

# TECHNISCHE UNIVERSITÄT MÜNCHEN

*Lehrstuhl für Proteomik und Bioanalytik*

## Quantitative chemical proteomics for cancer characterization

Zhixiang Wu

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: Univ.-Prof. Dr. D.Langosch

Prüfer der Dissertation:

1. Univ.-Prof. Dr. B. Küster

2. Univ.-Prof. Dr. J. Buchner

Die Dissertation wurde am 05.03.2012 bei der Technischen Universität München  
eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung,  
Landnutzung und Umwelt 16.04.2012 angenommen.



1. General Introduction.....	1
1.1 Cancer and kinase in cancer.....	1
1.2 Mass spectrometry-based proteomics.....	5
1.2.1 Sample preparation.....	7
1.2.2 MS analysis.....	9
1.2.3 Data processing.....	14
1.2.4 Mass spectrometry-based quantitative proteomics.....	15
1.3 Chemical proteomics.....	17
1.4 Kinome profiling.....	19
1.5 Aim and outline of this thesis.....	20
2. Establishment and optimization of Kinobead based quantitative chemical proteomics.....	21
2.1 Introduction.....	21
2.2 Material and methods.....	23
2.2.1 Cell culture and harvesting.....	23
2.2.2 Study for cell lysis with different detergents.....	23
2.2.3 Study for cell lysis with different concentrations of phosphatase inhibitors.....	23
2.2.4 Placenta lysis.....	24
2.2.5 Affinity purification and protein digestion.....	24
2.2.6 In-solution protein digestion.....	25
2.2.7 In-gel protein digestion.....	25
2.2.8 Dimethyl labeling of tryptic peptides.....	25
2.2.9 STAGE-tips purification.....	26
2.2.10 LC-MS/MS analysis.....	26
2.2.11 Peptide and protein identification and quantification.....	26
2.2.12 Statistical analysis.....	28
2.3 Result and Discussion.....	28
2.3.1 Kinobead establishment.....	28
2.3.2 Detergent selection.....	30
2.3.3 Kinobead capacity.....	31
2.3.4 Concentration of phosphatase inhibitors.....	33

## ***Table of contents***

---

2.3.5 Comparison of spectra counting based and intensity based label-free quantification .	34
2.3.6 Comparison of the dimethyl labeling and intensity based label-free quantification.....	37
2.3.7 Viability of intensity based label-free quantification for chemical proteomics.....	39
2.4 Conclusion .....	40
3. Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer .....	42
3.1 Introduction.....	42
3.2 Material and methods.....	44
3.2.1 HNSCC cell lines .....	44
3.2.2 Cell culture and harvesting.....	44
3.2.3 Affinity purification and protein digestion .....	45
3.2.4 LC-MS/MS analysis .....	45
3.2.5 Peptide and protein quantification and identification.....	45
3.2.6 Statistical analysis.....	45
3.2.7 Western blot analysis .....	46
3.2.8 Immunohistochemistry analysis.....	46
3.2.9 Sample preparation for DNA sequencing.....	47
3.2.10 Kinase inhibitor treatment .....	48
3.2.11 siRNA treatment.....	48
3.2.12 Ephrin A1 Fc treatment .....	48
3.3 Results and Discussion .....	49
3.3.1 Experimental strategy of identifying differential kinase protein expression in HNSCC cell lines .....	49
3.3.2 Targets profiling on HNSCC by chemical proteomics combined with quantitative mass spectrometry .....	50
3.3.3 Statistic analysis .....	52
3.3.4 Validation of the MS data by Western blot.....	54
3.3.5 Immunohistochemical analysis of EGFR on clinical sample .....	55
3.3.6 Confirmation of EGFR contributing to the proliferation of some but not all HNSCC.....	56
3.3.7 Synonymous substitutions in EGFR encoding sequence .....	59
3.3.8 Loss of function screening by siRNA highlights new candidate drug targets.....	60

3.3.9 Target validation of EPHA2 .....	62
3.3.10 EPHA2 was differentially expressed in the clinical samples.....	63
3.3.11 Combination treatment by targeting EGFR and EPHA2.....	65
3.3.12 Pathway analysis .....	66
3.3.13 Investigation of Src family kinases in HNSCC .....	67
3.3.14 Evaluation of c-MET as a target in HNSCC .....	69
3.4 Conclusion.....	71
4. Systematic identification of the HSP90 regulated proteome .....	73
4.1 Introduction .....	73
4.2 Material and methods.....	76
4.2.1 SILAC labeling and cell culture .....	76
4.2.2 Pulsed SILAC.....	76
4.2.3 Drug treatment and harvesting.....	76
4.2.4 Gel based sample preparation for full proteome .....	77
4.2.5 Affinity purification .....	77
4.2.6 Immunoprecipitation and GA-NHS affinity purification.....	77
4.2.7 LC-MS/MS analysis .....	78
4.2.8 Peptide and protein quantification and identification .....	78
4.2.9 Statistic analysis .....	78
4.2.10 GO enrichment/ Pathway analysis/complex analysis .....	79
4.2.11 Immunoblot analysis.....	79
4.3 Results and discussion.....	80
4.3.1 Strategy of kinome and proteome wide investigation of HSP90 regulation .....	80
4.3.2 GA reduces cell viability .....	81
4.3.3 Proteome wide mass spectrometric identification and quantification of protein changes induced by GA.....	82
4.3.4 Data evaluation .....	83
4.3.5 High quality map of HSP90 regulated proteins.....	85
4.3.6 Kinome-wide quantitative profiling reveals new potential clients of HSP90 .....	86
4.3.7 GO analysis reveals subgroups of proteins with diverse response to the HSP90 inhibition. ....	88

## ***Table of contents***

---

4.3.8 Cancer specific pathways and networks involved in different cancers .....	91
4.3.9 Co-regulation in protein complex induced by GA .....	92
4.3.10 Pharmacological intervention with HSP90 regulated pathways .....	94
4.3.11 Protein turnover time is a main gradient of HSP90 inhibition induced hierarchical destabilization .....	96
4.3.12 Proteins with protein kinase activity have shorter half-life .....	99
4.3.13 Identification of HSP90 physical interactions by immune precipitation and affinity purification .....	100
4.3.14 DDR1 is a novel bona fide client of HSP90 .....	102
4.4 Conclusion .....	104
Summary .....	105
Zusammenfassung.....	107
Reference .....	109
Abbreviations .....	124
Appendix .....	127
Acknowledgments.....	135
Curriculum Vitae .....	136

## **1. General Introduction**

### **1.1 Cancer and kinase in cancer**

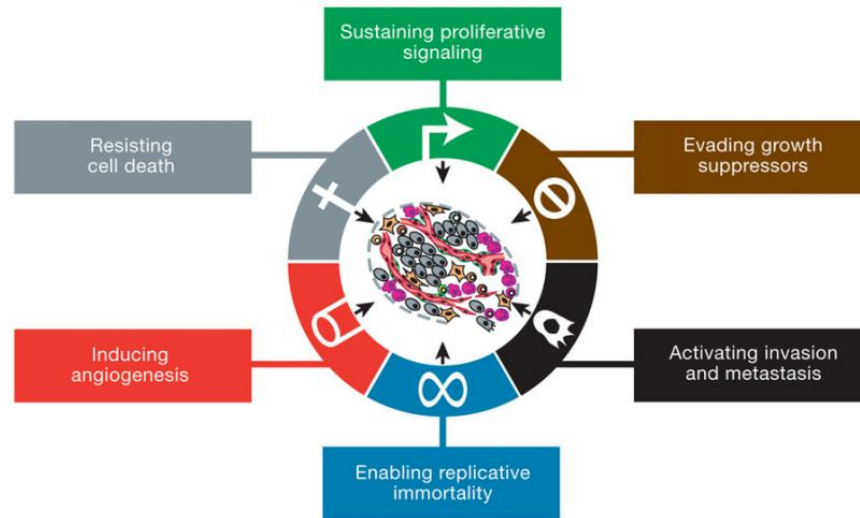
Cancer refers to a large group of different diseases characterized by failure of regulation of cell growth, which ultimately results in uncontrolled expansion of cancer cells and malignant tumor formation (1). Despite the slight decline of the incidence rate (by 0.6% per year) in men and cancer death rates (by 1.7% per year) in total, cancer remains a worldwide major public health problem. The American Cancer Society estimates that a total of 1,638,910 new cancer cases are projected to occur in the United States in 2012 and 577,190 people will die from cancer, contributing to one quarter of the total death in US (2). The prevalence of tumor sites remains different between the sexes. For example, the five most common forms of cancer for men are prostate, lung, colorectal, bladder and melanoma, whereas in women the top five are breast, lung, colorectal, uterine corpus, thyroid and melanoma. Cancers of the lung, prostate, colorectal, pancreas and liver in men and cancers of the lung, breast, colorectal, pancreas and ovary in women are the most lethal forms of cancer. The general 5-year relative survival rates have been notably improved for all cancer types since 1975, which was around 67% between 2001 and 2007. However it varies widely between the cancer types. For instance, the 5-year survival rate of prostate cancer patients has been improved to almost 100% while only 6% of patients diagnosed with pancreas cancer have a more than 5 year survival rate (2). Besides, for most of the cancers, the five year survival rate highly depends on the stage of the cancer at point of diagnosis. Early detection of cancer greatly increases the chances for successful treatment and consequently improves the prognosis. For example, the statistic data from the National Cancer database of the US showed that the 5-year survival rate for patients diagnosed with 0 stage breast cancer was 93%, while it dramatically decreased to 15% for those patients diagnosed with stage IV breast cancer. Despite the notable improvement, the 5-year survival rates for most types of cancers diagnosed at late stages is still less than 20%; for some even less than 4%. Unfortunately, for several types of cancer including lung, ovary and pancreas, more than half are only detectable at the distant stage. Therefore, there is an urgent need of novel diagnostic approaches to improve the chance to diagnose these types of cancer at an early stage, including development of novel markers for early diagnosis.

## ***General introduction***

---

In most cases, cancer is a complex disease rising from multiple genetic and epigenetic changes in genes involved in the cancer-associated signaling pathways typically acquired over a longer period of time. In general, two broad categories of altered genes are required for a normal healthy cell to transform into a cancer cell, oncogenes (promoting cell growth and reproduction) and tumor suppressing genes (inhibiting cell division and survival) (3). Occasionally, a single gene alteration could also result in manifestation of cancer, such as chronic myelogenous leukemia (CML), most of which are driven by the constitutive activation of ABL kinase due to the “Philadelphia” chromosome translocation. Currently, phosphorylation-based signaling pathways are the most common pathways and research hotspots, of which many pathways have been revealed to be involved in diverse types of cancers. Given the fact that kinases are in the central nodes of mediating diverse intracellular phosphorylation-based signaling pathways, it may be not surprising that numerous kinases have been revealed to be *bona fide* oncogenes—drivers of tumor growth, including growth factor receptors (e.g., EGFR, v-KIT, MET) and signal transducers (e.g., SRC, ABL and AKT). The accumulation of multiple genetic lesions of kinase genes leads the normal cells to acquire the cancer-related phenotypes as summarized by Hanahan and Weinberg. This basically includes six characteristics as shown in Fig. 1, most of which could be promoted or exacerbated by kinase-involved dysregulation signaling (4, 5). (i) Self-sufficiency in growth signals. Mitogenic growth signals are absolutely required for a cell to change from the quiescent state into the proliferative state. In normal cells, these signaling cascades are initiated by stimulating the transmembrane receptors with different ligands (e.g., diffusible growth factors and extracellular matrix components) and transmitted into the cells via the coupled intracellular mediators to promote cell proliferation. Protein kinases are the primary components in these signaling pathways, including the transmembrane receptor tyrosine kinases (e.g., EGFR and MET) (6) the tightly coupled nonreceptor tyrosine kinases (e.g., SRC and YES) and further downstream serine/threonine kinases (e.g., AKT and AURKA). Aberrant activation of these kinases promotes the uncontrolled proliferation, which has been implicated in tumor initiation, progression, and acquired resistance. (ii) Insensitivity to antigrowth signals; the proliferation of the normal cell is under tight control of not only mitogenic growth signals mentioned above, but also by anti-proliferative signals. These signaling pathways are initiated by growth inhibitors and are largely directed by the cyclin-dependent kinases (CDKs), such as growth inhibitor TGF $\beta$ . Loss of TGF $\beta$  responsiveness could be a result of down-regulation of the TGF $\beta$  receptor (serine/threonine





*Figure 1. The six Hallmarks of Cancer, Adapted from Hanahan D and Weinberg AR, Cell. 2000*

kinase receptor) or failure of CDK4 inhibition (7). In addition, in some tumors carrying aneuploidy and other genomic aberrations, the mitosis related kinases, like the Aurora family and polo-like kinase 1 (PLK1), are often found to be up-regulated to promote cell division. (iii) Evading apoptosis. Complementary to the uncontrolled proliferation, evading apoptosis further allows the cancer to expand its population. Many RTKs and intracellular kinases contribute to this escape mechanism, including the TGF- $\beta$  pathway mentioned above. Another important signaling pathway involved in the acquisition of this capability is the PI3K-PTEN-AKT pathway. Either loss of function of PTEN or increased kinase activities of PI3K and/or AKT allow constitutive activation of this pathway for the cells to evade apoptosis. (iv) Limitless replicative potential. In distinction to normal cells that have a limit to the number of successive cell growth-and-division cycles to pass through, cancer cells attain the ability of limitless replicative potential in order to generate macroscopic tumors. Expression of telomerase to protect the ends of chromosomes is one major mechanism for this ability. However several CDKs are also required to overcome replicative senescence. (v) Sustained angiogenesis. Supply of oxygen and nutrients and elimination of the metabolites via the vasculature are absolutely required for the growth of tumors. So far, numerous of receptor tyrosine kinases and other protein kinases have been identified to contribute to the vascular recruitment and development, including the receptors of the growth factors VEGF, PDGF, FGF and HGF, as well as AKT and mTOR kinases. Many of these kinases have been developed as targets for cancer therapy. (vi) Tissue invasion and metastasis. This capability

## ***General introduction***

---

allows the cancer cells to invade adjacent tissues and migrate to distal sites representing a very late stage of cancer development, which significantly reduces the prognosis of the cancer patients as indicated earlier. Therefore, either diagnosis at the early stages or the inhibition of invasion and metastasis, avoiding the cancer to develop into the later stages, are very important for therapeutic management. The aberrant activation of several kinases (e.g., MET and CDK5) have been shown to promote the cell migration and metastasis (8, 9). The order and number of these six capabilities cells can acquire varies widely, even among tumors of the same types. Genome instability and variability allows the expression of these aberrant kinases involved in the six acquired capabilities of cancers. However, proper folding is essential for these aberrant kinases to be functional, which requires a higher activated chaperone machinery. So far, many of these kinases mentioned above have been identified as the *bona fide* clients of HSP90. In addition, overexpressed and/or hyperactivated state HSP90 has been found in many types of cancer indicating its oncogenic role in the process of tumorigenesis (10, 11).

