2.5). Thus, these extensive kinome changes prioritize a pool of biologically relevant kinases, of which several are accessible with current pharmacological matter.

Phosphorylation rewiring in resistance extends the mechanism of PIK3 reactivation and uncovers many de novo activation events

After having established the global extent of acquired resistance, I next asked how phosphorylation is changing in resistance. To study signaling recovery, the 204 sites which are initially inhibited by lapatinib in the parental cell line were arbitrarily divided into the categories “remain inhibited” (112 sites) and “reactivated” (92 sites) by requiring at least a 50 % phosphorylation recovery compared to the initial phosphorylation change (e.g. a site which was originally inhibited with a \log_2 FC of -2 would be called reactivated if the \log_2 FC in resistance is > -1). Extracted pathway information of connected phosphoproteins clearly shows that the “reactivation” network is strongly enriched in sites associated with the KEGG annotated PIK3/AKT/mTOR pathway. In contrast, not a single phosphoprotein with such an annotation remains inhibited (Figure 4A). Some selected, activity associated sites and kinase substrates shown in Figure 4B corroborate this finding. Whereas the activity of kinases belonging to the MEK/ERK signaling pathway (e.g. MAP2K1 and MAPK1) remains inhibited, the activity of those from the PIK3/AKT/mTOR pathway (e.g. AKT or RPS6KB1) fully recovers in lapatinib resistant compared to sensitive cells. Hence, as one would expect, the phosphoproteomic data readily captured the previously described recovery of PIK3/AKT/mTOR signaling and adds many new observations (e.g. the transcription factor JUN is reactivated whereas MYC remains inhibited). In addition, the data suggests a detailed mechanism for adaptor/scaffolding protein rewiring downstream of the altered receptor tyrosine kinases. First, strong, ERBB2 independent (the site is not responsive to 30 min lapatinib treatment) phosphorylation of the receptor tyrosine kinase adaptor GAB2 at position pY-476 (\log_2 FC of 2.9) is capable of inducing interaction with PIK3R1 (which is also phosphorylated at pY-580 with a \log_2 FC of 2.9), the regulatory subunit of PIK3 (54). Second, despite their initial sensitivity towards lapatinib, several sites of the adaptor protein IRS1 are reactivated in resistance (e.g. pT-446, pT-453 and pS527). The protein IRS1 serves as a docking site and scaffold for many different SH2 domain containing proteins, including the p85 subunit of PIK3, and might in turn play an important role for signal rewiring. In contrast, the adaptor proteins IRS2 (pS-1176) and SHC1 (pY-427) remain highly responsive to lapatinib inhibition also in the resistant cell line.

Next, I focused on sites and kinase substrates which are upregulated in resistance compared to the untreated parental cell line and found a clear activation of several kinases known to interact with each other (SRC, PAK1/4, RPS6KA1/3, JAK2, CK 1/2, CDK1) and altered phosphorylation of transcription factors located downstream of some of those kinases (e.g. RPS6KA family is

known to phosphorylate STAT3-pS727 (55)) (Figure 4C). Excessive activation of phosphorylation sites located on proteins belonging to the initially targeted pathway (mTOR and ERBB signaling as the top two KEGG pathways; Figure 4D) further suggests that the cell tries to surmount inhibitor action by excessive activation of initially targeted pathways. Strikingly, this is also apparent for sites belonging to the spliceosome (Figure 4D), which has already been found to be affected in response to short term lapatinib treatment (Figure 2D).

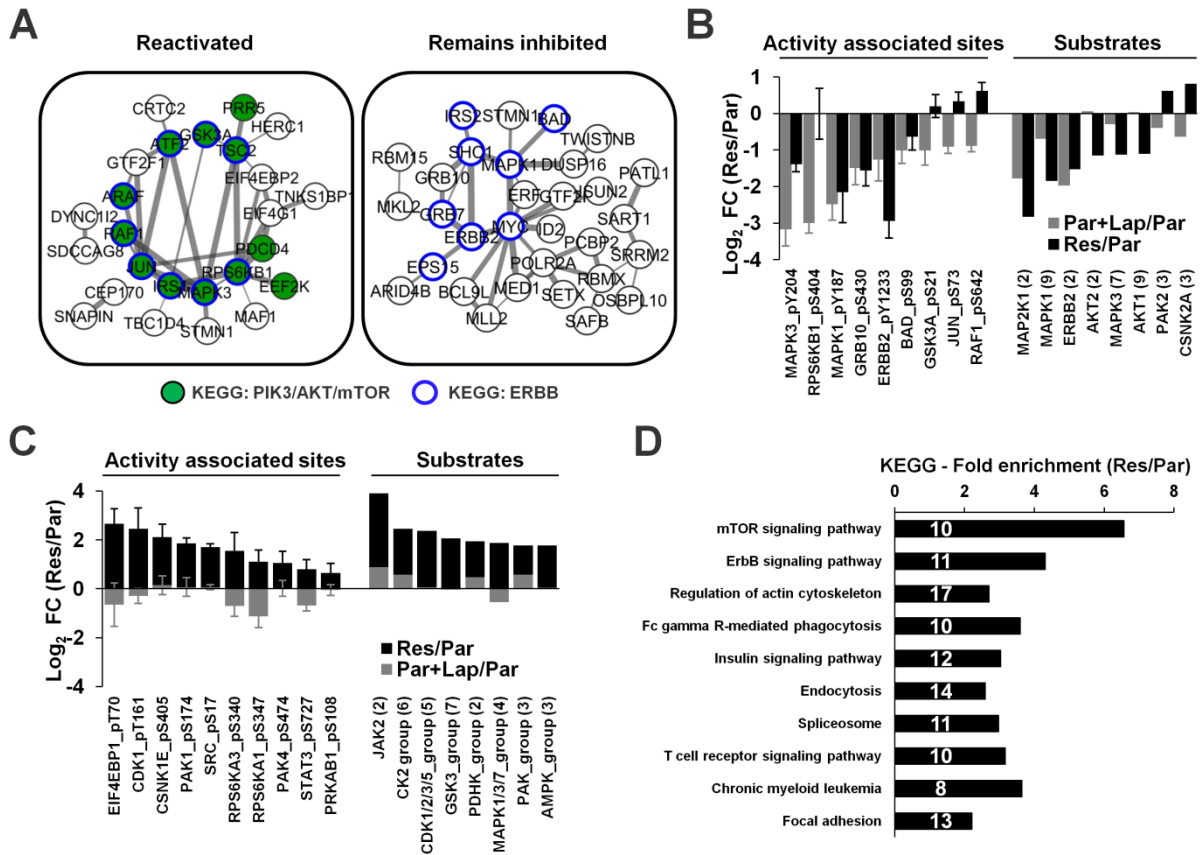


Figure 4. Phosphorylation changes in resistance. (A) An arbitrary classification in reactivated sites (112 sites) and such that remain largely inhibited (95 sites) in resistant cells reveals that sites on proteins which are part of the PIK3/AKT/mTOR pathway (annotated using KEGG) are exclusively found reactivated. (B) The lapatinib induced \log_2 FC of selected activity associated sites and kinase substrate sites is plotted for the parental (Par) and the resistant (Res) cell line. (C) A selection of activity associated sites and kinase substrates which are seen reproducibly upregulated in resistance and do not show statistically significant inhibition upon short term lapatinib treatment. For kinase substrate enrichment, the average \log_2 FC of the substrate sites is plotted and the number of associated substrates is indicated in brackets. (G) Global analysis of KEGG pathways enriched among phosphoproteins which are upregulated in resistance. Numbers inside the bars indicate the associated number of (phospho) proteins.

A multi proteomic model of resistance reveals a wide range of deregulated drug targets

As shown above, the multilayered dataset paints a complementary and comprehensive picture of resistance acquisition which is summarized in the pathway model shown in Figure 5A. It features kinases, transcription factors and scaffolding proteins which were found to be reactivated (e.g. PIK3, AKT, JUN, IRS1), remain inhibited (e.g. MEK, MYC, SHC1) or are activated/overexpressed (e.g. CDK1, STAT3, GAB2) after long term lapatinib treatment. Notably, all acquired datasets provide unique information which further emphasizes the value of a multi proteomic perspective. Next, it was of interest if those newly discovered resistance features can be pharmacologically exploited. In contrast to continuously inhibited sites, sites that are initially responsive to treatment and reactivated in resistance should have a higher chance of being functionally relevant for survival. Indeed, Liu *et al.* showed that the resistant cells are highly sensitive towards AKT, mTOR and PIK3 inhibition (28), whereas no reduction in viability upon dose dependent treatment with the MEK inhibitor selumetinib was found (data not shown). The proteomic model also captured many additional alterations and a prominent one is the strong activation of CDK1 (pT-161, log₂ FC 2.5) and moderate activation of CDK2 (pY-15, log₂ FC 0.6). Strikingly, both the parental and the resistant cell line are highly sensitive towards BMS-387032 (IC₅₀ parental ~200 nM, resistant ~300 nM; Figure 5B) and SCH-727965 (IC₅₀