Given the critical roles that kinases play in the development of cancer as described above, targeting these aberrant kinases present in a particular cancer may allow the selective killing of tumor cells or inhibition of their carcinogenic behavior. Currently, there are three modalities employed to inhibit the kinases activity; small molecular inhibitors, antibodies and small interference RNA (siRNA). Small molecular inhibitors refer to a group of small molecular weight organic compounds, either synthesized or natural products. Most of these small molecular inhibitors are designed to competitively bind to the ATP pocket and subsequently block the hydrolysis of ATP to inhibit the kinases activity, such as the first FDA approved kinase inhibitor drug Iressa, which specifically targets the EGFR kinase and inhibits its kinase activity. However these types of drugs are currently limited in their specificity and there is an emerging risk of acquiring resistance. In addition, some small molecular inhibitors targeting the outside of ATP pockets are currently under different clinical trials, for example, the allosteric Akt inhibitor MK-2206, which can specifically bind to the pleckstrin domain (PH) and consequently abrogate the interaction between AKT and PIP3 to inhibit the activity of the AKT-PI3K pathway (12). To date, all the FDA approved therapeutic antibodies are monoclonal and recognize a single epitope on a signal target, such as the anti-HER2 (ERBB2) antibody trastuzumab (Herceptin®) and the anti-EGFR antibody cetuximab (Erbix®)(13). Despite the success of clinical trials, their utility with regard to kinases is restricted to the receptors or their ligands on the cell surface due to their

impenetrability of the cells membrane. siRNA has provided a powerful tool for inhibiting the gene expression. Theoretically speaking, it is possible to design a rather specific siRNA drug against each protein including kinases. The lack of an effective delivery approach however limits the clinical application so far, however it becomes an important technique and platform for the validation of new drugs (14). Besides, distinct with the three approaches mentioned above that targets the kinases directly, inhibition of the chaperon machinery by targeting HSP90 offers an alternative strategy to inhibit the aberrant kinases indirectly. Numerous of the small molecular inhibitors designed to bind to the ATP pocket to inhibit the ATPase activity of HSP90 are currently under different stages of clinic trials.

With the efforts of past decades, a great progress has been made in understanding the role of specific kinases in ordinary cell biology as well as tumor carcinogenesis and the subsequent success of developing small molecule inhibitors and antibody-directed drugs for basic research, clinical studies and clinical therapy. However, a global level to understand the signaling networks remains challenging, such as the relevant questions of this thesis: how to classify the patient and predict which patient can benefit from a given therapeutic drug or strategy, which kinases can be developed as therapeutic targets, how does the cell maintain aberrant signaling networks, and so on. Numerous systematic approaches have been developed to address these questions including the genomics and proteomics approaches. Comparatively speaking, the proteomics approaches could provide a straightforward understanding of disease-relevant kinase function and signaling pathway on protein level regardless of the diversity of molecular mechanisms underlying the de-regulated kinase function in human cancers.

### **1.2 Mass spectrometry-based proteomics**

The proteome is the final product of genome expression and comprises all the proteins present in a cell at a particular time (15). Genome sequencing revealed that the human genome contains around 20,000 genes, while the whole human proteome has more than one million different proteins with an average of five post translational modifications, such as phosphorylation and acylation, on each protein (16). Complementary to genomics, proteomics is the large-scale study of gene and cellular function directly at protein level, including the identification and quantification of proteins, the study of post-translational modifications (PTMs) and protein-protein interactions, which could offer a better understanding of an organism (17, 18). Mass

## General introduction

spectrometry (MS) is an analytical technique developed early last century to measure the mass-to-charge ratio of charged particles. Due to two main breakthroughs: the availability of gene and genome sequence databases and the development of ionization techniques for large and polar molecules (matrix-assisted laser desorption ionization (MALDI), Electrospray, Tanaka, Fenn, Nobel Prize 2002). Protein or peptide sequencing via tandem mass spectrometry (MS/MS) becomes one of most comprehensive and versatile tools in large-scale proteomics (19, 20). According the different pre-processing of the protein samples, MS-based protein identification can be mainly divided into two categories: bottom-up proteomics and top-down proteomics (Fig. 2). In the bottom-up proteomics, sample proteins are proteolytically or chemically digested prior to the MS analysis, which is widely used in the proteomics field currently. In contrast in the top-down proteomics, the intact protein is typically submitted directly into MS after separation. Each of the approaches has its own advantages and disadvantages. For the bottom up approach, the sample will be digested first to get soluble proteolytic peptides that are subsequently separated with a powerful separation LC system. This approach is quite suitable for complex samples, which is considered to be more difficult to study on intact proteins. In addition, tandem mass spectra

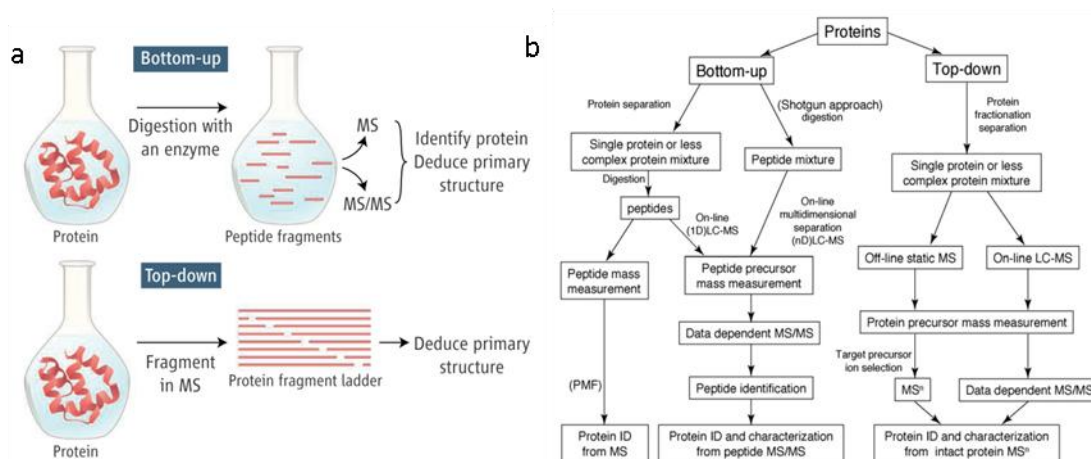


Figure 2. Schematics of MS based proteomics strategies. a) top-down versus bottom-up approaches in proteomics. Adapted from Chait, B. T. *Science*, 2006. b) outlines of diverse approaches for MS-based protein identification and characterization. Adapted from Han et al., *Curr Opin Chem Biol*, 2008

acquired from these proteolytic peptides often yield sufficient information to identify the peptides and proteins (21). However, because of the different physicochemical properties of the proteolytic peptides, only a small fraction of these peptides can be detected and generate useful MS/MS spectra for their identification. Consequently, the coverage of many identified proteins is

low, which might not be a problem for the parent protein identification or even quantification, but it's an obvious drawback for the characterization of post-translational modifications (PTMs) on parent proteins. In the top-down approach, the ionized intact protein is injected and subsequently fragmented in the mass spectrometer. A complete description of the primary structure of the protein can reveal all of its modifications, even any correlations that exist between these modifications could be deduced from the informative fragment ions acquired by MS (21). However this approach requires a higher performance MS, for example the Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry, to get enough resolution to generate de-convoluted spectra of the highly charged ions (22). Besides, the low throughput limits its application in large-scale proteomics (23). Nevertheless, both strategies contain three basic pillars: (i) the 'front end' fractionation of complex mixtures, (ii) mass spectral data acquisition and (iii) protein identification and characterization by database searching. Among these, the 'front end' fractionation of complex mixtures, separation of peptides or proteins, are a key elements for both approaches (20, 23). Fig. 3 illustrates a typical workflow of the MS based bottom-up proteomics experiment (24).

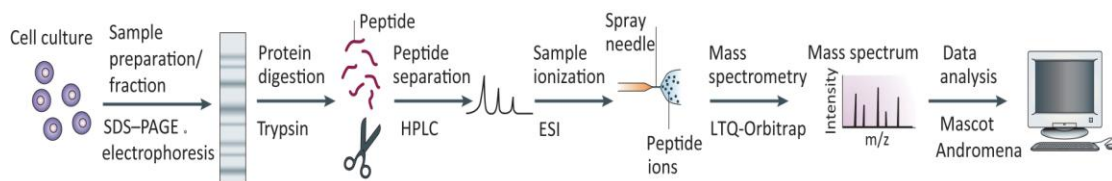


Figure 3. Workflow of a typical shotgun proteomics experiment. Modified from Steen, et al., *Nature Reviews Molecular Cell Biology*, 2004.

### 1.2.1 Sample preparation

Sample preparation includes the multiple steps of protein extraction, pre-separation and separation. However, concerning that the relevant chemical proteomics for the pre-separation in this thesis will be introduced later, the introduction here starts from the protein separation step prior to the coupled LC-MS. After extraction and solubilization from cells or tissue, the protein mixtures can be separated via two primary approaches: gel based and gel free. Gel based separation represents an effective and common approach in shotgun proteomics. For mass spectrometric analysis by ESI-MS, 1D gel electrophoresis separation is often used to separate the

## ***General introduction***

---

protein mixture according to their different masses and normally followed by in-gel digestion. A typical procedure of in-gel digestion includes destaining, reduction, and alkylation of cysteines, enzymatic cleavage of proteins into peptides, and extraction of peptides from the gel (25). Despite the diversity of available proteases, for large scale proteomics analysis, trypsin remains the most commonly used enzyme which allows generating well defined peptide fragments. Trypsin specifically and effectively cleaves C-terminal of lysine and arginine residues, except when either of them is followed by a proline residue. Tryptic peptides are mainly doubly charged at opposite ends of the peptide, which can be fragmented into both charged ions: C-terminal fragment ions and N-terminal fragment ions. The  $m/z$  value difference between successive members of each series can be used to deduce the amino acid sequence (26). The gel-free approach includes diverse fundamental different techniques recently summarized by Ahmed (27). In the large scale proteome research, most of the time tryptic peptides from 1D gel fractions are still far too complex for the capacity of any given mass spectrometer. Therefore, coupling with other small-scale preparative separation methods is often required to reduce the complexity of tryptic peptide mixtures, for example the ion-pairing reversed phase high-pressure liquid chromatography (RP-HPLC). In principle, any inert non-polar substance that achieves sufficient packing can be used for reversed-phase chromatography. The most popular column in proteomics is the octadecyl carbon chain (C18) bonded silica. Generally, tryptic peptides are first dissolved with aqueous buffer A (e.g., water) and loaded to the column. The retained peptides will be sequentially eluted by increasing the organic solvent buffer B (e.g., acetonitrile) due to their different hydrophobicity. Normally, acid is added to both buffers (e.g., 0.1% formic acid) to improve the chromatographic peak shape and to provide a source of protons for the peptides ionization for the following mass spectrometry analysis. For the same material, such as C18, high loading capacity, sensitivity, and dynamic range of the RPLC can be achieved by packing long, narrow capillary columns. However on the other hand this requires high pressure, which is limited by the HPLC. With decades of efforts, the resolution, efficiency, reproducibility and analysis speed of RPLC has been significantly improved. Together with the ESI compatible buffer, RP-HPLC is the most widely used on-line coupling technique in large-scale proteomics (28).

### 1.2.2 MS analysis

The next step is the analysis of the eluted tryptic peptides by MS. A typical mass spectrometer generally consists of three parts: an ion source to produce the gas-phase ions, a mass analyzer to separate the ionized analytes according to their mass-to-charge ratio ( $m/z$ ), and a detector that registers the number of ions at each  $m/z$  value. Once the peptides or proteins are ionized, they are transferred into the mass analyzer via the electric potential differences between the ion source and mass analyzer, and subsequently separated according to their different mass-to-charge ratios. Alternatively, to acquire the structural information, an extra collision reaction cell is implemented in the tandem mass spectrometer, where the peptides or proteins can be fragmented by different dissociation methods. The resulting fragment ions are subsequently separated in the mass analyzer. In the end, all separated ions with different mass-to-charge ratio can be detected by different kinds of detectors and translated into a mass spectrum containing the signal intensity against  $m/z$  values. The acquired raw mass spectra will be pre-processed to extract the signal out of the noise and followed by searching against the proper database to identify the peptides or proteins.

The analytes for MS analysis have to be converted into gas-phase charged ions. Although many ways of producing ions in the gas phase have been developed, two major techniques soft ionization techniques are most commonly used in proteomics research (19, 29, 30): Electrospray ionization (ESI) and MALDI. Electrospray ionization is a method by which analytes present in a solution can be ionized into gas phase, which was first introduced by Malcolm *et al.* and its underlying physical effect was first described by Geoffrey Taylor (31). However, the significance

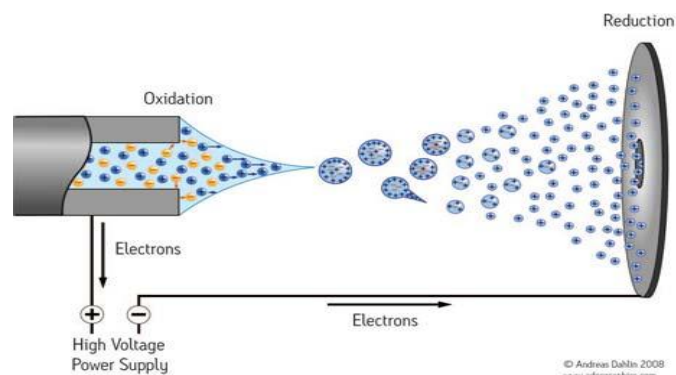


Figure 4. Schematic of major processes occurring in the atmospheric pressure region of electrospray. Adapted from <http://www.adorgraphics.com/galleri.html> (access date: 23.01.2012)

## General introduction

of ESI-MS in proteomics was only recognized later last century by John Fenn (29, 32) and afterwards it became a routinely used technique in proteomics research. Although the full mechanism of ESI has not been resolved yet (33), three major steps in the production of gas-phase ions were summarized by Kebarle *et al.*: (a) production of charged droplets at the ES capillary tip; (b) shrinkage of the charged droplets by solvent evaporation and repeated droplet disintegrations leading ultimately to very small highly charged droplets capable of producing gas-phase ions; and (c) the actual mechanism by which gas-phase ions are produced from the very small and highly charged droplets (Fig. 4)(34). Because of the ability of ionizing the analytes from the liquid phase, ESI is particularly suited for coupling with the liquid based separation tools (for example, liquid chromatography). Therefore integrated liquid-chromatography ESI-MS systems

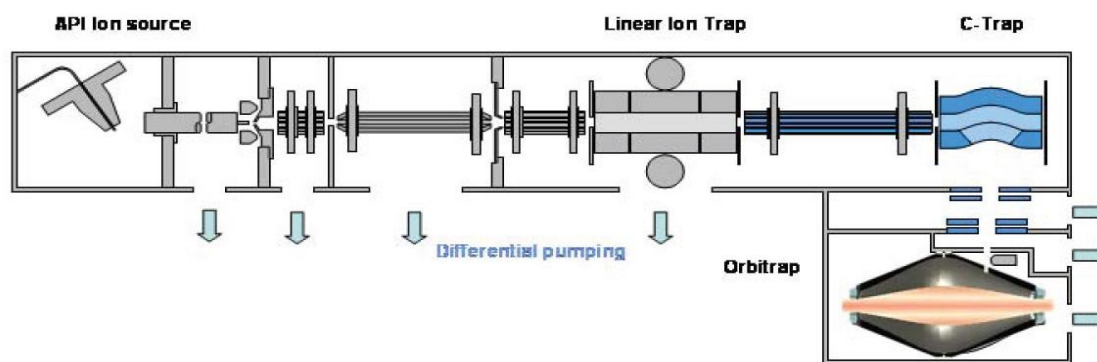


Figure 5. Schematics of LTQ-Orbitrap. In the LTQ-Orbitrap, an ion trap is combined to the Orbitrap analyzer for the ions selection and fragmentation. From Thermol fisher company.

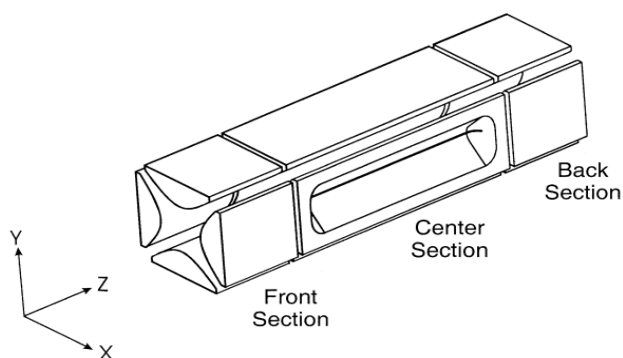
are preferred to be used for the analysis of complex samples. Another major advance enabling ESI-MS to sensitively detect protein and peptides is the development of nano-flow ESI (35, 36), which is particularly important to applications in proteomics because of the limited quantity of the samples. In matrix-assisted laser desorption ionization (MALDI), the sample is spotted and co-crystallized with the matrix solution on a plate (30, 37, 38), which can be ionized softly by laser stimulation. The feature that MALDI can retain the 2-dimension coordinate of the lased spots enables MALDI to carry the absolute advantage over electrospray in the imaging mass spectrometry field, where mass spectrometry is applied to visualize the spatial distribution of proteins, peptides, drug candidate compounds and their metabolites, and so on (39-42).

The commonly used mass analyzers in proteomics research can be cataloged into several types based on different physical principles: linear ion trap (LIT), time-of-flight (TOF), and Fourier-transform ion cyclotron resonance (FTICR) and Orbitrap mass analyzer (19, 20). Given the fact



that the hybridization of LIT and Orbitrap mass analyzer (LTQ-Orbitrap, Fig. 5) is the only relevant machine used in the thesis, a short overview about the LIT and Orbitrap mass analyzer is given here:

The linear quadrupole ion trap consists of four hyperbolic rods with parabolic shaped inner edges creating a hollow space where the ions are trapped and controlled, and each rod comprises of three parts: front section, center section and back section (Fig. 6) (43). This construction allows the trap to confine the ions radially by applying the radio frequency (RF) voltage to the four quadrupole rods to form a 2D RF field, and axially by applying DC potentials to end of each electrode. To function as an  $m/z$  analyzer, two extra phases of supplemental AC voltage are applied across the x-rods for ion isolation, activation, and mass-selective ejection, which also



*Figure 6. 2D quadrupole ion trap. Application of discreet DC voltages on each part of the rods allows trapping the ions along the z axis in the central section of the device. Adapted from Schwartz, et al., J Am Soc Mass Spectrom, 2002*

allow performing the tandem mass spectrum in the trap (44). First the precursor ions can be selected by isolation operation and fragmented by collision induced dissociation (activation). Subsequently, the fragmented ions are ejected out of the trap and detected by the detector. Compared to the three-dimensional (3D) Paul traps, linear traps have higher injection efficiencies and higher ion storage capacities, which consequently result in increased sensitivity, resolution and mass accuracy and faster scan speeds(45).

The Orbitrap, a new type of mass analyzer, was invented by Makarov and applied into proteomics research by Hu *et al.* (46, 47). The main attributes of this analyzer are high mass accuracy and resolution, which are relatively comparable with the FT-ICR instrumentation (48). In the Orbitrap, squeezed ions from C-trap are trapped in an electrostatic field generated by an outer barrel-like and an inner spindle-like electrode and oscillating along the central electrode

## General introduction

---

and the z-axis. The oscillating ions induce an image current into the two outer halves of the orbitrap, which is inversely proportional to the square root of the  $m/z$  and can be detected using a differential amplifier (Fig.7). The signals are amplified and transformed into a frequency spectrum by fast Fourier Transformation which is finally converted into a mass spectrum.

For the application in proteomics and related fields, the Orbitrap mass analyzer is typically combined with a low resolution linear ion trap (LTQ-Orbitrap), which performs the function of collecting ion populations in MS mode or selecting and fragmenting the ions within a certain  $m/z$  window in MS/MS mode. This combination has gained a huge success in proteomics research in the past years. In a typical proteomics experiment with LTQ-Orbitrap, the high resolution MS scan in Orbitrap and the low resolution MS/MS scans in the linear ion can be performed in parallel which results in obtaining high mass accuracy spectra for precursor peptides and fast fragmentation (49, 50).

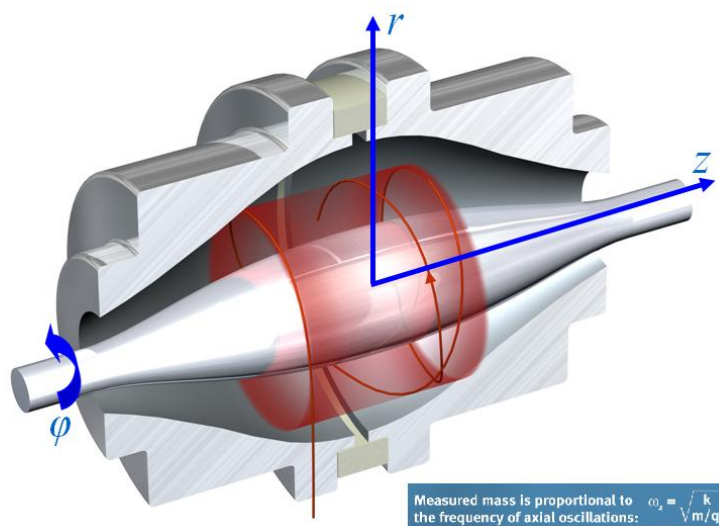
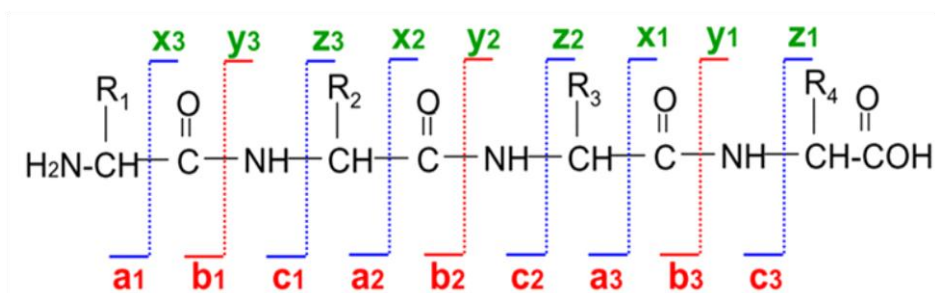


Figure 7. A cut away model of Orbitrap mass analyzer and the formula which axial oscillation frequency follows. Adapted from Scigelova and Makarov, *Proteomics*, 2006

The detector is the final part of a mass spectrometer, which is used to record either the charge induced or the current produced when an ion passes by or hits a surface. Two types of detectors are commonly implemented in the mass spectrometer for proteomics research: destructive detector and non-destructive detector. In the destructive detector, the ion impact generates secondary electrons which can amplify the original signaling, therefore can achieve a high sensitivity. Typically, in the LIT, the detector consists of the conversion dynode and the dynode electron multiplier. While, in Orbitraps, the detector is a non-destructive detector consisting of a

pair of metal surfaces within the mass analyzer/ion trap region which the ions only pass near as they oscillate. Only a weak AC image current rather DC current is produced by the analyte ions when circuiting between the electrodes. Compared to the multiplier, Orbitraps need more ions to generate the AC image; therefore it has a relatively low sensitivity.

Tandem mass spectrometry (MS/MS or MS<sup>2</sup>) is a key technique for protein or peptide sequencing as well as PTM analysis (20) consisting of multiple events: The first step is ion selection, the precursor ions from the full scan are isolated and accumulated; the second step is fragmentation which is performed on ions to break them apart and in the final step the fragment ions are separated by the mass analyzer and recorded by the detector. After isolation the target ions are fragmented by colliding with the inert gas or electron donor resulting in bond cleavage and generation of various fragment ions for peptides identification and PTM characterization. Several different fragmentation methods are currently available on different machines, among which Collision Induced (CID) and Electron Transfer Dissociation (ETD) are two widely used methods in proteomics research with different mechanisms and consequently generate different fragment ions. As shown in Fig 8, there are three possible types of breakage between two residues, NH-CH, CO-CH and CO-NH. Two species are generated after the backbone breakage and only charged species can be detected afterwards by MS. According to cleavage and charging status of the new



*Figure 8. Nomenclature of peptide fragment ions by Roepstorff and Fohlman.*

species, the fragment ions are indicated by a<sub>n</sub>, b<sub>n</sub>, or c<sub>n</sub> if the charge is retained on the N-terminus and x<sub>n</sub>, y<sub>n</sub> or z<sub>n</sub> if the charge is maintained on the C-terminus (51). The subscript n indicates the number of amino acid residues in the fragment. In the CID, kinetic energy from the collision of ions with inert gas is converted into internal (vibrational) energy mainly resulting in peptide backbone breakage which mainly consists of b- and y-ions. In ETD, peptide ions can react with the electron donor (e.g. Anthracen anion) and consequently the ion charge is reduced and paired

## ***General introduction***

---

electron configuration changes to unpaired electron configuration leading to fragmentation. Different from CID, the fragment ions are mainly c- and z-ions (52).

### **1.2.3 Data processing**

For the large-scale proteomics research, possibility-based database searching is the most effective way to interpret the tandem spectra (24, 53). However, the premise of this approach is the peptide of interest should be in the database. Profiting from the advance of DNA sequencing technology, many databases are available for this purpose. As introduced above, each peptide fragment in a series differs from its neighbor by one amino acid, which will be used as an input spectrum. Proteins from the database are first digested virtually into peptides and further generate the theoretical spectrum. By comparing the input spectrum with the theoretical spectrum, the most likely peptide is reported as output. Numerous commercial or free software is currently available for this comparison, such as Mascot and Andromeda (54). They are using different scoring algorithms, but are based on two common criteria (i) the peptide mass and (ii) the number and intensity of the peaks matched by the theoretically-computed  $m/z$  values of the fragment ions (55). The possibilities are presented as the identification scores: the higher the score, the higher possibility that the match is correct. With this possibility-based identification, the number of misidentifications (false positives) can easily increase to a high level by adding up the small error rates for each of the corresponding peptides, particularly for large-scale proteomics research. Therefore, stringent criteria are required to get high confident identifications for all of the proteins. However, this would be achieved at the cost of losing the confident identification of the proteins which are present in the sample (false negative)(24). To solve this problem, many searching engines allow using defined protein mixtures or randomized databases to evaluate the false positive rate (FDR), and adjust the criteria to control the FDR at a acceptable level (e.g., 1%). *De novo* is another way to interpret the MS/MS spectrum, which does not require any prior knowledge of the amino acid sequence of the proteins present in the sample (56-58). However, it remains a challenge even with the help of powerful bioinformatic tools because of the quantity and complexity of the tandem spectra, for example, missing peaks, different modifications and so on (58-60).