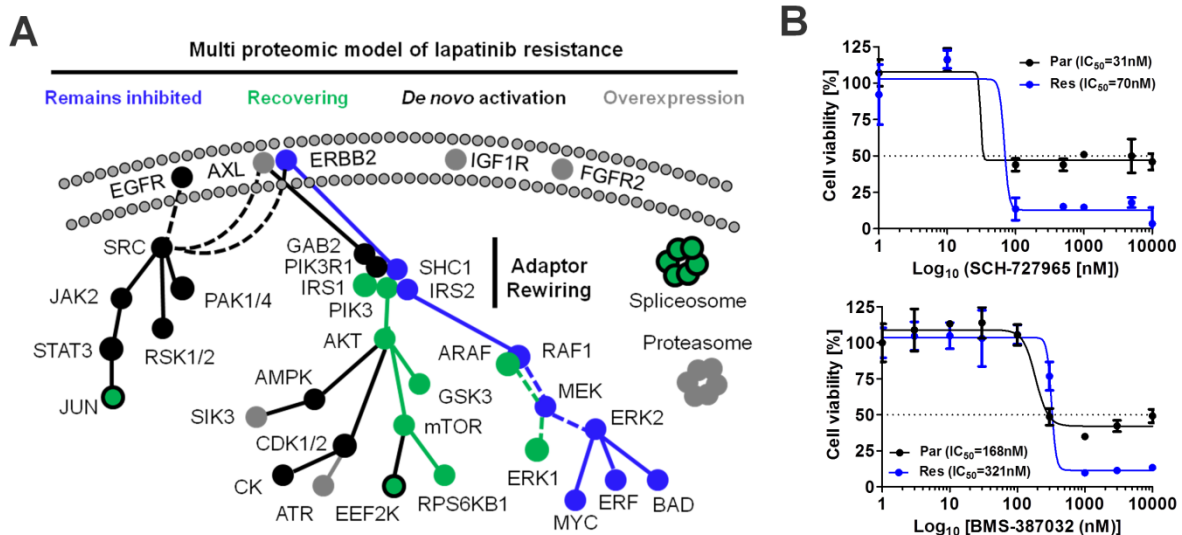


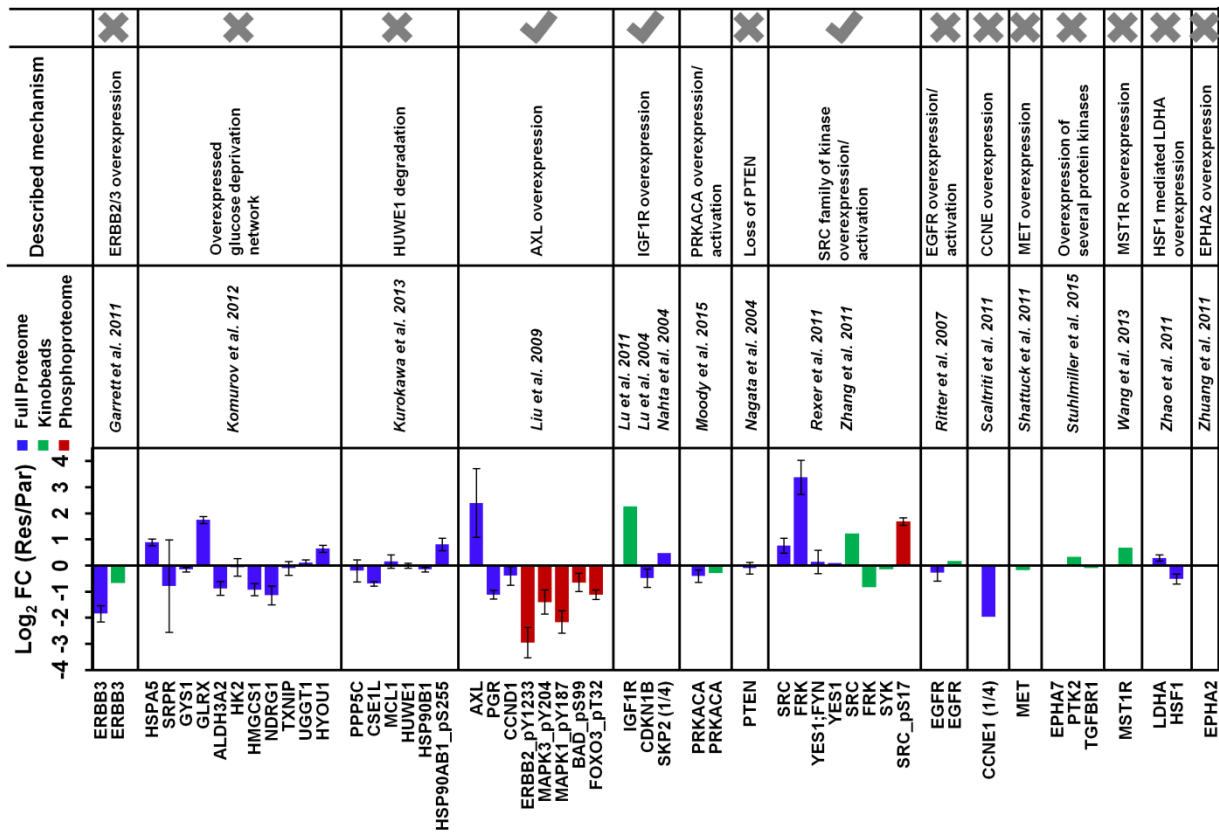
Figure 5. Identification of pharmacologically actionable targets and phenotypes in lapatinib resistant breast cancer (A) A pathway model of overexpression, activation and rewiring events in resistance. Phosphoproteomic analysis identified nodes which remain responsive to lapatinib also in resistance (blue), nodes that are reactivated and might thus re-establish proliferation (green) and nodes which are de novo upregulated in resistance (black). (B) Dose-response curves using two CDK inhibitors (BMS-387032, SCH-727965) show that viability of the resistant cells is highly dependent on CDK signaling.

parental ~30 nM, resistant ~70 nM; Figure 5B), two inhibitors of the CDK family. The fact that roughly 45 % of the parental cells survive CDK inhibition indicates that a higher proportion of the resistant cell line population is dependent on CDK1/2 signaling. Another example is the kinase EE2FK, which is an emerging target in certain breast cancer subtypes and can be pharmacologically addressed (56). It contains two sites which are responsive to lapatinib treatment in parental cells (pS-72, pS-74; \log_2 FC -1.7, -1.3; p -values of 2.7E-04 and 3.7E-05) and is found to be highly overexpressed in resistance (see Figure 3C). In effect this leads to the full recovery of absolute phosphorylation levels (pS-72 and pS-74, \log_2 FC (Res/Par) 0.07 and 0.05). Intriguingly, this observation might imply that the cell compensates inhibitor elicited loss of phosphorylation by protein overexpression. Of note, also 38 different ribosomal components are overexpressed in resistant cells, which might be a means of counterbalancing lapatinib induced mitigation of translation (which is a known effect of mTOR inhibitor). Owing to time constraints, no further studies directed towards examining the functional role of EE2FK and other prioritized targets (e.g. CK, SIK3 or JAK2) in resistance could be carried out but corresponding experiments are currently ongoing. In conclusion, the multi proteomics model of resistance made it apparent that resistant cells acquire many different and targetable alterations.

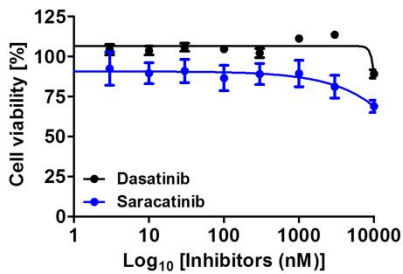
A comparison to previously described mechanisms of ERBB2 resistance suggest very diverse molecular paths of acquisition

Motivated by the fact that many different activation and overexpression events exist, I next asked if some common alterations can be found. Fortunately, resistance against ERBB2 inhibition in breast cancer is one of the best studied model systems and many different routes leading to resistance have been uncovered. Strikingly, the comprehensive dataset contained quantitative information for virtually all of the molecules and phosphorylation events described in 18 different studies (Figure 6A). As expected, large-scale proteomic measurement readily confirmed Western Blot-based observations from the laboratory which established the resistant cell line which was also used for this study (28) (loss of PGR expression and ERK pY-204, BAD pS-99 and FOXO pT-32 phosphorylation in resistance; Figure 6A). However, alterations found in other studies, such as the compensatory upregulation of ERBB family members (ERRB3 (10), EGFR (9)), loss of PTEN (23) or the overexpression of CCNE (22) and PRKACA (16) were not identified. Except for AXL and IGF1R, there was also no evidence for previously described upregulation of different receptor tyrosine kinases (MET (14), MST1R (15), EPHA2 (11), EPHA7, PTK2, TGFBR1 (57)). Interestingly and despite nearly five fold overexpression, IGF1R

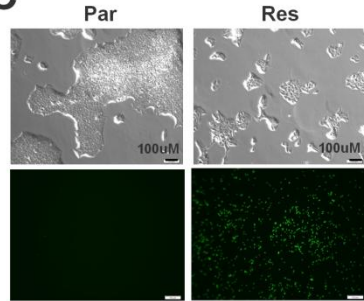
A



B



C



D

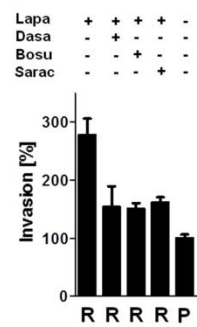


Figure 6. Comparison of previously described mechanisms of resistance against ERBB2 targeted therapy to the dataset acquired in this study. (A) 18 different molecular mechanisms of resistance were compiled from literature and compared against the quantitative protein (blue bars), phosphorylation (red bars) and kinase (green bars) measurements obtained by exploratory mass spectrometry. The described molecular mechanism of resistance is indicated above (a tick indicates confirmation and a cross indicates discordance) (B) Dose-response curves of two SRC family of kinase inhibitors dasatinib and saracatinib shows that viability of the resistant cells is not dependent on SRC signaling. (C) Left panel: Resistant cells (Res/R) are morphologically distinct compared to the parental cell line (Par/P). Right panel: Matrigel invasion assay shows that resistant cells are more invasive than their parental counterpart. Notably, the invasive phenotype is particularly pronounced upon lapatinib removal. (D) In contrast to proliferation, reduced invasive potential upon SRC family of kinases inhibition with dasatinib (250 nM), saracatinib (250 nM) or bosutinib (300 nM) provides evidence that SRC signaling is important for invasion in the cell line model used for his study.

was not required for proliferation (cells are not responsive to the IGF1R inhibitor linsitinib; data not shown). This is in line with the absence of significant CDKN1B and SKP2 expression changes, which were previously shown to be important for IGF1R mediated resistance (12). Consequently it can be argued that the plain observation of an activation or expression event in resistance might not necessarily be a good indicator of functional relevance.

Independent studies in BT-474 cells have previously identified SRC and members of the SRC family of kinases as common drivers of lapatinib resistance, which are required for sustained proliferation (18, 17). As SRC activation (and FRK overexpression in the Kinobead dataset; Figure 6A) was also observed in BT-474-J4 cells, I intended to reproduce these findings, but, surprisingly, found that resistant cells are not responsive to the SRC family of kinase inhibitors dasatinib and saracatinib (Figure 6B). However, due to the observed morphological changes of the resistant cell line (Figure 6C), the enrichment of the KEGG pathway “regulation of actin cytoskeleton” (Figure 4D) among upregulated phosphoproteins and the known role of the *de novo* activated SRC pathway in invasion (58), it was investigated if the cell line might have an altered invasive behavior. Indeed, Figure 6C shows that invasion in the resistant cell line is increased by roughly three fold compared to the parental cell line. This acquired phenotype can be reduced by low doses of the SRC family of kinase inhibitors dasatinib, saracatinib and bosutinib (Figure 5D). This result corroborates conclusions drawn for IGF1R (i.e. functional relevance is context specific) and further suggest that previously identified resistance drivers can have distinct functional roles in each individual case (e.g. instead of being relevant for proliferation in the model system used in this study, SRC overexpression/activation is responsible for an invasive phenotype). Interestingly, the removal of lapatinib increases the invasiveness to roughly six fold (Figure 6C), which calls discontinued inhibitor exposure upon the development of resistance into question. Mechanistically, SRC might directly phosphorylate and activate EGFR/ERBB2 on pY-727/pY-735, a site which is unresponsive to lapatinib treatment, known to interact with SRC (59) and highly upregulated in resistance (\log_2 FC of 2.9; notably it is only found in one out of four replicates). As soon as lapatinib is removed, EGFR might get reactivated and can thus contribute to enhanced invasion.

Collectively, the results show that the mechanisms of resistance acquisition are very heterogeneous and that previously identified resistance drivers might either not be functionally relevant (e.g. IGF1R) or have several, context specific functions (e.g. SRC family of kinases). This heterogeneity might pose a great challenge in terms of unifying treatment options. Thus, it might be a more efficient strategy to search for molecular integrators of different, heterogeneous resistance mechanisms or common alterations.

Lapatinib resistance is accompanied by a phosphorylation-based shift towards increased glucose consumption

Two other described molecular mechanisms of resistance are based on glycolytic addiction caused by metabolic reprogramming (Figure 6A). In the first case, overexpression of twelve different proteins was found to activate a glucose deprivation response which rendered resistant cells sensitive towards glycolysis inhibition (26). Although significant overexpression of HSPA5 (p -value of $1.7E-06$) and the glucose transporter GLRX (p -value of $1.1E-05$) was observed, there was no such evidence for the remaining nine proteins (one was not covered by the acquired dataset; Figure 6A). In fact, abundance of three of those proteins even significantly decreased in the resistant cell line (ALDH3A2, HMGCS1 and NDRG1). A second mechanism involves upregulation of the transcription factor HSF1 which was found to increase LDHA expression, a metabolic protein that enhances glycolytic flux (60). Whereas LDHA overexpression was only marginally elevated in my cell line (\log_2 FC of 0.3, p -value of $1.5E-03$), I found a 62-fold increase of LDAH pY-10 phosphorylation, a site which is known to cause

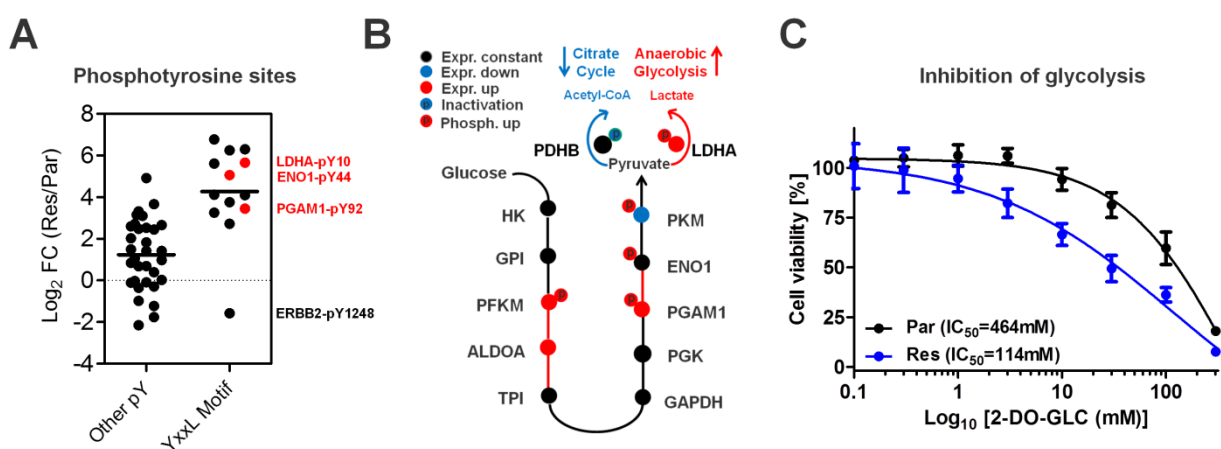


Figure 7. Lapatinib resistant cells are addicted to glycolysis which represent a targetable phenotype (A) The \log_2 FC of phosphotyrosine sites between parental and resistant cells is plotted separately for YXXL-motif containing proteins and such that do not contain this motif (phosphorylation changes are corrected for altered protein expression). The clear quantitative separation suggests that one activated kinase or deactivated phosphatase is responsible for these changes (the outlier YXXL motif site is ERBB2 pY-1248 which is inhibited due to lapatinib treatment). Importantly, also three pY-sites from major glycolytic enzymes (LDHA, ENO1, PGAM1) carry the YXXL motif (red colored dots). (B) A metabolic pathway model summarizing expression and phosphorylation changes detected in this study. Metabolic rewiring of resistant cells is primarily driven by phosphorylation. Posttranslational activation of LDHA via pY-10 and inactivation of PDHA1 via pS-300, promotes conversion of lactate to pyruvate rather than acetyl-CoA, which in turn enhances anaerobic glycolysis. (C) Cell viability assay where resistant and parental cells are treated with increasing doses of 2-deoxy-glucose, an inhibitor of glycolysis, for 96 h. Compared to parental cells, lapatinib resistant cells show an increased addiction to glucose.

enzymatic activation (61) (p -value of $8.1E-03$; Figure 7A). Also ENO1 pY-44 and PGAM1-pY92, phosphotyrosine sites on two other important glycolytic enzymes, showed a 31- and 13-fold increase, respectively (p -value of $4.7E-06$ and $4.0E-05$; Figure 7A). In search for the responsible tyrosine kinase/phosphatase, an overrepresented phosphotyrosine motif was discovered (YXXL, where X stands for an arbitrary amino acid and L for the amino acid leucine). Sites that contain this motif are quantitatively separated from the rest (except for one site which represents an inhibited ERBB2 site; see Figure 7A) and comprise all the pY sites found on glycolytic enzymes (Figure 7A). Hence, it is very likely that one specific kinase/phosphatase mediates for posttranslational activation of glycolysis, although it is at present unclear which one this is. Efficient conversion of pyruvate to lactate can additionally be enhanced by reduction of alternative pyruvate metabolism. Indeed, the phosphorylation on the enzymatic activity inhibiting site pS-300 on pyruvate dehydrogenase alpha (PDHA1) (62), the enzyme which irreversibly converts pyruvate to acetyl-CoA and thus prepares it for its entry into the citric acid cycle, is 27 fold increased in resistance (Figure 7B). Consequently, anaerobic pyruvate conversion by concerted, posttranslational activation of LDHA and inactivation of PDHA1, should strongly promote the Warburg effect (i.e. predominant energy production via anaerobic glycolysis which is energetically much less favorable) and concomitant glycolytic addiction. In line with this model, the resistant cells are much more sensitive to glycolysis inhibition by 2-deoxy glucose (Figure 7C). Taken together, the study uncovers a posttranslational activation/de-activation of key enzymes within the glycolytic pathway which seems to be a viable route towards metabolic reprogramming and glucose addiction in resistance. The fact that an antibody against pY10, a biomarker for LDHA activity, exists and that glycolysis inhibitors such as 2-deoxy glucose or the anti diabetic drug metformin are actively evaluated in clinical trials renders the treatment of this phenotype a viable therapeutic option.

Conclusion

In this study a large scale, multi proteomic analysis of the mechanism of lapatinib action in parental and resistant cells was conducted. The mass spectrometry-based approach readily captured and vastly extended the previously described mechanism of AXL mediated resistance and signaling recovery. Moreover, the combination of proteomics, phosphoproteomics and chemoproteomics uncovered additional, pharmacologically targetable phenotypes. Specifically, the data suggests that glycolytic addiction of resistant cells might be a common characteristic which can be targeted in a clinical setting (e.g. by 2-deoxy-glucose or metformin). The fact that this alteration cannot necessarily be observed by expression changes emphasizes the importance to study posttranslational modifications in general and phosphorylation in particular. Based on data from other studies and my own analysis, there might be great merit in searching for unifying features of resistance rather than solely relying on the identification of potentially interchangeable molecular events which might converge on the same molecule, pathway or phenotype (e.g. several resistance mechanisms observed in ERBB2 overexpressing breast cancer, whether it is AXL, EPHA2 or MET, lead to the reactivation of PIK3, which might consequently be a more attractive node to target). Given the many laboratories working on very similar models of resistance, the findings in this study hopefully encourage researchers to test those for the existence of previously described resistance drivers and especially for potential molecular integrators or convergence points. This should ultimately result in a more robust management and treatment of resistance.

Acknowledgements

I would like to thank Mrs. Andrea Hubauer for excellent technical assistance with preparation of proteomic samples and conduction of invasion and proliferation assays. Moreover, I wish to thank Dr. Simone Lemeer for mass spectrometry measurements.