### **1.2.4 Mass spectrometry-based quantitative proteomics**

Rather than just giving a list of identified proteins, quantitative proteomics also provides quantitative information of the analyzed samples, as a major advance in proteomics, which allows researchers to quantitatively compare the samples. Relative and absolute quantification represent two forms of quantitative data. In relative quantification, the protein amount is relatively calculated to the same protein present in a reference sample, such as how many fold of changes after a certain treatment or relative changes between diseased and healthy state, whereas absolute quantification determines exactly how many copies or amount of a protein there is per cell (61). Compared to the classic methods of differential protein gel or blot staining by dyes and fluorophores, mass spectrometry-based approaches enable to identify and quantify many proteins in parallel, which has gained a huge success in the large-scale proteomics research field. Label free and isotope labeling are two basic strategies with a different basis. Label free quantification approaches achieve the relative or absolute quantification by correlating to the mass spectrometric signal of intact peptides (intensity based) or the number of peptide sequencing events (spectral counting). In contrast, isotope labeling approaches introduce a differential mass tag to the samples that can be recognized by a mass spectrometer and at the same time provide the basis for quantification. Fig. 9 summarizes the existing methodologies used for quantification. MS-based label-free quantification methods have emerged recently, consisting of two fundamentally different strategies: spectral counts and intensity. In the former approach, the number of MS/MS events from a given peptide were used to estimate the amount of the peptide, which is based on the empirical observation that the more peptide are presents in the samples, the more frequent a tandem MS spectra for this peptide can be acquired. In the intensity based approach, the extracted ion chromatogram (XIC) for a given peptide is extracted from different samples for the quantification. Numerous software packages have been developed to extract this information in batch, such as Progenesis LCMS and Maxquant (62). Most of the label-free quantification based strategies are focusing on relative quantification, however absolute quantification based on label-free approaches have also been investigated (63). Despite the fact that both systematic and non-systematic errors can be introduced between experiments, label-free quantification is worth considering compared to the isotope labeling approach for the following reasons: (i) No principle limit to the number of experiments that can be compared. (ii) More analytical depth (more identification) due to the relative simpler spectral

## General introduction

complexity. (iii) Higher dynamic range of quantification (64). However, stable LC-MS system and elaborate experiment procedure are absolutely required for the label-free quantification.

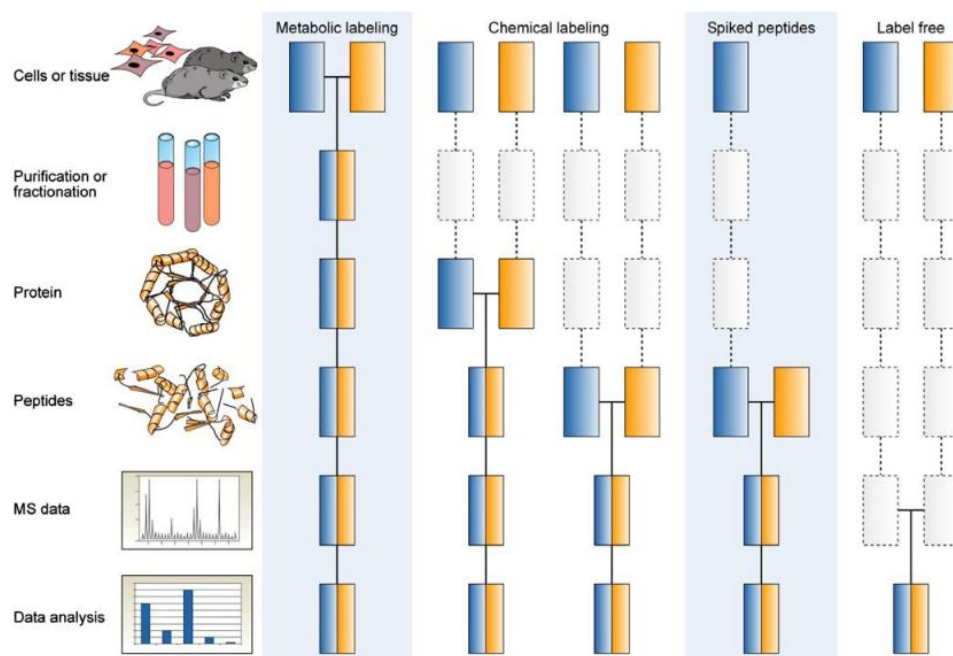


Figure 9. Workflows of common quantitative mass spectrometry. Boxes in blue and yellow represent two experimental conditions. Horizontal lines indicate when samples are combined. Dashed lines indicate points at which experimental variation and thus quantification error can occur. Adapted from Bantscheff et al., *Anal Bioanal Chem*, 2007

Metabolic incorporation of stable isotopes has been widely used in quantitative proteomics in the past ten years, such as stable Isotope Labeling with Amino acids in Cell culture (SILAC) where the normal arginine and lysine in the medium are replaced by  $^{13}\text{C}/^{15}\text{N}$  isotope labeled arginine and lysine which can be incorporated into all cellular proteins through cell growth and protein turnover. Cells cultured in heavy and light medium generate two cell populations and enable the relative quantification by comparing the intensities of isotope clusters of the intact peptide in the survey spectrum (61, 65). Profiting from the unique experimental design that the samples from different labeled channels can be mixed before lysis avoids bringing in errors in the following steps. Metabolic labeling based approaches represent the highest accurate method of MS-based quantification. Currently, the main approach of metabolic labeling is SILAC which is limited to higher eukaryote cells, like human cancer cell lines (64). Chemical tagging is another class of isotope labeling based approach, which brings in the isotope tags on protein level or peptide level via different chemical reaction, such as ICAT (66) and iTRAQ (67). Spiking the sample with

isotope labeled peptides is often used for absolute quantification because of the ascertain amount of spiked peptides which can be used to deduce the absolute amount of the corresponding peptides from the sample, such as the isotope labeled absolute quantification (AQUA) peptides which can be spiked in to the sample for absolute quantification (68).

### 1.3 Chemical proteomics

The human proteome contains between 20,000 and 25,000 non-redundant proteins and around 50,000 to 500,000 unique protein species due to RNA splicing, proteolysis and PTM events (69, 70) with a dynamic range of expression of more than six orders of magnitude (71), which is still an unmet technical challenge. Therefore, pre-separation of the complex protein mixture is absolutely required to get a deep investigation of a certain group of proteins. Numerous emerged strategies and tools devote to separation of a subproteome according to their unique characteristics shared by a group of proteins. For example, Gel Filtration can be used to separate the proteins according to their molecular weight, membrane proteins can be separated from the cytoplasmic proteins and immuno-precipitation (IP) can be used to separate the proteins sharing a common epitope (e.g., protein level phosphotyrosine IP), and so on. Among these, chemical proteomics is a promising tool focusing on the exploration of protein function using synthetic small molecules which target a subgroup of proteins related by binding site shape (properties compound-centric chemical proteomics, CCCP) or by chemical reactivities (activity-based probe profiling, ABPP)(72, 73). Compound-centric chemical proteomics basically is derived from the classical drug affinity chromatography developed over forty years ago by Cuatrecasas *et al.* (74). A typical chemical probe used for CCCP is based on affinity to ascertain target protein and contains two parts: a capture part which can be recognized by a subgroup of proteins carrying a common domain, such as the ATP pocket in the kinase domain, and an flexible linker ending with an amino-, carboxyl- or hydroxyl group, which can be used to immobilize the probes to the matrix such as agarose beads, without interfering with the biological activity of the other part. The emergence of employing high-performance mass spectrometry for protein analysis remarkably expands its application scope in proteomics, which allows identification of a subgroup of proteins either directly or indirectly bound to the probes (75-78). Fig. 10 demonstrates a typical strategy for CCCP: after incubating the sample with the immobilized probes, the target proteins can be co-eluted from the matrix, and can be further separated through different ways, for example 1D or

## General introduction

2D gel, and followed by LC-MS analysis. Normally, the chemical probes used here need a high affinity against their targets, for example with the  $K_d$  in nM scale, which allows one to wash the matrix stringently to reduce the unspecific binding. The washing conditions are another rather important parameter for this affinity based chemical proteomics. Improper washing either yields too less targets (too stringent washing) or high background from the unspecific binding proteins (too soft washing), which thereof has to be optimized beforehand (79).

Besides the huge success in drug discovery (75, 80, 81), CCCP has also been used to characterize the biological function of the target protein where normally a highly selective immobilizable inhibitor is required. The interactors of the target protein could be co-purified out by using proper washing conditions, which can be used to unravel the biological roles of the target protein.

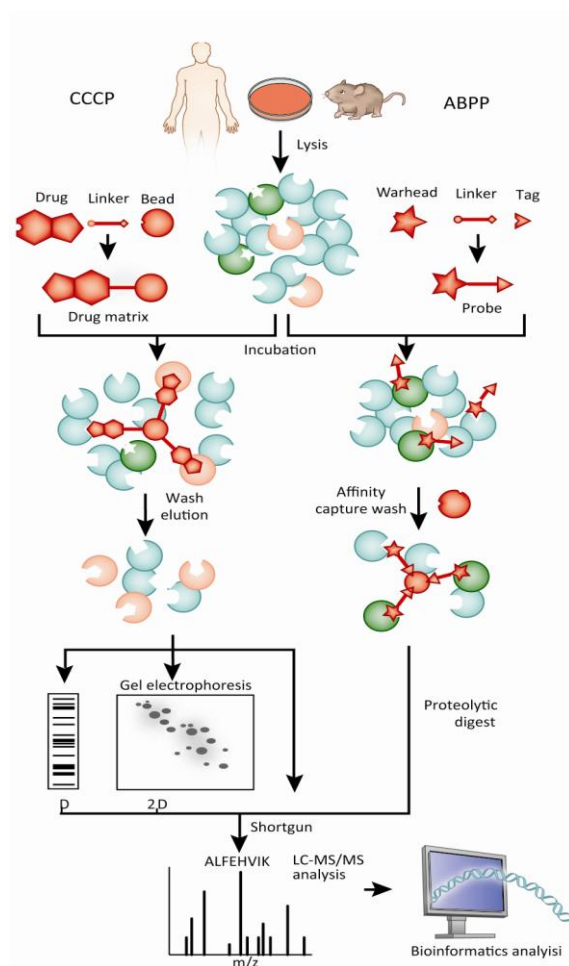


Figure 10. Strategies of MS-based chemical proteomics, Modified from Rix and Superti-Furga, *Nat Chem Biol*, 2009.



Recently, Moullick *et al.* applied this strategy to identify the HSP90 dependent oncogenic proteins and successfully revealed several cancer-specific networks coordinated by Hsp90 (77). Comparing to the affinity based probes, activity-based probes contain two main components: a reactive group or warhead which can covalently label the target proteins, typically an electrophilic or photo-reactive group (82), and a tag working as a reporter for the detection, enrichment and identification, such as biotin or a dye (83). However, more often there is a third part, which could be a pure linker between the abovementioned two groups or it serves as the site-directing function which directs the reactive element towards specific targets (82).

### 1.4 Kinome profiling

As introduced at the beginning of this chapter, protein kinases are involved in almost all the cellular processes by phosphorylating up to 30% of the proteome. Gene amplification, overexpression and point mutation induced aberrant kinases expression and/or activation have been found in numerous pathophysiological events including cancer. Profiling the kinome of a particular tumor allows the discovery of deregulated signaling pathways induced by the aberrant kinases, which may offer the drug target for the therapeutic intervention and therefore may provide biological markers for future guidance in personalized medicine. To date, numerous strategies have been designed to analyze the kinome status (84, 85). Among these, the kinase-centric chemical proteomics technology is one of the most promising approaches (75, 84). By immobilizing the ATP mimetics to agarose-beads via a linker, it enables to purify hundreds of protein kinases from human and rodent cell lines and tissues by competitively binding to the ATP pocket within the protein kinases (75). Conjugation with quantitative mass spectrometry, it allows the quantitative comparison of the enriched kinases from different samples. Originally this technique was developed for drug targets profiling. Cell or tissue lysate are pre-incubated with either vehicle (e.g., DMSO) or drug over a range of concentrations (e.g., from 100 pM to 10 mM) and subsequently subjected to Kinobead precipitation. Depending on the affinity of the kinases to the drug, different amounts of the free kinases is available for capturing on Kinobead. Targets with high affinity will be blocked by the drug, whereas non targeted kinases and other proteins are unaffected (75). Recently it has been used to profile the clinic samples in response to the different therapeutic interventions to reveal the potential drug targets in kinome wide. Cells were isolated from cancer patients and treated with various concentration of drug targeting the

## ***General introduction***

---

kinases. Subsequently the samples were profiled via the Kinobeads to investigate the potential targets of the testing drugs and the targets mostly contributing the cell apoptosis (86).

### **1.5 Aim and outline of this thesis**

This thesis comprises three projects with the common goal of using Kinobead based chemical proteomics in conjunction with quantitative mass spectrometry for cancer characterization. The aim of first project was to establish and optimize the Kinobead technique for target profiling from different samples and systematically evaluate the performance of label free quantification in a kinase-centric chemical proteomics, which is described in Chapter 2.

Chapter 3 describes the project in which 34 head and neck cancer cell lines were profiled employing the established quantitative chemical proteomics strategy in the first project. 30 targets out of the 42 statistical significant kinases were submitted for siRNA mediated knockdown for further validation and resulted in 9 potential therapy targets in head and neck cancer. The result suggested that this quantitative chemical proteomic strategy is indeed a useful approach for therapeutic targets discovery.

In chapter 4, the Kinobead approach together with a full proteome strategy was applied for the characterization of the HSP90 regulated proteome which may offer some new insights for cancer intervention by targeting HSP90. The result implied numerous kinases to be novel clients of HSP90. With the assistance of bioinformatics tools, different HSP90 coordinated pathways were revealed to be involved in different cancer cell lines. Pulsed-SILAC experiments also provided a new interpretation for the differential dependence of the clients on HSP90.

## **2. Establishment and optimization of Kinobead based quantitative chemical proteomics**

### **2.1 Introduction**

Kinase (alternatively called phosphotransferase) is an enzyme which catalyzes the transfer of the terminal phosphate from ATP to the specific substrates. With the completion of the human genome sequencing, the whole set of human protein kinases were catalogued by Sudarsanam *et al.* In total, the kinome comprises 518 protein kinases constituting about 1.7% of all human genes (87). Coordinating with phosphatases, an enzyme which removes the phosphate group from its substrates, kinases mediate most of the signal transduction in eukaryotic cells and consequently regulate many cellular processes, including cell cycle progression, apoptosis, and differentiation. Dysregulated kinases have been indicated in many pathological events, such as cancer. Profiling the kinome of a particular tumor thus allows the discovery of deregulated signaling pathways induced by the aberrant kinases, which may offer the drug targets for the therapeutic intervention and therefore may provide biological markers for future guidance in personalized medicine.

Advanced genomic techniques, such as microarray or next generation sequencing techniques, enable the rapid and sensitive analysis of the transcriptome. However, information from mRNA levels only offers an indirect evaluation of protein abundance and function due to either poor correlation between mRNA and protein concentration or the post translational modifications which cannot be predicted from the deduced amino acids sequence (88-90). In contrast, protein level result gives a straightforward assessment of the kinase quantity and/or activity, such as the classic kinase activity screen assay. The kinase activity can be estimated according to the phosphorylation level of itself or its specific substrates (91). Employing similar strategies important improvements have been made using protein array technology which enables the analysis of many kinases or substrates in parallel, however this strategy mainly suffers from the availability of specific antibodies and/or specific substrates (85). Chemical proteomics in conjunction with mass spectrometry, as introduced above, allows focusing on a certain subproteome sharing similar physiochemical properties. In this thesis, focus lies on the kinome, a group of proteins with kinase domains, and a structurally conserved protein domain containing the catalytic function of the phosphotransferase. There are two indispensable conserved regions in the kinase domain: an ATP binding region in the N-terminal extremity of the catalytic domain

## ***Kinobead establishment***

---

and a conserved aspartic acid residue in the central part of the catalytic domain which is important for the catalytic activity of the enzyme (92). The former region (ATP pocket) also allows the selective binding of ATP analogs, which can be further developed as probes for affinity-based chemical proteomics for drug discovery or biological research (75, 93). In a typical experiment, these compounds are immobilized to agarose beads via a linker, resulting in the probe matrix subsequently used to affinity capture the kinases from cell or tissue extracts. Considering the fact that human cells contains high concentrations of ATP (~ 2 mM) and high ATP affinity for most of the kinase ( $K_{d(ATP)} < 0.1$  mM), an extreme high affinity is absolutely required for these compounds to effectively bind to the ATP pocket instead of ATP. The binding efficiency can be estimated by the equation:  $IC_{50} = (K_d(\text{inhibitor}) (1 + ([ATP]/K_d(ATP))))$ , where the  $K_d$  value is determined by competition binding assay experiments. Most of the currently designed compounds have an nM scale  $K_d$  for their targets which can effectively bind to the ATP pocket according to the aforementioned equation (94, 95). In addition, several other factors also affect the efficiency of the whole approach, including the expression level of target proteins and extraction methods, both of which have an effect on the concentration of the target proteins; the compound coupling density, could be too low and too high resulting in low binding capacity and high unspecific binding to the probe matrix, respectively (96); the kinases activity, it has been found that some of the target kinases only can bind to the probes in the active form (75). Recently, another activity-based probe using ATP or ADP as the site-directing group has been developed containing three parts: the ATP or ADP group, an acyl phosphate reactive group, which can react with the adjacent lysine residue, and a biotin group for enrichment (97, 98). However, the requirement of a lysine nearby the ATP pocket and plenty of non-kinase proteins containing an ATP binding domain remain challenging for this strategy.

Kinobead is one of the affinity-based chemical proteomics technology originally developed by Kuster and colleagues at Cellzome (75). By immobilizing pan-kinase inhibitors to agarose beads via a linker, it enables the purification of around 350 protein kinases from human and rodent cell lines and tissues by competitively binding to the ATP pocket within the protein kinases (75). The technique was originally developed for drug target proofing. In this study however, serial parameters of the Kinobead based chemical proteomics were optimized in order to apply the approach to profile the kinome for cancer characterization. Two quantification methods were compared was leading to the selection of intensity based label free quantification for the

following experiments. Systematic evaluation of the variance of the whole strategy revealed that less than 10% and 5% of the variance was induced by biological and technical replicates respectively.

## **2.2 Material and methods**

### **2.2.1 Cell culture and harvesting**

K562 cells were cultured in Roswell Park Memorial Institute medium 1640 (RPMI1640) with stable glutamine (PAA, Pasching, Austria) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS, PAA, Pasching, Austria) at 37 °C in the incubator (Cell Genix, Brand) with humidified air and 5% CO<sub>2</sub>. K562 cells were washed 3 times with pre-cooled PBS after harvesting by centrifugation.

The entire 34 head and neck cell lines used in this study are listed in the appendix table S1 representing HNSCC of the tongue and were kindly supplied by Dr. Stephan Feller from Oxford University, Dr. Suzanne A. Eccles from McElwain Laboratories, Dr. Susanne M. Gollin from University of Pittsburgh and Prof. Reidar Grenman, University of Turku. With the exception of the two cell lines OSC-19 and OSC-20, all were originally obtained from primary tumors (see appendix table S1). Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) with high glucose and glutamine (PAA, Pasching, Austria) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS, PAA, Pasching, Austria), 1x non-essential amino acids (NEAA, PAA, Pasching, Austria) at 37 °C in humidified air with 5% CO<sub>2</sub>.

### **2.2.2 Study for cell lysis with different detergents**

Harvested K562 were lysed in 50 mM Tris/HCl pH 7.5, 5% glycerol, 1.5 mM MgCl<sub>2</sub>, 150 mM NaCl, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 25 mM NaF, and protease inhibitors (SIGMAFAST, Sigma-Aldrich) containing different detergents (See Results part). Homogenates were centrifuged at 6000x g at 4 °C for 10 min to remove cell debris. Protein concentration in lysates was determined by the Bradford assay according to the manufacture's instruction.

### **2.2.3 Study for cell lysis with different concentrations of phosphatase inhibitors**

The whole panel of cell lysate for target profiling was kindly supplied by Dr. Stephan Feller, Jessica B. Doondeea and Melanie C. Janning from Oxford University and prepared according to the following protocol: Prior to harvesting, cells were cultured in medium for 48 h. Cells were

## ***Kinobead establishment***

---

grown to 90-100% confluence while avoiding substantial overgrowth. Before lysis, cells were washed 3 times with ice cold phosphate buffered saline (PBS). Radio immunoprecipitation assay (RIPA) 100 buffer (mixed micelle buffer with 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.5% deoxycholic acid, 0.1% SDS) with freshly added protease (2x protease inhibitor cocktail, Roche Applied Science, Mannheim, Germany) and different concentrations of phosphatase inhibitors (phosphatase inhibitor cocktail 1, Sigma-Aldrich, Munich, Germany, phosphatase inhibitor cocktail 2, Sigma-Aldrich, Munich, Germany, 1 mM sodium ortho-vanadate and 1 mM sodium molybdate) was used for cell lysis. One ml of ice cold lysis buffer per 175 cm<sup>2</sup> cell culture flask was then added to the cells. Cells were scraped immediately, collected in a pre-cooled micro-centrifuge tube and incubated on an orbital mixer for 30 min at 4 °C. Lysates were then centrifuged at 2 °C for 30 min at 20,000x g, supernatants collected, aliquoted, frozen in liquid nitrogen and stored at -80 °C until further use.

### **2.2.4 Placenta lysis**

Human placenta or cells was de-frozen at 37°C in water bath and washed with pre-cooled PBS to get rid of the blood, and subsequently lysed in 50 mM Tris/HCl pH 7.5, 5% glycerol, 0.8% NP-40, 1.5 mM MgCl<sub>2</sub>, 150 mM NaCl, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 25 mM NaF, 1 mM DTT and protease inhibitors (SIGMAFAST, Sigma-Aldrich) and 5x phosphatase inhibitor cocktail I and II (ChemBiochem, Merk). Homogenized and lysed cells were centrifuged at 6000x g and 4 °C for 15 min to remove cell debris. The supernatant was transferred into polycarbonate ultracentrifuge tubes and cleared by ultracentrifugation at 145,000x g and 4 °C for 1 h. Protein concentration in lysates was determined by the Bradford assay according to the manufacturer's instructions. The lysate was aliquoted and stored at -80 °C for further use.

### **2.2.5 Affinity purification and protein digestion**

Kinobead pulldowns were performed as described previously (75). Briefly, cell lysates were diluted with equal volumes of 1x compound pulldown (CP) buffer (50 mM Tris/HCl pH 7.5, 5% glycerol, 1.5 mM MgCl<sub>2</sub>, 150 mM NaCl, 20 mM NaF, 1 mM sodium ortho-vanadate, 1 mM DTT, 5 mM calyculin A and protease inhibitors). Lysates were further diluted if necessary to a final protein concentration of 5 mg/ml using 1x CP buffer supplemented with 0.4% NP-40. Kinobead (100µl suspension) was incubated with lysates (total of 5 mg protein) at 4 °C for 4 h. Subsequently, beads were washed with 1x CP buffer and collected by centrifugation. Bound

proteins were eluted with 2x NuPAGE® LDS Sample Buffer (Invitrogen, Darmstadt, Germany) and eluates were reduced and alkylated by 10 mM dithiothreitol (DTT) and 55 mM iodoacetamide (IAA). Samples were then run into a 4–12% NuPAGE® gel (Invitrogen, Darmstadt, Germany) for about 1 cm to concentrate the sample prior to in-gel trypsin digestion.

### **2.2.6 In-solution protein digestion**

Universal Proteomics Standard set 2 (UPS2, Sigma-Aldrich, Germany) was denatured with 8 M urea 50 mM Triethylammonium bicarbonate (TEAB) pH 8.5 for 5 min followed by thiol groups reduction by incubation with 10 mM DTT (final conc.) for 1 h at 56 °C. The resulting cysteine residues were alkylated with 55 mM iodoacetamide (IAA) at room temperature, in the dark for 45 min. The reduced and alkylated proteins were diluted into 2 mM urea with 50 mM TEAB, pH 8.5 and digested overnight with trypsin (1:100, trypsin/substrate ratio) at 37 °C. The reaction was stopped by adding 1% formic acid (FA).

### **2.2.7 In-gel protein digestion**

To avoid keratin contamination, the modified in-gel digestion procedures were performed in the clean room as described (99, 100). Briefly, the protein containing lanes were cut and destained with 50% 50 mM TEAB/50% ethanol. Gel pieces were dehydrated with ethanol and rehydrated with 5 mM TEAB containing trypsin (1:100 wt/wt) and incubated overnight at 37 °C. Reaction was stopped by adding 5% FA and the reduced/alkylated residual peptides were extracted by incubating twice with 1 % FA and one time with 60% acetonitrile (ACN) in 0.1% FA followed by two incubations with 100 % ACN. The extraction was frozen at -80 °C and lyophilized with SpeedVac (UNIVAPO 150 ECH).

### **2.2.8 Dimethyl labeling of tryptic peptides**

Dimethyl labeling of tryptic peptides was performed as described previously (101). After in-gel digestion, lyophilized tryptic peptide was dissolved with 100 µl reaction buffer (100 mM TEAB) and 4 µl of 4% formaldehyde was added (Formaldehyde for light channel, D-formaldehyde for medium channel). After briefly vortexing and spinning down, 4 µl sodium cyanoborohydride (600 mM) was added to the reaction system followed by general shaking for 1 h at room temperature using a tube shaker (Eppendorf, Germany). The reaction was stopped by adding 16 µl of 1%

## ***Kinobead establishment***

---

ammonia. Equal amounts of labeled peptides from two channels were mixed together and the excess was removed reagent using the following STAGE-tips purification method.

### **2.2.9 STAGE-tips purification**

Samples were re-suspended in loading buffer (5% FA). After wetting and equilibration steps, acidified samples were loaded to the “GELoader” pipette tip (C18 material, Proxeon) slowly and followed by washing 2 times with loading buffer. Peptides were eluted 2 times with 20 µl elution buffer (80% ACN/5% FA) and frozen at -80 °C and lyophilized with SpeedVac (UNIVAPO 150 ECH) for further MS analysis.