Abbreviations

FC	Fold change
FDR	False discovery rate
hSAX	hydrophilic strong anion exchange chromatography
I.D.	Inner diameter
IMAC	Immobilized metal ion affinity chromatography
KBy	Kinobeads gamma
KEGG	Kyoto Encyclopedia of Genes and Genomes
Lap	Lapatinib
LC	Liquid chromatography
LFQ	Label free quantification
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
Par	Parental cell line
PCA	Principal component analysis
ppm	Parts per million
PSM	Peptide spectrum match
p-Y/S/T	Phosphotyrosine, -serine, -threonine
Res	Resistant cell line

References

1. Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., and McGuire, W. L. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235, 177–182
2. Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., and Ullrich, A. (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244, 707–712
3. Fiore, P. D., Pierce, J. H., Kraus, M. H., Segatto, O., King, C. R., and Aaronson, S. A. (1987) erbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. *Science* 237, 178–182
4. Cobleigh, M. A., Vogel, C. L., Tripathy, D., Robert, N. J., Scholl, S., Fehrenbacher, L., Wolter, J. M., Paton, V., Shak, S., Lieberman, G., and Slamon, D. J. (1999) Multinational Study of the Efficacy and Safety of Humanized Anti-HER2 Monoclonal Antibody in Women Who Have HER2-Overexpressing Metastatic Breast Cancer That Has Progressed After Chemotherapy for Metastatic Disease. *J. Clin. Oncol.* 17, 2639–2639
5. Hudziak, R. M., Lewis, G. D., Winget, M., Fendly, B. M., Shepard, H. M., and Ullrich, A. (1989) p185HER2 monoclonal antibody has antiproliferative effects in vitro and sensitizes human breast tumor cells to tumor necrosis factor. *Mol. Cell. Biol.* 9, 1165–1172
6. Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J., and Norton, L. (2001) Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *N. Engl. J. Med.* 344, 783–792
7. Xia, W., Mullin, R. J., Keith, B. R., Liu, L.-H., Ma, H., Rusnak, D. W., Owens, G., Alligood, K. J., and Spector, N. L. (2002) Anti-tumor activity of GW572016: a dual tyrosine kinase inhibitor blocks EGF activation of EGFR/erbB2 and downstream Erk1/2 and AKT pathways. *Oncogene* 21, 6255–6263
8. Geyer, C. E., Forster, J., Lindquist, D., Chan, S., Romieu, C. G., Pienkowski, T., Jagiello-Gruszfeld, A., Crown, J., Chan, A., Kaufman, B., Skarlos, D., Campone, M., Davidson, N., Berger, M., Oliva, C., Rubin, S. D., Stein, S., and Cameron, D. (2006) Lapatinib plus Capecitabine for HER2-Positive Advanced Breast Cancer. *N. Engl. J. Med.* 355, 2733–2743
9. Ritter, C. A., Perez-Torres, M., Rinehart, C., Guix, M., Dugger, T., Engelman, J. A., and Arteaga, C. L. (2007) Human Breast Cancer Cells Selected for Resistance to Trastuzumab In vivo Overexpress Epidermal Growth Factor Receptor and ErbB Ligands and Remain Dependent on the ErbB Receptor Network. *Clin. Cancer Res.* 13, 4909–4919
10. Garrett, J. T., Olivares, M. G., Rinehart, C., Granja-Ingram, N. D., Sánchez, V., Chakrabarty, A., Dave, B., Cook, R. S., Pao, W., McKinely, E., Manning, H. C., Chang, J., and Arteaga, C. L. (2011) Transcriptional and posttranslational up-regulation of HER3 (ErbB3) compensates for inhibition of the HER2 tyrosine kinase. *Proc. Natl. Acad. Sci.* 108, 5021–5026
11. Zhuang, G., Brantley-Sieders, D. M., Vaught, D., Yu, J., Xie, L., Wells, S., Jackson, D., Muraoka-Cook, R., Arteaga, C., and Chen, J. (2010) Elevation of receptor tyrosine kinase EphA2 mediates resistance to trastuzumab therapy. *Cancer Res.* 70, 299–308
12. Lu, Y., Zi, X., Zhao, Y., Mascarenhas, D., and Pollak, M. (2001) Insulin-Like Growth Factor-I Receptor Signaling and Resistance to Trastuzumab (Herceptin). *J. Natl. Cancer Inst.* 93, 1852–1857
13. Nahta, R., Yuan, L. X. H., Zhang, B., Kobayashi, R., and Esteva, F. J. (2005) Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. *Cancer Res.* 65, 11118–11128
14. Shattuck, D. L., Miller, J. K., Carraway, K. L., and Sweeney, C. (2008) Met Receptor Contributes to Trastuzumab Resistance of Her2-Overexpressing Breast Cancer Cells. *Cancer Res.* 68, 1471–1477
15. Wang, Q., Quan, H., Zhao, J., Xie, C., Wang, L., and Lou, L. (2013) RON confers lapatinib resistance in HER2-positive breast cancer cells. *Cancer Lett.* 340, 43–50
16. Moody, S. E., Schinzel, A. C., Singh, S., Izzo, F., Strickland, M. R., Luo, L., Thomas, S. R., Boehm, J. S., Kim, S. Y., Wang, Z. C., and Hahn, W. C. (2015) PRKACA mediates resistance to

- HER2-targeted therapy in breast cancer cells and restores anti-apoptotic signaling. *Oncogene* 34, 2061–2071
17. Rexer, B. N., Ham, A.-J. L., Rinehart, C., Hill, S., Granja-Ingram, N. de M., González-Angulo, A. M., Mills, G. B., Dave, B., Chang, J. C., Liebler, D. C., and Arteaga, C. L. (2011) Phosphoproteomic mass spectrometry profiling links Src family kinases to escape from HER2 tyrosine kinase inhibition. *Oncogene* 30, 4163–4174
 18. Zhang, S., Huang, W.-C., Li, P., Guo, H., Poh, S.-B., Brady, S. W., Xiong, Y., Tseng, L.-M., Li, S.-H., Ding, Z., Sahin, A. A., Esteva, F. J., Hortobagyi, G. N., and Yu, D. (2011) Combating trastuzumab resistance by targeting SRC, a common node downstream of multiple resistance pathways. *Nat. Med.* 17, 461–469
 19. De Luca, A., D'Alessio, A., Gallo, M., Maiello, M. R., Bode, A. M., and Normanno, N. (2013) Src and CXCR4 are involved in the invasiveness of breast cancer cells with acquired resistance to lapatinib. *Cell Cycle Georget. Tex* 13,
 20. Campbell, I. G., Russell, S. E., Choong, D. Y. H., Montgomery, K. G., Ciavarella, M. L., Hooi, C. S. F., Cristiano, B. E., Pearson, R. B., and Phillips, W. A. (2004) Mutation of the PIK3CA Gene in Ovarian and Breast Cancer. *Cancer Res.* 64, 7678–7681
 21. Saal, L. H., Holm, K., Maurer, M., Memeo, L., Su, T., Wang, X., Yu, J. S., Malmström, P.-O., Mansukhani, M., Enoksson, J., Hibshoosh, H., Borg, Å., and Parsons, R. (2005) PIK3CA Mutations Correlate with Hormone Receptors, Node Metastasis, and ERBB2, and Are Mutually Exclusive with PTEN Loss in Human Breast Carcinoma. *Cancer Res.* 65, 2554–2559
 22. Scaltriti, M., Eichhorn, P. J., Cortés, J., Prudkin, L., Aura, C., Jiménez, J., Chandarlapaty, S., Serra, V., Prat, A., Ibrahim, Y. H., Guzmán, M., Gili, M., Rodríguez, O., Rodríguez, S., Pérez, J., Green, S. R., Mai, S., Rosen, N., Hudis, C., and Baselga, J. (2011) Cyclin E amplification/overexpression is a mechanism of trastuzumab resistance in HER2+ breast cancer patients. *Proc. Natl. Acad. Sci.* 108, 3761–3766
 23. Nagata, Y., Lan, K.-H., Zhou, X., Tan, M., Esteva, F. J., Sahin, A. A., Klos, K. S., Li, P., Monia, B. P., Nguyen, N. T., Hortobagyi, G. N., Hung, M.-C., and Yu, D. (2004) PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* 6, 117–127
 24. Nahta, R., Takahashi, T., Ueno, N. T., Hung, M.-C., and Esteva, F. J. (2004) P27kip1 Down-Regulation Is Associated with Trastuzumab Resistance in Breast Cancer Cells. *Cancer Res.* 64, 3981–3986
 25. Zhao, Y., Liu, H., Liu, Z., Ding, Y., LeDoux, S. P., Wilson, G. L., Voellmy, R., Lin, Y., Lin, W., Nahta, R., Liu, B., Fodstad, O., Chen, J., Wu, Y., Price, J. E., and Tan, M. (2011) Overcoming Trastuzumab Resistance in Breast Cancer by Targeting Dysregulated Glucose Metabolism. *Cancer Res.* 71, 4585–4597
 26. Komurov, K., Tseng, J.-T., Muller, M., Seviour, E. G., Moss, T. J., Yang, L., Nagrath, D., and Ram, P. T. (2012) The glucose-deprivation network counteracts lapatinib-induced toxicity in resistant ErbB2-positive breast cancer cells. *Mol. Syst. Biol.* 8,
 27. Xia, W., Bacus, S., Hegde, P., Husain, I., Strum, J., Liu, L., Paulazzo, G., Lyass, L., Trusk, P., Hill, J., Harris, J., and Spector, N. L. (2006) A model of acquired autoresistance to a potent ErbB2 tyrosine kinase inhibitor and a therapeutic strategy to prevent its onset in breast cancer. *Proc. Natl. Acad. Sci.* 103, 7795–7800
 28. Liu, L., Greger, J., Shi, H., Liu, Y., Greshock, J., Annan, R., Halsey, W., Sathe, G. M., Martin, A.-M., and Gilmer, T. M. (2009) Novel mechanism of lapatinib resistance in HER2-positive breast tumor cells: activation of AXL. *Cancer Res.* 69, 6871–6878
 29. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. R. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* 4, 484–494
 30. Ruprecht, B., Koch, H., Medard, G., Mundt, M., Kuster, B., and Lemeer, S. (2015) Comprehensive and Reproducible Phosphopeptide Enrichment Using Iron Immobilized Metal Ion Affinity Chromatography (Fe-IMAC) Columns. *Mol. Cell. Proteomics* 14, 205–215
 31. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896–1906
 32. Ritorto, M. S., Cook, K., Tyagi, K., Pedrioli, P. G. A., and Trost, M. (2013) Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *J. Proteome Res.* 12, 2449–2457

33. Médard, G., Pachl, F., Ruprecht, B., Klaeger, S., Heinzlmeir, S., Helm, D., Qiao, H., Ku, X., Wilhelm, M., Kuehne, T., Wu, Z., Dittmann, A., Hopf, C., Kramer, K., and Kuster, B. (2015) Optimized Chemical Proteomics Assay for Kinase Inhibitor Profiling. *J. Proteome Res.* 14, 1574–1586
34. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* 68, 850–858
35. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372
36. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* 10, 1794–1805
37. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S., and Mann, M. (2010) Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* 3, ra3
38. Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., Palma, A., Cesareni, G., Jensen, L. J., and Linding, R. (2014) KinomeXplorer: an integrated platform for kinome biology studies. *Nat. Methods* 11, 603–604
39. Schwartz, D., and Gygi, S. P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* 23, 1391–1398
40. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., Mering, C. von, and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808–D815
41. Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30
42. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504
43. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270
44. Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4, P3
45. Oltean, S., and Bates, D. O. (2014) Hallmarks of alternative splicing in cancer. *Oncogene* 33, 5311–5318
46. van Alphen, R. J., Wiemer, E. a. C., Burger, H., and Eskens, F. a. L. M. (2008) The spliceosome as target for anticancer treatment. *Br. J. Cancer* 100, 228–232
47. Eustáquio, A. S., Janso, J. E., Ratnayake, A. S., O'Donnell, C. J., and Koehn, F. E. (2014) Spliceostatin hemiketal biosynthesis in *Burkholderia* spp. is catalyzed by an iron/ α -ketoglutarate-dependent dioxygenase. *Proc. Natl. Acad. Sci. U. S. A.* 111, E3376–E3385
48. He, H., Ratnayake, A. S., Janso, J. E., He, M., Yang, H. Y., Loganzo, F., Shor, B., O'Donnell, C. J., and Koehn, F. E. (2014) Cytotoxic Spliceostatins from *Burkholderia* sp. and Their Semisynthetic Analogues. *J. Nat. Prod.* 77, 1864–1870
49. Rajapaksa, G., Nikolos, F., Bado, I., Clarke, R., Gustafsson, J.-Å., and Thomas, C. (2015) ER β decreases breast cancer cell survival by regulating the IRE1/XBP-1 pathway. *Oncogene* 34, 4130–4141
50. Wei, W., Liu, W., Serra, S., Asa, S. L., and Ezzat, S. (2015) The breast cancer susceptibility FGFR2 provides an alternate mode of HER2 activation. *Oncogene*,
51. Abdel-Fatah, T. M. A., Middleton, F. K., Arora, A., Agarwal, D., Chen, T., Moseley, P. M., Perry, C., Doherty, R., Chan, S., Green, A. R., Rakha, E., Ball, G., Ellis, I. O., Curtin, N. J., and Madhusudan, S. (2015) Untangling the ATR-CHEK1 network for prognostication, prediction and therapeutic target validation in breast cancer. *Mol. Oncol.* 9, 569–585
52. Meric-Bernstam, F., Chen, H., Akcakanat, A., Do, K.-A., Lluch, A., Hennessy, B. T., Hortobagyi, G. N., Mills, G. B., and Gonzalez-Angulo, A. (2012) Aberrations in translational regulation are

- associated with poor prognosis in hormone receptor-positive breast cancer. *Breast Cancer Res. BCR* 14, R138
53. Yao, Z., Li, J., Liu, Z., Zheng, L., Fan, N., Zhang, Y., Jia, N., Lv, J., Liu, N., Zhu, X., Du, J., Lv, C., Xie, F., Liu, Y., Wang, X., Fei, Z., and Gao, C. (2016) Integrative bioinformatics and proteomics-based discovery of an eEF2K inhibitor (cefatrizine) with ER stress modulation in breast cancer cells. *Mol. Biosyst.*,
 54. Crouin, C., Arnaud, M., Gesbert, F., Camonis, J., and Bertoglio, J. (2001) A yeast two-hybrid study of human p97/Gab2 interactions with its SH2 domain-containing binding partners. *FEBS Lett.* 495, 148–153
 55. Zhang, Y., Cho, Y.-Y., Petersen, B. L., Bode, A. M., Zhu, F., and Dong, Z. (2003) Ataxia telangiectasia mutated proteins, MAPKs, and RSK2 are involved in the phosphorylation of STAT3. *J. Biol. Chem.* 278, 12650–12659
 56. Russnes, H. G., and Caldas, C. (2014) eEF2K—a new target in breast cancers with combined inactivation of p53 and PTEN. *EMBO Mol. Med.* 6, 1512–1514
 57. Stuhlmiller, T. J., Miller, S. M., Zawistowski, J. S., Nakamura, K., Beltran, A. S., Duncan, J. S., Angus, S. P., Collins, K. A. L., Granger, D. A., Reuther, R. A., Graves, L. M., Gomez, S. M., Kuan, P.-F., Parker, J. S., Chen, X., Sciaky, N., Carey, L. A., Earp, H. S., Jin, J., and Johnson, G. L. (2015) Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep.* 11, 390–404
 58. Guarino, M. (2010) Src signaling in cancer invasion. *J. Cell. Physiol.* 223, 14–26
 59. Thelemann, A., Petti, F., Griffin, G., Iwata, K., Hunt, T., Settinar, T., Fenyó, D., Gibson, N., and Haley, J. D. (2005) Phosphotyrosine signaling networks in epidermal growth factor receptor overexpressing squamous carcinoma cells. *Mol. Cell. Proteomics MCP* 4, 356–376
 60. Zhao, Y. H., Zhou, M., Liu, H., Ding, Y., Khong, H. T., Yu, D., Fodstad, O., and Tan, M. (2009) Upregulation of lactate dehydrogenase A by ErbB2 through heat shock factor 1 promotes breast cancer cell glycolysis and growth. *Oncogene* 28, 3689–3701
 61. Fan, J., Hitosugi, T., Chung, T.-W., Xie, J., Ge, Q., Gu, T.-L., Polakiewicz, R. D., Chen, G. Z., Boggon, T. J., Lonial, S., Khuri, F. R., Kang, S., and Chen, J. (2011) Tyrosine phosphorylation of lactate dehydrogenase A is important for NADH/NAD(+) redox homeostasis in cancer cells. *Mol. Cell. Biol.* 31, 4938–4950
 62. Korotchikina, L. G., and Patel, M. S. (2001) Probing the mechanism of inactivation of human pyruvate dehydrogenase by phosphorylation of three sites. *J. Biol. Chem.* 276, 5731–5738

Chapter 6

General discussion

General discussion

Mass spectrometry has emerged as the key technology for the explorative study of phosphorylation events. Advances in sample preparation, phosphopeptide enrichment and mass spectrometric instrumentation now enables the routine identification and quantification of 10,000s of sites in a single study. Still, several challenges such as missing workflow reproducibility, absence of automation capabilities, enrichment material complementarity, and incomplete sequence coverage remain. Hence, one central aim of this thesis was to develop, implement and benchmark an efficient and robust phosphoproteomic workflow which addresses these limitations.

In chapter 2, I described the development of a purification approach based on chromatographic phosphopeptide enrichment using commercially available IMAC columns charged with Fe^{3+} ions. I found that Fe-IMAC columns enable comprehensive enrichment without biases introduced by insufficient capacity. Importantly, bound phosphopeptides are also efficiently eluted off the material, a fact that is hard to address for materials where the metal ions and thus the bound analyte molecules, cannot be easily stripped of the carrier material. The chromatographic approach with its multiple interaction events vastly improves enrichment selectivity and thus overcomes co-purification of non phosphorylated peptides as a major drawback of previous Fe-IMAC formats (1). With this enrichment approach at hand, I initially aimed at sequential depletion of phosphopeptides by means of different enrichment materials (TiO_2 and Ti-IMAC) which have previously been shown to purify distinct parts of the phosphoproteome (2). Surprisingly however, missing recovery of phosphopeptides by other materials (and also upon repeated Fe-IMAC column enrichment) from Fe-IMAC column flow throughs clearly showed that the IMAC column completely depleted the sample of phosphopeptides which was not the case *vice versa*. Hence, the developed workflow is *per se* comprehensive and does neither require additional enrichment materials nor consecutive enrichment steps. Moreover, this suggests that the complementarity which has previously been observed is only of apparent nature and is primarily caused by biases introduced by insufficient capacity and inefficient elution. Although the initially intended fractionation of phosphopeptides (i.e. according to the number of phosphorylation sites per peptide or other factors influencing affinity to the immobilized metal ion) did not succeed using the Fe-IMAC column approach, it turned out to be highly beneficial for enriching phosphopeptides prior to separation, a concept which saves a tremendous amount of time and is leveraged by high capacity and comprehensiveness (3). Such pre-enrichments also pave the way for high throughput

applications which would be extremely cumbersome to achieve if phosphopeptides are enriched from each fraction separately. In concert with devices such as an autosampler or a fraction collector, and a routine means for quality control (recorded absorbance), the whole process can be efficiently automatized. Given the recent debate about reproducibility (4, 5, 6), the column approach also offers the potential to decrease the inter-experimental and inter-laboratory variation. The commercial availability and the chromatographic principle of the column format should increase uniformity across laboratories and reduce the variation introduced by the manual assembly of spin columns or the synthesis of functionalized beads. Another advantage is that Fe-IMAC columns enrichment employs relatively mild buffer conditions and does not require selectivity enhancing additives which supports phosphopeptide integrity (e.g. phosphoester hydrolysis) and avoids additional desalting steps. It also facilitates re-application of IMAC column flow throughs for the enrichment of multiple PTMs, which would be a very appealing workflow extension that should be easy to implement (7). However, it should also be noted that bead- and tip-based approaches are certainly advantageous if a laboratory does not have access to a HPLC device. The increased sensitivity of such tip-based approaches is also beneficial when it comes to the enrichment of limited amounts of sample (e.g. below 100 µg). In this respect the column is less effective which might make it less applicable to the enrichment of clinical samples. Moreover, my own experience shows that reproducibility of tip and bead formats is not considerably worse if all experiments are performed within one batch.