### **2.2.10 LC-MS/MS analysis**

Nanoflow LC-MS/MS was performed by coupling an Eksigent nanoLC-Ultra 1D+ (Eksigent, Dublin, CA) to a LTQ-Orbitrap XL ETD (Thermo Scientific, Bremen, Germany). Tryptic peptides were dissolved in 20 µl 0.1 % formic acid and 10 µl were injected for each analysis. Peptides were delivered to a trap column (75 µm i.d. × 2 cm, packed with 5 µm C18 resin, Reprosil PUR AQ, Dr. Maisch, Ammerbuch, Germany) at a flow rate of 5 µL/min in 100% buffer A (0.1% FA in HPLC grade water). After 10 min of loading and washing, peptides were transferred to an analytical column (75 µm×40 cm C18 column Reprosil PUR AQ, 3 µm, Dr. Maisch, Ammerbuch, Germany) and separated using a 225 minute gradient from 2% to 35% of buffer B (0.1% FA in acetonitrile) at 300 nL/minute flow rate. The LTQ-Orbitrap was operated in data dependent mode, automatically switching between MS and MS2. Full scan MS spectra were acquired in the Orbitrap at 60,000 resolution. Internal calibration was performed using the ion signal  $(\text{Si}(\text{CH}_3)_2\text{O})_6\text{H}^+$  at  $m/z$  445.120025 present in ambient laboratory air. Tandem mass spectra were generated for up to 8 peptide precursors in the linear ion trap for fragmentation by using collision-induced dissociation (CID).

### **2.2.11 Peptide and protein identification and quantification**

#### **2.2.11.1 Spectra count base label free quantification**

MS raw files were loaded to Distiller (v.2.2.1, Matrix Science, London, UK) for peak processing and peak picking using the following setting: un-centroiding of tandem MS spectra, re-calculating precursor charge state, and minimum ten peaks for tandem MS spectrum. Subsequently, peak lists were searched against protein sequence database IPI human (v. 3.68, 87,061 sequences)



using Mascot (v.2.2, Matrix Science, London, UK) using carbamidomethyl C as a fixed modification and S, T, Y phosphorylation and M oxidation as variable modifications, trypsin as proteolytic enzyme with up to 2 miss cleavages, precursor ion mass tolerance of 5 ppm, fragment ion mass tolerance of 0.6 Da, decoy search enabled. Data interpretation was performed with Scaffold2, (v2.1.03, Proteome Software, Portland). Number of assigned spectra used for spectra counting based label free quantification was filtered using a minimal protein identification probability of 99% and minimal peptide identification probability of 95%. Exponentially modified protein abundance index (emPAI) was directly extracted from the Mascot searching result by an in house developed script (programmed by Dr. Kurt Fellenberg).

### **2.2.11.2 Intensity based label free quantification**

The Progenesis software (version 3.1, Nonlinear Dynamics, Newcastle, UK) was used for intensity-based label-free quantification. Briefly, after selecting one sample as a reference, the retention times of all eluting precursor m/z values in all other samples within the experiment were aligned to create a large list of 'features' representing the same peptide in each sample. Features with 2-6 charges were included for further analysis. Features with 2 or less isotopes were excluded. After alignment and feature filtering, replicate samples were grouped together and raw abundances of all features were normalized to determine a global scaling factor for correcting experimental variation such as differences in the quantity of protein loaded into the instrument. Briefly, for each sample, one unique factor is calculated and used to correct all features in the sample for experimental variation as described previously (102). Given that multiple MS/MS spectra are frequently collected for the same feature (precursor ion) across all the samples, the precursor intensities were ranked and the MS/MS spectra of the five most intense precursors for each feature were transformed into peak lists and exported to generate Mascot generic files (mgf). The mgf files were searched against the protein sequence database IPI human (v. 3.68, 87,061 sequences) using Mascot (v.2.2, Matrix Science, London, UK). Search parameters were as follows: fixed modification of carbamidomethylation of cysteine residues, variable modification of S,T,Y phosphorylation and M oxidation, trypsin as proteolytic enzyme with up to 2 miss cleavages, precursor ion mass tolerance of 5 ppm, fragment ion mass tolerance of 0.6 Da, decoy search enabled. Search results for spectrum to peptide matches were exported in .xml format and then imported into the Progenesis software to enable the combination of peptide

## ***Kinobead establishment***

---

quantification and identification. Peptides with mascot ion scores less than identity threshold ( $p < 0.05$ ) were filtered out, and only unique peptides for corresponding proteins were used for identification and quantification. For protein quantification, the feature intensities of all unique peptides of a protein were summed up.

### **2.2.11.3 Dimethyl labeling for protein quantification**

Maxquant, version 1.1.1.25 (<http://www.Maxquant.org/>) (62), was used for the dimethyl labeling based protein quantification. All MS raw files were processed with Maxquant to generate the peak lists and searched against the IPI human (v. 3.68, 87,061 sequences) combined with 262 common contaminants and concatenated with the reversed versions of all sequences using the Andromeda search engine (54). Cysteine carbamidomethylation was set as a fixed modification. S, T, Y phosphorylation and M oxidation were allowed as variable modifications. Peptide identification was based on a search with an initial mass deviation of the precursor ion of up to 7 ppm. The fragment mass tolerance was set to 20 ppm on the  $m/z$  scale. The minimum required peptide length was 6 amino acids and up to 3 missed cleavage sites and three isotopically labeled amino acids were permitted. The accepted FDR was 1 % for both protein and peptide identifications.

### **2.2.12 Statistical analysis**

The data analysis was in collaboration with Dr. Amin Moghaddas Gholami, Kuster Lab, TUM. To detect differential kinase protein expression between multiple cell lines, a moderated linear model (F-statistic) or anova model (anova test) was applied using the limma package (103) in Bioconductor (104). G-test was performed as previously described (105) for spectra counting and emPAI based quantification data. p-values were adjusted for multiple testing to control the False Discovery Rate (FDR) at 5% using the algorithm of Benjamini and Hochberg (106). Adjusted p-values were used to identify differentially expressed kinases.

## **2.3 Result and Discussion**

### **2.3.1 Kinobead establishment**

The first step of this project was to establish the Kinobead technology, which was originally developed by Kuster and colleagues at Cellzome (75). Fig. 11a shows the scheme of Kinobead with the ATP mimics immobilized via either  $-NH_2$  group or  $-COOH$  group. Since low density

## Kinobead establishment

coupling results in low capacity while high density coupling leads to high background due to the unspecific binding, the coupling concentration of individual compounds was optimized for the Kinobead matrix based on the original protocol. Each compound was coupled with 1  $\mu\text{mol/ml}$ , 2  $\mu\text{mol/ml}$  and 4  $\mu\text{mol/ml}$  followed by pulldown experiments on placenta samples to investigate the affinity purification efficiency and unspecific binding. Two compounds are illustrated in Fig. 11b as examples. For compound TUM001, all three coupling densities allowed for effective fishing of many proteins. The intensities of many bands were mildly increased from 1 to 4  $\mu\text{mol/ml}$ , but accompanied by a strong increase of many unspecific bands. To avoid the high background of unspecific binding, 1  $\mu\text{mol/ml}$  coupling concentration was used for this compound.

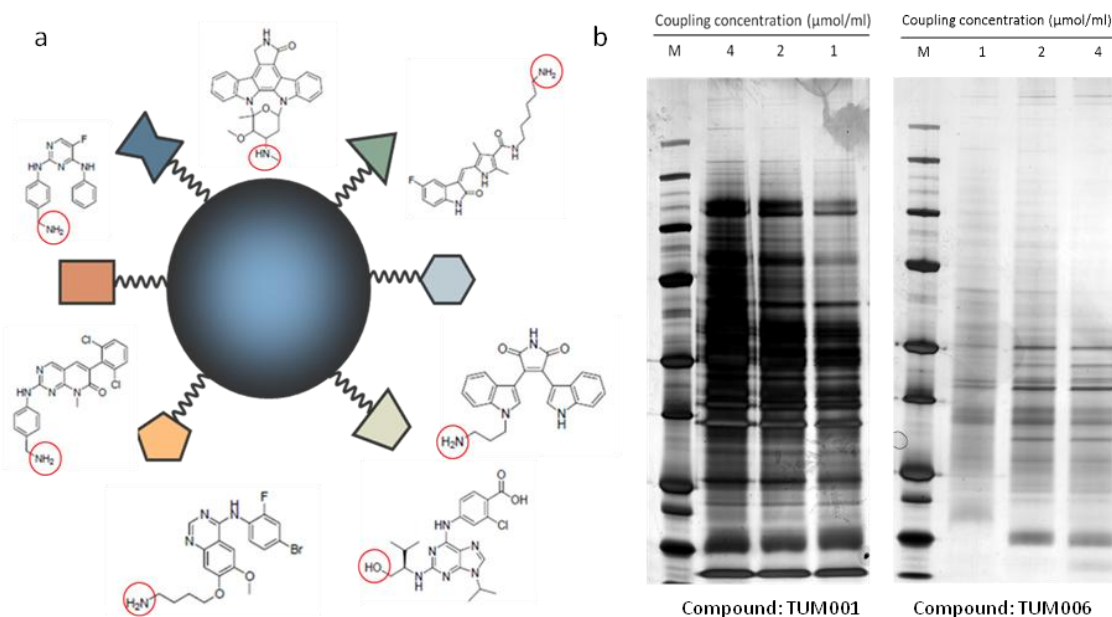


Figure 11. a) Schematic of the kinobeads. Red circles indicate the attachment site. b) Eluted proteins from the beads coupled with different concentration of compound 1 and 6. Pulldowns were performed with 5 mg total protein from human placenta. The elution was separated by SDS-PAGE gel and followed by silver stain.

While, in compound TUM006, a significant increase of the capture efficiency was observed from 1 to 2  $\mu\text{mol/ml}$ , almost no change was found from 2 to 4  $\mu\text{mol/ml}$ . Consequently, 2  $\mu\text{mol/ml}$  coupling concentration was selected for this compound. However, further MS analysis is required to acquire the accurate result, which was not performed here due to the previous study (75). The details of the whole 8 compounds used for the Kinobead are listed in appendix table S2.

## ***Kinobead establishment***

---

### **2.3.2 Detergent selection**

Detergents are amphipathic molecules that possess both a hydrophobic (water-fear) and a hydrophilic (water-friend) group that allow them to spontaneously form highly organized spherical structures called micelles, which have similar properties as the membrane lipid bilayers. Consequently, it can be used to release the proteins by creating a mimic of the natural lipid bilayer. However the micelles only begin to form at a certain concentration which is known as the critical micelle concentration (CMC). This is an important property for selecting detergents, besides varies further conditions including pH, ionic strength, temperature as well as the presence of protein and lipid. Overall, there are three classifications: Non-ionic, ionic and zwitterionic detergents. Keeping proteins biologically active is absolutely required to enable the kinases to bind to the Kinobead. Hence ionic detergents will not be considered here, since they result in denatured proteins. In this study, five non-ionic detergents and one zwitterionic detergent were compared in order to enable the inhibitor matrix to capture more kinases and to decrease the unspecific binding. Equal numbers of K562 ( $10^7$ ) cells were lyzed with the same lysate buffer but containing different detergents, 0.8% NP-40, 0.5% Triton-100, 0.8% Brij35, 0.5% Tween-20, 0.8% digitonin and 1% CHAPS, followed by Kinobead pulldown to enrich the kinases which were subsequently analyzed by MS (in duplicates). Overall, 98 to 108 kinases were identified from each individual pulldown. Although slightly difference of total kinase identification was observed among different conditions, 0.8% NP40 and 0.5% Triton generally resulted in the most kinase identifications and particular most spectra assigned to kinases (ca. 20% more compared to 1% CHAPS). Beside, both conditions allowed more identification of receptor

*Table 1. Comparison of different detergents. Average number from two replicates is shown here. Data were filtered with 99% minimum protein possibility and 95% peptide possibility. Unspecific binding was calculated by dividing the spectra assigned for non-kinase protein with total assigned spectra*

<b>Detergents:</b>	<b>0.5% Triton100</b>	<b>0.8% Brij35</b>	<b>0.8% NP40</b>	<b>0.5% Tween20</b>	<b>1% CHAPS</b>	<b>0.8% Digitonin</b>
<b>Total Kinase protein</b>	105.5	101	105.5	103.5	101.5	104
<b>Total kinase spectra</b>	1663	1467.5	1602	1445	1336	1400.5
<b>Receptor kinase</b>	14	6.5	14.5	8	13.5	12.5
<b>Receptor kinase spectra</b>	260.5	55	278	99.5	258	200
<b>Non-kinase protein</b>	242.5	187	236	219.5	238.5	249
<b>Non-kinase spectra</b>	1690	1369.5	1509	1595	1511	1774
<b>% of unspecific binding</b>	50.4	48.2	48.5	52.5	53	55.8

kinases (Table1). 1% CHAPS, a zwitterionic detergent, gave similar number of total kinases and receptor kinases, but together with lowest kinases derived spectra and more unspecific binding. 0.8% Brij35 led to the lowest background at the cost of decreasing identifications, especially receptor kinases. Considering the fact that receptor tyrosine kinases locate on the top of the signaling pathways and play more important role in many diseases related signaling pathways, 0.8% NP-40 was selected for the following experiments.