Work done in this thesis exclusively relied on HCD fragmentation for the identification of phosphopeptides. Improvement of phosphopeptide localization by means of alternative fragmentation methods or hybrids thereof is an active area of research which might hold potential for the future. Promising approaches are EThcD (8) or the recently published UV induced dissociation (9). In addition, there is room for substantial improvement on the computational side of site assignment with a special focus on a more efficient interpretation of fragmentation patterns and a concomitant advance in the scoring of site localization (10).

The last four years also saw several other phosphoproteomic workflow combinations published. For example, Loroche *et al.* compiled a workflow based on solid-phase extraction and electrostatic repulsion-hydrophilic interaction chromatography which offered vastly increased sensitivity and is in turn suitable for low µg input material (11). In addition, in another study, the Gerber laboratory found that chemical tagging of phosphopeptide eluates obtained after single stage enrichment (e.g. by TMT) does not compromise robustness (12). Consequently, less of the expensive labeling reagent is required. The recently published “EasyPhos” workflow extends the approach introduced by Kettenbach *et al.* (3) and features sample processing in a 96-well

plate format which enhances throughput and increases sample recovery due to more infrequent changes of reaction vessels and smaller reaction volumes (13). An enhanced reproducibility facilitates label-free quantification of enriched phosphopeptides in a single measurement approach akin to what I have described in chapter 2. Clear advantages of label-free methods are the large number of samples that can be compared and their applicability to serum and tissue. An emerging alternative is the multiplexed quantification of TMT10-plex labeled phosphopeptides which offers the advantage of less missing values, increases throughput tremendously but also suffers from severe ratio compression (9, 10). This can, in principle, be solved by a recently introduced MS3 approach (16). However, judging from my own experience, also MS3 spectra are not entirely free of ratio compression which renders the call of absence problematic (a feature might be entirely absent but shows a quantitative signal due to co-isolation of a peptide species with similar mass eluting at the same time). To mitigate this problem, one might want to confine MS2 or MS3-based quantification approaches to the study of relative or gradual changes (e.g. upon stimulation of signaling for different periods of time) within the same biological system (e.g. one cell line). In contrast, the higher dynamic range and more robust estimation of absence which is offered by MS1-based quantification approaches might be more suited to study heterogeneous systems such as different cell lines or tissues.

During the time of writing, several improvements which are related to previously mentioned advantages of the Fe-IMAC column workflow and were not described in chapter 2 have been implemented. First, the Fe-IMAC column enrichment time was reduced by a factor of eight (15 min turnaround instead of the 120 min as described in chapter 2). The time for column recharging was decreased to roughly 1 h (with further room for improvement) with a tested maximum of twelve consecutive enrichments (this number could well be higher, but this remains to be verified). Second, I have combined the enrichment with TMT labeling (pre and post column) and an efficient and sensitive high pH reversed-phase micro-column fractionation. In concert, these workflow extensions patch some previous shortcomings, add versatility and increase the sample throughput tremendously. To ensure efficient dissemination to other users and, perhaps even more importantly to other laboratories, the finalized workflow has been summarized in a step-by-step protocol which will be made more broadly available in the near future. Not least because of these efforts, several other groups have successfully implemented the methodology I have developed in this thesis. This shows that the workflow is easy to adapt and means that the idea of increased inter-laboratory reproducibility of phosphoproteomic studies might not just remain a theoretical one. Akin to a recent comparability study (17), a next step might be the analysis and variability assessment of phosphopeptide enrichments

conducted in different laboratories. Importantly, the described phosphoproteomic workflow is now also an integral part of various studies conducted in our laboratory. Examples include the phosphorylation site analysis of dozens of cell lines, tissues and distinct species such as mice and plants. Also relative changes, such as cell line specific phosphoproteome perturbations triggered by inhibiting or stimulating agents are now routinely conducted at a competitive depth and with high throughput. The workflow has also proven to be compatible with all common labeling and quantification approaches. Here, the simplicity and different modules which can be flexibly combined and individually tailored depending on the question at hand are particularly beneficial. This versatility is important since no one streamlined workflow will be sufficient to address every experimental question at hand.

Despite impressive technological improvements and a reported quantification of >50,000 phosphopeptides in a single cell line (18), the coverage of the phosphoproteome is far from complete. In chapter 2, I have shown that different enrichment materials are not *per se* suited to study different parts of the phosphoproteome. But even if only of apparent nature, this orthogonality which is caused by a combination of format shortcomings and insufficient mass spectrometric sequencing speed can in principle be used to isolate distinct parts of the phosphoproteome. Yet, a more efficient means to boost sequence coverage is the unbiased enrichment of phosphopeptides followed by multidimensional peptide separation. Given the superior retention of negatively charged groups, hSAX has proven to be particularly powerful for this purpose. However, other studies have nicely shown that also other fractionation techniques are highly suited for phosphoproteomic studies (e.g. high-pH reversed-phase columns (19), HILIC (20) or SCX (21)). The initial dataset from chapter 2 was acquired on a moderately performing Orbitrap Velos. In light of the vast advances in mass spectrometric instrumentation and the concomitant increase in sequencing speed it would be exciting to see which depth can be achieved nowadays (22). As bottom-up proteomics workflows are very much streamlined in the sense that they almost exclusively rely on trypsin for protein digestion and nESI for peptide ionization, chapter 3 examines how alterations in these workflow steps might enhance phosphoproteomic coverage. The study predominantly focuses on the complementarity of nESI and MALDI ionization for phosphoproteomics and describes the global characteristics of preferentially detected phosphopeptides. The comparison is based on thousands of phosphorylation events and as a consequence extends and strengthens observations made in previous studies, where samples of low complexity or artificially synthesized peptides were used. While the fundamental and global characteristics of differentially detected

phosphospecies are noteworthy and interesting, MALDI-based workflows are unfortunately not implemented in many laboratories and thus not broadly applicable. Still, the combination of phosphotyrosine immuno-affinity enrichment and MALDI-MS/MS, which favors the detection of tyrosine phosphorylated peptides, might be useful given the pivotal biological role of phosphotyrosine signaling. Moreover, we know from our own experience that the application of MALDI MS/MS can be crucial for the identification of phosphorylation sites of biological significance which are otherwise inaccessible (23). In chapter 3, I also showed that MALDI increases the sequence coverage equally as effective as orthogonal digestion enzymes such as Asp-N or Glu-C. As the latter approach is easy to implement in routine LC-MS/MS workflows and is readily compatible with Fe-IMAC column enrichment, it might represent a more straightforward way to increase phosphosite coverage. Another widespread method for the detection of tyrosine phosphorylated peptides is the enrichment using commercially available, phosphotyrosine specific antibodies. Typically, 10ths of milligrams of proteins amounts are required to get high numbers of identifications, confirming that the modification is of very low abundance. A more targeted approach recently described in literature is the use of motif-specific antibodies (24, 25, 26). Their application has led to the identification of kinase substrates that were not yet identified in large-scale studies (26), as well as to the identification of phosphopeptides containing a specific substrate consensus motif, previously not associated with a specific kinase (24).

A major motivation of phosphoproteomic studies is to infer on cellular activity which often carries more functional relevance than plain protein abundance. Since all phosphorylation events (at least as far as we know) stem from kinases, the analysis of substrate changes is only an indirect indicator of such activities. In search for a more direct method which might complement signaling studies, I examined to which extent chemoproteomic kinase enrichment is influenced by the kinase activation state. The work described in chapter 4 was mainly motivated by repeated claims from another laboratory which established and published (27) the concept that immobilized inhibitor bead mixes exclusively capture the active kinase conformation. Hence, I activated kinases using phosphatase inhibition and compared their chemoproteomic binding to phosphorylation site changes which were measured using the previously established Fe-IMAC column workflow. Unfortunately for the field, these claims did not hold true and I was able to show that this matter is much more complex than previously anticipated. Whereas only a minor fraction is fished in one preferential conformation, the binding of vast majority of kinases is not altered upon activation. This chapter came with two important scientific lessons. First,

confirmation of experimental results is an important part of the scientific process. Sometimes false conclusion drawn from observations are stated as a fact, generalized and used as a sound basis for further research. Such self propagation can only be addressed by independent research. Second, it is considerably easier to criticize a concept than to provide or at least suggest an adequate alternative, solution or extension. To address the latter I am currently working on the enrichment of phosphopeptides from pre-enriched kinome samples using the Fe-IMAC column (an idea that is based on previous work (11,12)). The in depth analysis of the (phospho) kinome renders the combined presence of the phosphopeptides and its non-phosphorylated counterpart more likely. Consequently, absolute stoichiometries (so-called phosphorylation site occupancies (30) of a large number of kinase activation loop sites can be calculated. Although this approach is not as direct as initially aimed for, it is a facile means for a deeper interrogation of kinase activity and should be further explored in the future. Still, for the few kinase inhibitor and bead combinations where activity dependent binding was found to be true, kinase activity can also be determined directly. Moreover, our laboratory is now examining the concept of activity independent binding for a possibility to profile small molecule kinase inhibitors against distinct activation states. According to the results of a recent study, the identification of such conformation selective inhibitors might for instance be effective in reversing some cases of resistance (31).

Explorative phosphoproteomics can be used to answer a multitude of different biological questions. It provides insights into very fundamental cellular biology (e.g. the cell cycle (12, 13)), allows to infer on kinase activity (14, 15); it enables to record temporal signaling dynamics (34), to classify cell lines according to their origin (35), it can be used to predict phenotypic response (36), catalogue the phosphorylation events in different tissue types (37) and to study a drug's mode of action (38). Up to chapter 4 my thesis had a very technical character. Since new methodology always lives from application, chapter 5 describes the phospho- and chemoproteomic dissection of lapatinib resistance in ERBB2 overexpressing breast cancer. The study combined methodology which was examined and established in chapters 2 and 4 of this thesis and showed that Fe-IMAC column workflow is robustly applicable to address biological questions. Given the preparation of four biological replicates, the decreased amount of time spent for sample preparation and the high quantitative reproducibility were especially apparent. The enrichment has proven to be efficient and the peptide absorbance provided a facile means of quality control. Despite the increased complexity of MS1 spectra (triple dimethyl labeling), the combination with a simple tip fractionation approach (six fractions) rendered the coverage of the

phosphoproteome deep enough to identify several key players within the ERBB signaling pathway. Still, as already pointed out above, future efforts should clearly be directed towards an improved coverage of the phosphoproteome. Of note, the hSAX approach, which has been described in chapter 2, was confirmed to be powerful for the orthogonal and high resolution separation of non phosphorylated peptides obtained from Fe-IMAC column flow throughs. The systems level view on signaling recovery and molecular deregulation upon the acquisition of resistance allowed me to gain new mechanistic insight and to find new ways of targeting this phenotype. It turned out that resistance, even against a very selective compound, is characterized by extensive molecular changes and many more vulnerabilities than previously anticipated. Overall, this study highlights one of the general strengths of discovery proteomics which is to create new insights from a systems level perspective and to tie together previous observations in an effort to gain new insights. In line with this strength, the recorded dataset contained quantitative data for virtually all previously described mechanisms of resistance against ERBB2 targeted therapy. This data revealed the enormous heterogeneity in resistance acquisition but also suggested common convergence points of rewiring which might hold potential in terms of unifying treatment options. With addition to glycolysis as a proposed common trait, it would certainly be exciting to examine if the other existing models of ERBB2 resistance in breast cancer corroborate this finding. Targeting glycolysis is especially promising given the availability of suited therapeutic agents such as the antidiabetic drug metformin or 2-deoxy-glucose which are currently evaluated in clinical trials.

Outlook

Developments in the field of proteomics have made the in-depth analysis of protein phosphorylation feasible within a reasonable amount of time. However, it is still completely unclear how much of the phosphoproteome we need to cover in order to comprehensively analyze sites of biological relevance, let alone how many phosphorylation sites are occurring within one single biological system. In the next years, developments in both enrichment methods and mass spectrometers will be aimed at bringing phosphoproteomic comprehensiveness within reach. From a technical point of view, important next steps have to be the routine integration of several different proteases for digestion, improved sensitivity as sample amounts are often limited, integration of consecutive PTM enrichments and an increase in throughput as well as multiplexing capabilities. Apart from purely technical advances, progress in the functional annotation of the identified phosphorylation sites is crucial. A facile first step might be the identification and quantification of the “responsive” phosphoproteome. This could prioritize sites which are actively changing upon a variety of external perturbations and in turn provide valuable biological information. Moreover, these perturbations would reveal tightly connected phosphorylation sites and provide insights into the hierarchical composition of cellular signaling. Ultimately this should lead to frameworks of individual and dynamic cellular pathway maps that extend or replace “traditional” static models. Finally, from a data management and analysis point of view, stringent standards in data reporting and experimental design will increase inter-laboratory comparability and the establishment of data repositories. Such repositories should not only catalogue datasets but make the information accessible for exploration and use it as a foundation for functionally annotated, higher level insight. Moreover, new analysis tools that incorporate PTM data should put the analyzed data in a biological perspective thereby simplifying systems level research.

Abbreviations

ESI	Electro-spray ionization
ETD	Electron transfer dissociation
EThcD	HCD supplemented ETD
HCD	Higher energy C-trap dissociation
HILIC	Hydrophilic interaction chromatography
HPLC	High-performance liquid chromatography
hSAX	Hydrophilic strong anion exchange
IMAC	Immobilized metal affinity chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS1	Precursor mass spectrum
MS2	Fragment mass spectrum
MS/MS	Tandem mass spectrometry
MS3	Isolated fragment ion spectrum
PTM	Post-translational modification
SCX	Strong cation exchange
TiO ₂	Titaniumdioxide
TMT	Tandem mass tag
LC	Liquid chromatography
UV	Ultraviolette

References

1. Dunn, J. D., Watson, J. T., and Bruening, M. L. (2006) Detection of Phosphopeptides Using Fe(III)–Nitrilotriacetate Complexes Immobilized on a MALDI Plate. *Anal. Chem.* 78, 1574–1580
2. Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., and Aebersold, R. (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* 4, 231–237
3. Kettenbach, A. N., and Gerber, S. A. (2011) Rapid and Reproducible Single-Stage Phosphopeptide Enrichment of Complex Peptide Mixtures: Application to General and Phosphotyrosine-Specific Phosphoproteomics Experiments. *Anal. Chem.* 83, 7635–7644
4. The long road to reproducibility (2015) *Nat. Cell Biol.* 17, 1513–1514
5. Further confirmation needed (2012) *Nat. Biotechnol.* 30, 806–806
6. Announcement: Reducing our irreproducibility (2013) *Nature* 496, 398–398
7. Mertins, P., Qiao, J. W., Patel, J., Udeshi, N. D., Clauser, K. R., Mani, D. R., Burgess, M. W., Gillette, M. A., Jaffe, J. D., and Carr, S. A. (2013) Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nat. Methods* 10, 634–637
8. Frese, C. K., Zhou, H., Taus, T., Altelaar, A. F. M., Mechtler, K., Heck, A. J. R., and Mohammed, S. (2013) Unambiguous Phosphosite Localization using Electron-Transfer/Higher-Energy Collision Dissociation (ET_hcD). *J. Proteome Res.* 12, 1520–1525
9. Fort, K. L., Dyachenko, A., Potel, C. M., Corradini, E., Marino, F., Barendregt, A., Makarov, A. A., Scheltema, R. A., and Heck, A. J. R. (2016) Implementation of Ultraviolet Photodissociation on a Benchtop Q Exactive Mass Spectrometer and Its Application to Phosphoproteomics. *Anal. Chem.*,
10. Marx, H., Lemeer, S., Schliep, J. E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A. J. R., and Kuster, B. (2013) A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nat. Biotechnol.* 31, 557–564
11. Loroach, S., Zahedi, R. P., and Sickmann, A. (2015) Highly Sensitive Phosphoproteomics by Tailoring Solid-Phase Extraction to Electrostatic Repulsion-Hydrophilic Interaction Chromatography. *Anal. Chem.* 87, 1596–1604
12. Kettenbach, A. N., Sano, H., Keller, S. R., Lienhard, G. E., and Gerber, S. A. (2015) SPECHT - single-stage phosphopeptide enrichment and stable-isotope chemical tagging: quantitative phosphoproteomics of insulin action in muscle. *J. Proteomics* 114, 48–60
13. Humphrey, S. J., Azimifar, S. B., and Mann, M. (2015) High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nat. Biotechnol.* 33, 990–995
14. McAlister, G. C., Huttlin, E. L., Haas, W., Ting, L., Jedrychowski, M. P., Rogers, J. C., Kuhn, K., Pike, I., Grothe, R. A., Blethrow, J. D., and Gygi, S. P. (2012) Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* 84, 7469–7478
15. Paulo, J. A., McAllister, F. E., Everley, R. A., Beausoleil, S. A., Banks, A. S., and Gygi, S. P. (2014) Effects of MEK inhibitors GSK1120212 and PD0325901 in vivo using 10-plex quantitative proteomics and phosphoproteomics. *Proteomics*,
16. Ting, L., Rad, R., Gygi, S. P., and Haas, W. (2011) MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* 8, 937–940
17. Varjosalo, M., Sacco, R., Stukalov, A., van Drogen, A., Panyavsky, M., Hauri, S., Aebersold, R., Bennett, K. L., Colinge, J., Gstaiger, M., and Superti-Furga, G. (2013) Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MS. *Nat. Methods* 10, 307–314
18. Sharma, K., D'Souza, R. C. J., Tyanova, S., Schaab, C., Wiśniewski, J. R., Cox, J., and Mann, M. (2014) Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.* 8, 1583–1594
19. Batth, T. S., Francavilla, C., and Olsen, J. V. (2014) Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics. *J. Proteome Res.* 13, 6176–6186
20. McNulty, D. E., and Annan, R. S. (2008) Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol. Cell. Proteomics MCP* 7, 971–980
21. Villén, J., and Gygi, S. P. (2008) The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat. Protoc.* 3, 1630–1638