### 2.3.3 Kinobead capacity

To address the Kinobead capacity, 100  $\mu$ L settled Kinobead were empirically used to enrich the kinases from a serial amount of proteins from 1 to 30 mg of placenta lysate. The elute from Kinobead pulldowns was separated by SDS-PAGE gel and stained with coomassie blue. As shown in Fig 12, the intensities of the enriched proteins were increasing with the increase of the starting material from 1mg to 30 mg suggesting that the capacity of 100  $\mu$ L settled Kinobead

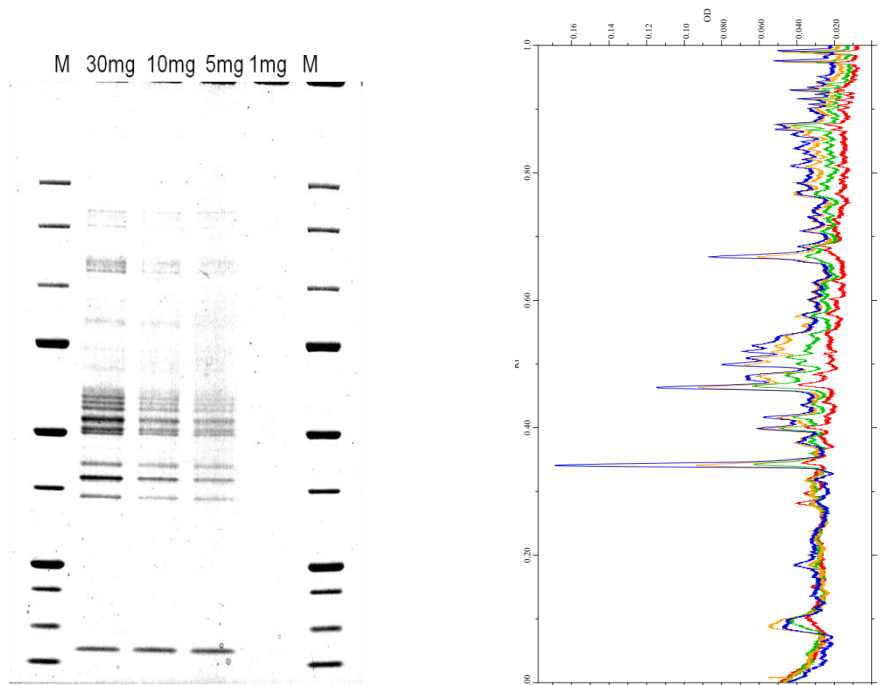


Figure 12. Kinobeads capacity estimation. Left) Kinobeadspulldown was performed by incubating 100  $\mu$ Lkinobeads with indicated different amounts of starting material (placenta lysate). Elution was separated by SDS-PAGE gel and stained with coomassie blue. Right) intensity from each lane was extracted from the left sample and overlaid together. X-axis and y-axis represent the band intensity and band positions respectively. Different samples were coded with color: red, green, yellow and blue indicates samples with 1mg, 5mg, 10mg and 30mg starting material.

## Kinobead establishment

should be enough to capture the kinases from more than 10 mg of lysate. Furthermore, the pattern of most peaks from 10 mg starting material (yellow line) showed approximately two times abundance compared to the corresponding peaks from 5 mg starting material (green peaks) indicating a linear correlation between the peak intensity and starting material. Whereas from 10 to 30 mg, most of the peak intensities were increased by only around a factor 2, suggesting 100  $\mu$ L settled Kinobead has a capacity to enrich the kinases from around 20 mg of lysate. However, this may vary from sample to sample depending on the heterogeneity of the sample. For example lysates from cell lines may be quite different from tissue lysates, which may contain lots of protein originating from blood (e.g., hemoglobin) and/or connective tissue (e.g., fibrin). To be conservative, 100  $\mu$ L settled Kinobead were selected for the pulldown with 5 mg whole lysate protein for further experiments.

The affinity of different kinases to the probes might vary, as a consequence competitive binding to the Kinobead may occur between the kinases: the high affinity kinases may competitively bind to the probes first to block the binding of the kinases with low affinity if the Kinobead capacity is not big enough. To address the problem, a serial volume of Kinobead from 12.5 to 100  $\mu$ L (settled) was used to purify the kinases from 12.5 mg placenta lysate, from which 100  $\mu$ L settled Kinobead should have enough capacity to capture the kinases according to the data described above. By using intensity based label free quantification (described below), 150 protein kinases were quantified out of 168 identified ones. To achieve a robust analysis, only the protein kinases quantified with at least two peptides were counted representing 122 protein kinases, as

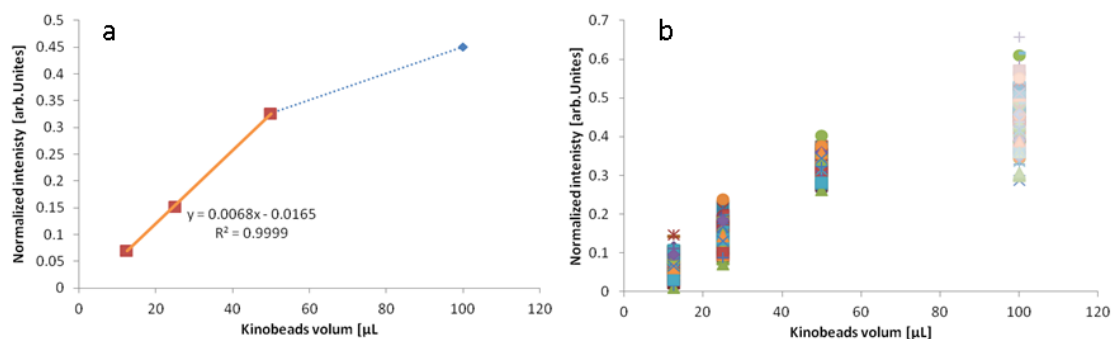


Figure 13. a) Total intensity of total 122 quantified protein kinases performed nice linearity among first three data points. b) Normalized intensity distribution of the 122 protein kinases from 4 pulldowns with indicated volume of kinobeads. Intensities of 119 kinases performed linear increase from 12.5 to 50  $\mu$ L beads. Only the kinases quantified by more 2 unique peptides were plotted here.

illustrated in Fig.13a, the average intensity of total 122 protein kinases performed a good linearity among 12.5, 25 and 50  $\mu\text{L}$  beads but not with 100 $\mu\text{L}$  beads suggesting some protein kinases have been completely depleted by using 100 $\mu\text{L}$  beads. For the individual kinases, 119 out of 122 protein kinases showed a linear behavior with  $R^2$  more than 0.96 with the increase of Kinobead from 12.5 to 50  $\mu\text{L}$  (Fig.13b) illustrating that no obvious competition among the different protein kinases was observed even in the samples with 12.5 and 25  $\mu\text{L}$  Kinobead (certainly having not enough capacity for catching all the kinases from 12.5 mg lysate). Together, these data indicated that the Kinobead matrix used here covered these quantified kinases without big variance of affinity, therefore the amounts of individual proteins captured by the probe matrix do not relate to the affinities of these proteins for the immobilized compounds but reflect their amounts present in the sample.

### **2.3.4 Concentration of phosphatase inhibitors**

From the previous research, some kinases have been shown to bind to the Kinobead only in an active conformation, such as Focal adhesion kinase (FAK/PTK2) (75). In this experiment, 14 head and neck cell lines lysed with 2x phosphatase inhibitor I & II were profiled with Kinobead (Fig. 14). However, comparison of the full lysate and Kinobead pulldown eluate by western blotting indicated that for one kinase, CSK, the MS profiling result was not consistent with full lysate (Fig.14a). Therefore it was hypothesized that active CSK is required to be captured by Kinobead, just the same as for FAK. CSK in the lysate might be inactivated due to the low concentration of phosphatase inhibitors (PPI). To validate this hypothesis and systematically investigate whether there are more kinases which bind to the Kinobead in active manner, the concentration of phosphatase inhibitor which can keep the kinase active by inhibiting the phosphatase activity was optimized. Two cell lines lysed with same buffer but different concentrations of phosphatase inhibitors (PPI) were used to compare the impact of PPI on the Kinobead pulldown product. As expected, the result confirmed that the FAK binding is activation dependent and two kinases CSK and TBK1 were found to show a similar behavior. Almost no CSK and TBK1 were fished by the Kinobead pull down from the lysate with low concentration of PPI (2x) but high amounts were observed from the super high PPI (5x) samples (Fig. 14b). Furthermore, when the same panel of samples from Fig. 14a was reanalyzed with super high concentrations of PPI, as shown in Fig. 14c, a comparable result was observed between Kinobead pulldown elution and whole lysate by

## Kinobead establishment

western blot suggesting that super high PPI is necessary for the Kinobead pulldown to achieve high fidelity, therefore high PPI concentrations were applied in all the Kinobead based experiments within this project.

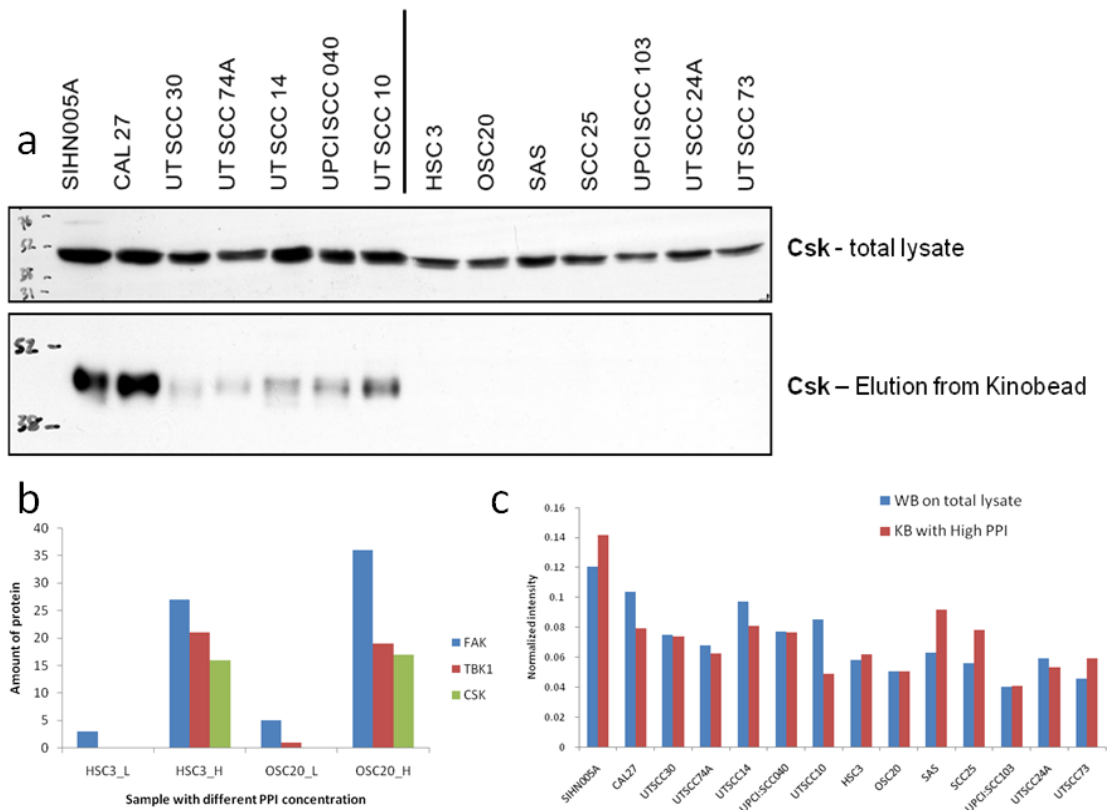
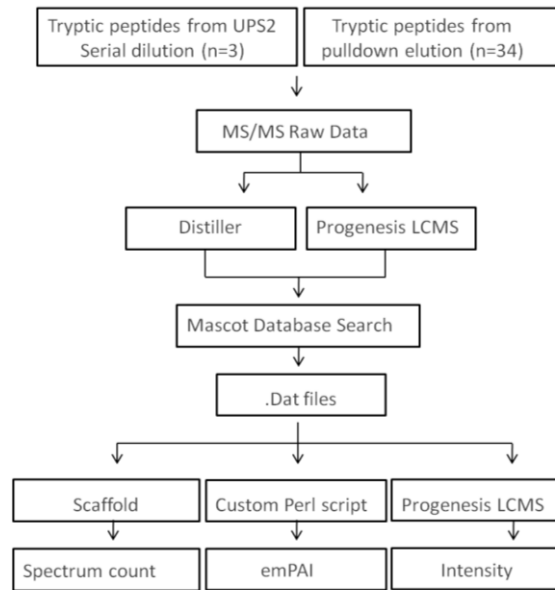


Figure 14. Phosphatase inhibitor effect on capture efficiency of Kinobeads a) Comparison of CSK from the Kinobeads elution and corresponding cell full lysate. Only two cell lines were in agreement with each other. b) Three kinases were identified only in the samples lysed with super high concentration of PPI by Kinobead pulldown but not the same samples lysed with a low concentration of PPI. c) Target profiling performed on the indicated cell lines lysed with super high PPI which was congruent with the full lysate.

### 2.3.5 Comparison of spectra counting based and intensity based label-free quantification

As described in the introduction, isotopic labeling based quantification has been widely used in quantitative proteomics research, but it is not suited for large-scale biomarker discovery mainly due to either difficult to label sample or limited available mass colors. Therefore, before making the decision on the quantification method for targets discovery, some comparisons of different label free quantification and also label free versus isotope labeled quantification methods were performed. First, the serial dilutions of the Universal Proteomics Standard set 2 were utilized to elucidate the performance of intensity or spectra count based label free quantification, which





*Figure 15. Comparison of different label free quantification methods. USP2 dilution samples were measured in triplicate for the properties evaluation of three label free quantification methods. And also one panel of 34 tumor cell lines was measured in unicate for data distribution plot (see below).*

contains 48 human proteins with dynamic range concentrations from 500 amoles to 50 pmoles. The raw files were either processed by Progenesis-LCMS (nonlinear, UK) for intensity based label free quantification or Mascot Distiller for spectral counting. Another approach called emPAI (107), which is also based on the spectral count but using the adjusted values was also compared here (Fig 15). As illustrated in Fig 16, with the spectral counting approach, all proteins showed a saturation behavior with increasing protein amounts. In contrast, the adjusted emPAI showed a much better linearity, especially for the high abundant and high molecular weight proteins which can generate more observed peptides. Intensity based label free quantification also enabled well linear quantification for the low abundant proteins and the response of the system was linear over more than two orders of magnitude. In addition, it comprises a lot fewer missing values than emPAI and spectral counting, which have a significant effect on the conclusions when performing subsequent statistical analysis.