22. Kelstrup, C. D., Jersie-Christensen, R. R., Batth, T. S., Arrey, T. N., Kuehn, A., Kellmann, M., and Olsen, J. V. (2014) Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field Orbitrap mass spectrometer. *J. Proteome Res.* 13, 6187–6195
23. Fernández-Sáiz, V., Targosz, B.-S., Lemeer, S., Eichner, R., Langer, C., Bullinger, L., Reiter, C., Slotta-Huspenina, J., Schroeder, S., Knorn, A.-M., Kurutz, J., Peschel, C., Pagano, M., Kuster, B., and Bassermann, F. (2013) SCFFbxo9 and CK2 direct the cellular response to growth factor withdrawal via Tel2/Tti1 degradation and promote survival in multiple myeloma. *Nat. Cell Biol.* 15, 72–81
24. Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., and Elledge, S. J. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316, 1160–1166
25. Moritz, A., Li, Y., Guo, A., Villén, J., Wang, Y., MacNeill, J., Kornhauser, J., Sprott, K., Zhou, J., Possemato, A., Ren, J. M., Hornbeck, P., Cantley, L. C., Gygi, S. P., Rush, J., and Comb, M. J. (2010) Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Sci. Signal.* 3, ra64
26. Giansanti, P., Stokes, M. P., Silva, J. C., Scholten, A., and Heck, A. J. R. (2013) Interrogating cAMP-dependent Kinase Signaling in Jurkat T Cells via a Protein Kinase A Targeted Immune-precipitation Phosphoproteomics Approach. *Mol. Cell. Proteomics MCP* 12, 3350–3359
27. Duncan, J. S., Whittle, M. C., Nakamura, K., Abell, A. N., Midland, A. A., Zawistowski, J. S., Johnson, N. L., Granger, D. A., Jordan, N. V., Darr, D. B., Usary, J., Kuan, P.-F., Smalley, D. M., Major, B., He, X., Hoadley, K. A., Zhou, B., Sharpless, N. E., Perou, C. M., Kim, W. Y., Gomez, S. M., Chen, X., Jin, J., Frye, S. V., Earp, H. S., Graves, L. M., and Johnson, G. L. (2012) Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* 149, 307–321
28. Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* 25, 1035–1044
29. Daub, H., Olsen, J. V., Bairlein, M., Gnad, F., Oppermann, F. S., Körner, R., Greff, Z., Kéri, G., Stemmann, O., and Mann, M. (2008) Kinase-Selective Enrichment Enables Quantitative Phosphoproteomics of the Kinome across the Cell Cycle. *Mol. Cell* 31, 438–448
30. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S., and Mann, M. (2010) Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* 3, ra3
31. Meyer, S. C., Keller, M. D., Chiu, S., Koppikar, P., Guryanova, O. A., Rapaport, F., Xu, K., Manova, K., Pankov, D., O'Reilly, R. J., Kleppe, M., McKenney, A. S., Shih, A. H., Shank, K., Ahn, J., Papalex, E., Spitzer, B., Socci, N., Viale, A., Mandon, E., Ebel, N., Andraos, R., Rubert, J., Dammassa, E., Romanet, V., Dölemeyer, A., Zender, M., Heinlein, M., Rampal, R., Weinberg, R. S., Hoffman, R., Sellers, W. R., Hofmann, F., Murakami, M., Baffert, F., Gaul, C., Radimerski, T., and Levine, R. L. (2015) CHZ868, a Type II JAK2 Inhibitor, Reverses Type I JAK Inhibitor Persistence and Demonstrates Efficacy in Myeloproliferative Neoplasms. *Cancer Cell* 28, 15–28
32. Casado, P., Rodríguez-Prados, J.-C., Cosulich, S. C., Guichard, S., Vanhaesebroeck, B., Joel, S., and Cutillas, P. R. (2013) Kinase-Substrate Enrichment Analysis Provides Insights into the Heterogeneity of Signaling Pathway Activation in Leukemia Cells. *Sci. Signal.* 6, rs6
33. Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., Palma, A., Cesareni, G., Jensen, L. J., and Linding, R. (2014) KinomeXplorer: an integrated platform for kinome biology studies. *Nat. Methods* 11, 603–604
34. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127, 635–648
35. Casado, P., Alcolea, M. P., Iorio, F., Rodríguez-Prados, J.-C., Vanhaesebroeck, B., Saez-Rodriguez, J., Joel, S., and Cutillas, P. R. (2013) Phosphoproteomics data classify hematological cancer cell lines according to tumor type and sensitivity to kinase inhibitors. *Genome Biol.* 14, R37

36. Klammer, M., Kaminski, M., Zedler, A., Oppermann, F., Blencke, S., Marx, S., Müller, S., Tebbe, A., Godl, K., and Schaab, C. (2012) Phosphosignature Predicts Dasatinib Response in Non-small Cell Lung Cancer. *Mol. Cell. Proteomics* 11, 651–668
37. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villén, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010) A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. *Cell* 143, 1174–1189
38. Oppermann, F. S., Grundner-Culemann, K., Kumar, C., Gruss, O. J., Jallepalli, P. V., and Daub, H. (2012) Combination of chemical genetics and phosphoproteomics for kinase signaling analysis enables confident identification of cellular downstream targets. *Mol. Cell. Proteomics MCP* 11, O111.012351

Appendix

Acknowledgement

Nach vier Jahren mit Höhen und Tiefen ist es hiermit geschafft. Für meine Doktorarbeit schulde ich einigen Menschen einen herzlichen Dank.

Zunächst geht mein Dank an Bernhard, meinen Doktorvater, der diese Arbeit erst ermöglicht hat, der stets ein offenes Ohr hatte, mich immer begeistern konnte und von dem ich, nicht nur wissenschaftlich, sehr viel gelernt habe. Ausserdem geht ein großes Dankeschön an Simone, die mich in den ersten beiden Jahren der Dissertation betreut und unterstützt hat.

Desweiteren gilt mein Dank Prof. Dr. Jesper Olsen, der sich bereit erklärt hat diese Arbeit als Zweitprüfer zu begutachten und Prof. Dr. Martin Kingenspor, der sich als Prüfungsvorsitzender zur Verfügung gestellt hat.

Außerdem möchte ich mich bei meinen Masterstudenten Heiner, Jana, Alex, Svenja und Sandra, sowie bei meiner Bachelorstudentin Lin und meiner Forschungspraktikantin Julia für die tatkräftige Unterstützung bedanken. Ihr alle habt einen wichtigen Beitrag zu dieser Dissertation geleistet.

Ein großer Dank geht auch an alle Leute vom Lehrstuhl für Proteomik und Bioanalytik: ihr habt stets für eine freundschaftliche, entspannte und unterstützende Atmosphäre gesorgt. Ganz besonders möchte ich mich bei Andrea, Micha, Gabi, Silvia und Andi bedanken. Danke an Jana, Stefan und Steffi fuer euer aufmerksames Korrekturlesen der Arbeit.

Bei meinen langjährigen Freunden möchte ich mich für die oftmals fehlende Zeit entschuldigen und mich für euer Verständnis bedanken.

Zum Schluss möchte ich mich bei meiner Familie und insbesondere bei meinen Eltern bedanken. Ihr habt mich den gesamten Weg über unterstützt, mir in jeglicher Hinsicht den Rücken freigehalten, mich immer aufgebaut und euch mit mir gefreut. Danke.

Publication record

First author

- MALDI-TOF- and nESI-Orbitrap-MS/MS identify orthogonal parts of the phosphoproteome, **Ruprecht B**, Roesli C, Lemeer S, Kuster B. *Proteomics*. Accepted for publication.
- Evaluation of Kinase Activity Profiling Using Chemical Proteomics, **Ruprecht B**, Zecha J, Heinzlmeir S, Médard G, Lemeer S, Kuster B. *ACS Chem Biol*. 2015 Dec 18; 10(12):2743-52.
- Comprehensive and reproducible phosphopeptide enrichment using iron immobilized metal ion affinity chromatography (Fe-IMAC) columns, **Ruprecht B**, Koch H, Médard G, Mundt M, Kuster B, Lemeer S, *Mol Cell Proteomics*. 2015 Jan; 14(1):205-15.
- Proteomic analysis of phosphorylation in cancer, **Ruprecht B**, Lemeer S., *Expert Rev Proteomics*. 2014 Jun; 11(3):259-67. Review.

Co-author

- Optimized chemical proteomics assay for kinase inhibitor profiling, Médard G, Pachi F, **Ruprecht B**, Klaeger S, Heinzlmeir S, Helm D, Qiao H, Ku X, Wilhelm M, Kuehne T, Wu Z, Dittmann A, Hopf C, Kramer K, Kuster B., *J Proteome Res*. 2015 Mar 6; 14(3):1574-86.
- Ion mobility tandem mass spectrometry enhances performance of bottom-up proteomics, Helm D, Vissers JP, Hughes CJ, Hahne H, **Ruprecht B**, Pachi F, Grzyb A, Richardson K, Wildgoose J, Maier SK, Marx H, Wilhelm M, Becher I, Lemeer S, Bantscheff M, Langridge JI, Kuster B, *Mol Cell Proteomics*. 2014 Dec; 13(12):3709-15.
- DMSO enhances electrospray response, boosting sensitivity of proteomic experiments, Hahne H, Pachi F, **Ruprecht B**, Maier SK, Klaeger S, Helm D, Médard G, Wilm M, Lemeer S, Kuster B., *Nat Methods*. 2013 Oct; 10(10):989-91.
- Characterization of a high field Orbitrap mass spectrometer for proteome analysis, Pachi F, **Ruprecht B**, Lemeer S, Kuster B., *Proteomics*. 2013 Sep; 13(17):2552-62.
- Characterization of a chemical affinity probe targeting Akt kinases, Pachi F, Plattner P, **Ruprecht B**, Médard G, Sewald N, Kuster B., *J Proteome Res*. 2013 Aug 2; 12(8):3792-800.
- Comparing immobilized kinase inhibitors and covalent ATP probes for proteomic profiling of kinase expression and drug selectivity., Lemeer S, Zörgiebel C, **Ruprecht B**, Kohl K, Kuster B., *J Proteome Res*. 2013 Apr 5; 12(4):1723-31..

Curriculum vitae

Personal information

Name	Benjamin Lorenz Robert Ruprecht
Date and place of birth	1985/10/12 in Altötting, Germany
Current address	Egilbertstrasse 37, 85354 Freising, Bavaria, Germany
Current position	PhD student at the Chair of Proteomics and Bioanalytics at the Technical University of Munich, Freising, Bavaria, Germany
Phone	+49-8161-984834 (private), +49-176-23575333 (mobile)
E-mail	Benjamin.Ruprecht@tum.de
Nationality	German

Education

1996/08 – 2005/06	High school diploma, Ruperti-Gymnasium, Mühldorf
2005/07 – 2006/03	Community service, district hospital, Mühldorf
2006/05 - 2006/07	Diploma, Mathematical Sciences, LMU
2006/09 – 2009/09	Bachelor of Science (B.Sc.), Molecular Biotechnology, TUM, passed with high distinction B.Sc. Thesis: “Synthesis of beta amyloid peptide derivatives”
2009/10 – 2011/11	Master of Science (M.Sc.), Molecular Biotechnology, TUM, passed with high distinction M.Sc. Thesis: “Systematic selectivity screening of small molecule AKT kinase inhibitors with chemical proteomics”
since 2012/01	Doctoral candidate at the Chair of Proteomics and Bioanalytics under supervision of Prof. Dr. Bernhard Kuster

Conference attendance including poster presentation

2013: HUPO, Yokohama
2013: Proteomic Forum, Berlin
2014: ASMS, Baltimore
2014: CIPSM, Wildbad Kreuth
2015: Proteomic Forum, Berlin
2015: HUPO, Vancouver