In order to choose the right statistical method for the following data analysis, one important parameter one needs to know is how the data are distributed. 34 Kinobead pulldown products from head and neck cancer cell lines were analyzed by three different quantification approaches. As shown in Fig. 17a, only the intensity data performed as a normal distribution, emPAI and spectral counting followed a chi-squared distribution. According to the different data

## Kinobead establishment

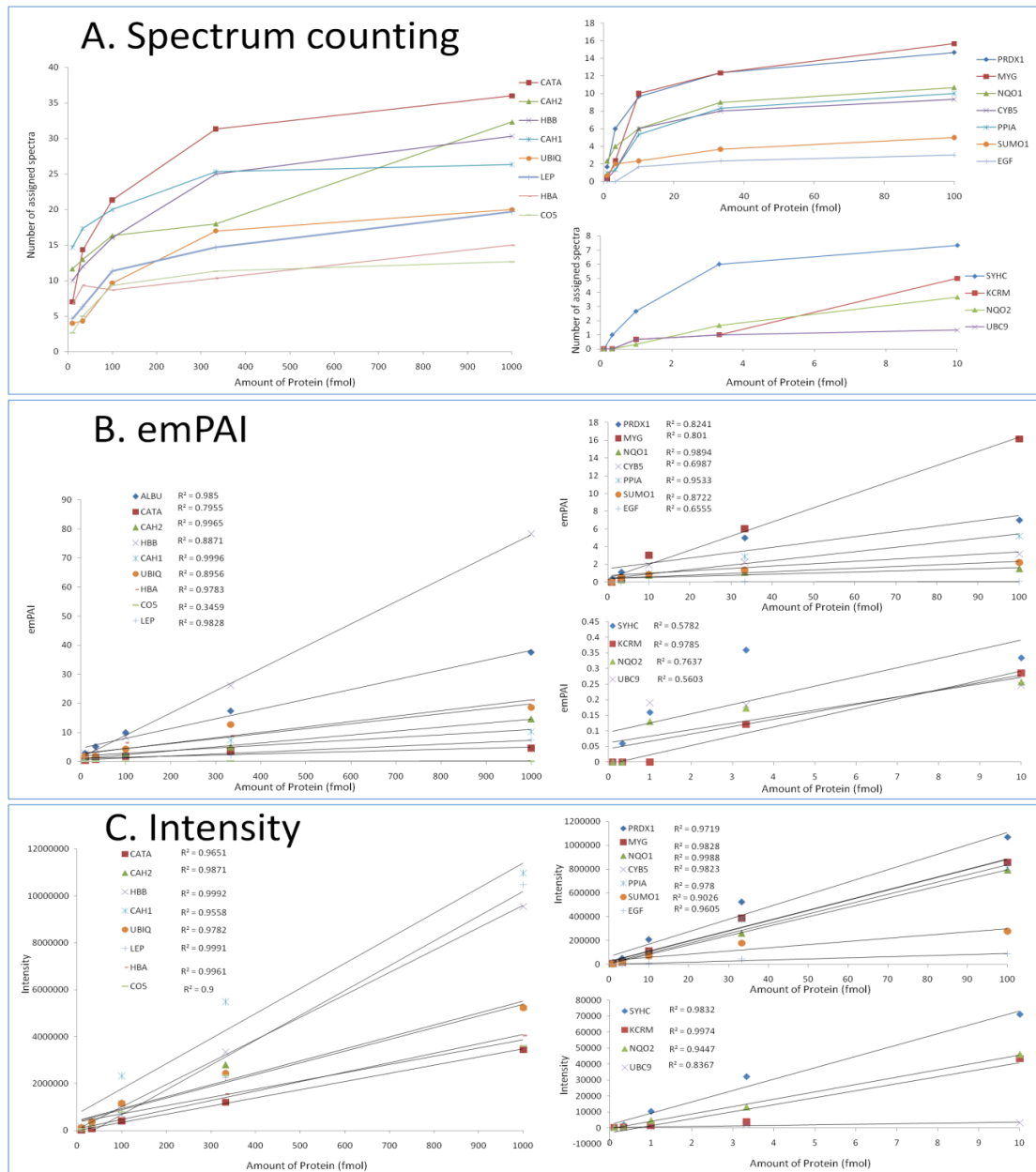


Figure 16. Comparison of three different label-free protein quantification methods. A) Spectrum counting: all proteins show saturation with increasing protein amounts. B) emPAI exported from Mascot search result performs much better than spectrum counting. C) Intensity: protein intensity is the sum of all identified peptide intensities of a given protein. This shows the best linearity for all proteins and all analyzed quantities.

distributions, three statistical tests on the different dataset were performed, more specifically: ANOVA and F-statistic were performed for intensity data and G-test was performed on spectral count data. The result suggested that many kinases out of a total of 160 kinases were

differentially expressed between the 34 cancer cell lines. The G-test identified 92 kinases whereas ANOVA and F-statistics identified 85 and 86 kinases respectively (Fig. 17b), which were quite comparable with each other. Given the fact that 50% of all data points in spectral counting test led to the suspicion that this may result from the more instable spectral counting quantification data compared to MS intensity values. In contrast, only 0.5% missing values are present in the MS intensity data. Based on this observation, moderated F-statistic was selected for the following intensity based data statistic analysis.

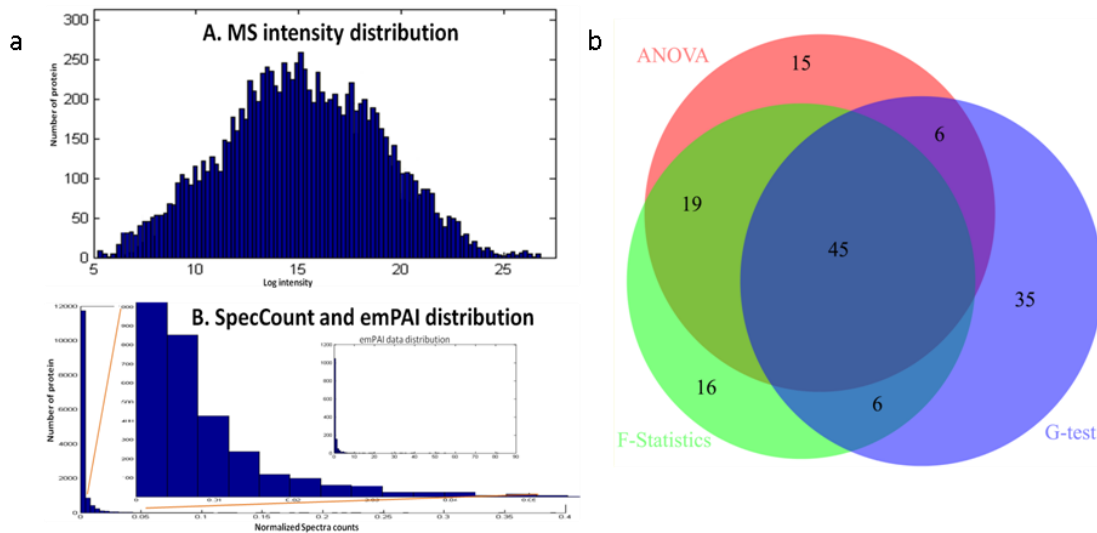


Figure 17. a) Data distribution plot generated from three different approaches. Label-free quantification of Kinobead affinity purifications show that MS intensity data is normally distributed while spectrum counting and emPAI data follow a chi-square distribution. b) Different statistic methods performed on the different datasets revealed many significant candidates.

### 2.3.6 Comparison of the dimethyl labeling and intensity based label-free quantification

For most of the biomarker discovery, metabolic labeling with isotopically labeled amino acids is scarcely possible not only because of the large number of samples but also because the clinical tissue samples cannot be labeled in this way. While the chemical isotope labeling based approach allows the researchers to label the clinical samples and still represents a relative quantification method with more accuracy from the former data. To evaluate the performance of the intensity based label free quantification, a comparison between the intensity based label free quantification and one of most often used isotopic chemical labeling approach, dimethyl labeling, was performed here. Kinases were purified from 5mg placenta lysate by Kinobead pulldown with triplicates. After in-gel digestion, half was used for label-free quantification and the other half

## Kinobead establishment

was labeled with isotopomeric dimethyl labels and followed by MS analysis. For the label-free quantification, protein intensity was extracted by Progenesis LCMS from the raw files while Maxquant was used for the dimethyl labeled samples. Generally, the proteins quantified with higher intensity were observed to be closer to 1:1 ratio (Fig. 18). Dimethyl labeling quantification still represents a more accurate relative quantification method with a standard deviation of the log 2 ratio 0.42 compared to 0.87 from the label free quantification. However, if just kinases which are supposed to be enriched by binding specifically are considered, the standard deviations were reduced to 0.32 and 0.45 respectively, making them more comparable with each other. For the identification, in total 335 proteins were identified including 94 kinases from three replicates of dimethyl labeled samples, while almost two times more proteins were identified from the same three replicates by label free quantification approach containing

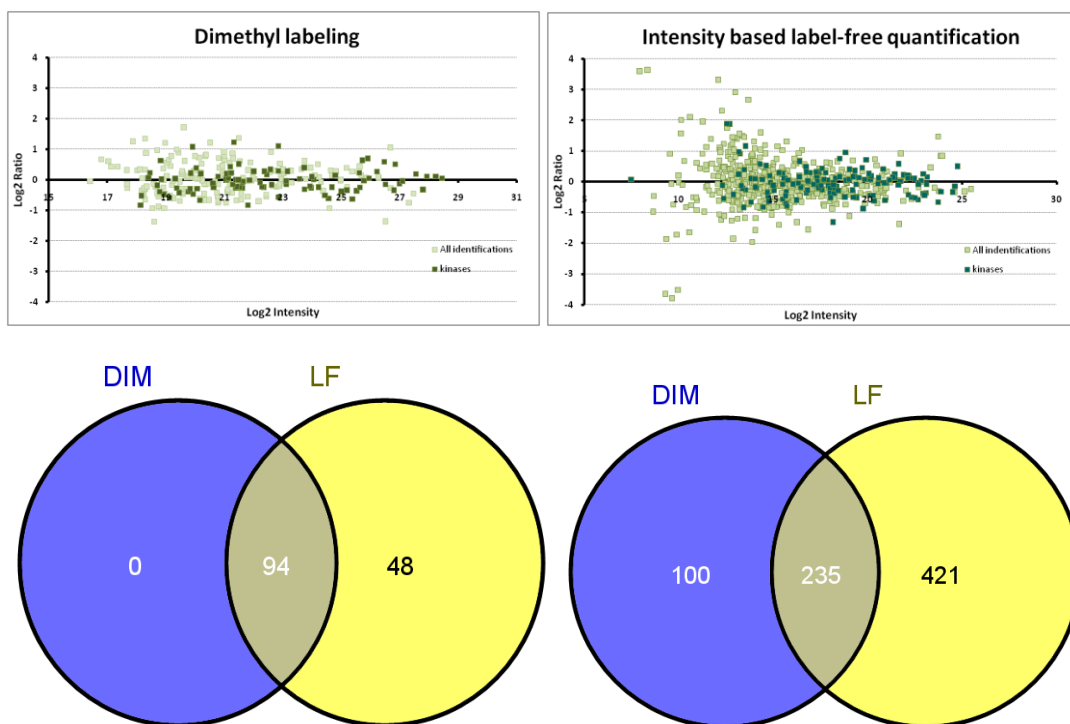


Figure 18. Protein quantification and identification comparison of dimethyl labeling and label free for chemical proteomics. Two samples with 1:1 mixture were quantified by either dimethyl labeling or intensity based label free quantification. Upper panel: plots of the results from two different quantification approaches. Most of the proteins quantified with less than two fold changes by both methods, but dimethyl labeling generally gives more accurate quantification results. Lower panel): Venn diagram comparing the identification of protein kinases (left) and total proteins (right) from the two methods.

additional 48 kinases compared to the dimethyl labeled sample. At least two reasons can lead to the less identification observed: first, extra steps were required to label the sample, such as the STAGE-TIP purification which can lead to the loss of tryptic peptides. Second, dimethyl labeling increased the complexity of the MS spectrum; most of the peptides had a pair after the labeling. Together, the intensity based label-free quantification approach resulted in slightly less accurate quantification but considerably more identifications. To further investigate how well the intensity based label-free quantification method can achieve, three biological and technical replicates were performed as described in the following section.

### **2.3.7 Viability of intensity based label-free quantification for chemical proteomics**

As mentioned in the introduction, compared to the isotope labeling based quantification methods, label-free is the least accurate quantification technique because all the independently performed experimental steps may introduce systematic and non-systematic variations between experiments which are finally reflected in the acquired data (64). However, the data described in the previous section also suggested that label-free can result in comparable quantification results and higher identification rates. Therefore, the intensity based label-free quantification approach was selected for the next project and the technical merits of the overall approach were evaluated before applying this strategy to real samples. For biological replicate analysis, three independent Kinobead pulldowns were performed on the cell lysate from the same cell line CAL27 and analyzed by LC-MS/MS and subsequent intensity based label free quantification by Progenesis-LCMS. The remainder of the tryptic peptides of the biological replicates were pooled and analyzed in 3 MS runs for technical variance evaluation. First, in term of the identifications, as expected, more peptides were identified for a given protein. In addition proteins representing a higher abundance or molecular weight will be more robustly found, which is the fundamental principle for spectra counting based label-free quantification (Fig. 19a). 95% of these proteins were identified with more than 3 peptides reproduced in all 3 replicates. Only around 60% of the proteins identified with one peptide can be found in the three replicates. A direct consequence is the effect of missing data which may generate serious effect on the statistical results. Then, the same file was processed with progenesis LCMS to evaluate the viability of using the intensity. As shown in Fig.19b, 8% and 4% of the overall variance were from biological and technical replicates respectively, suggesting that the intensity based quantification is a robust way for label-free

## Kinobead establishment

quantification. Furthermore, the relation between the proteins identified by different number of peptide and protein intensity variance was also investigated (Fig.19c). Generally speaking, coefficient of variations (CVs) and standard deviations of protein quantification decreases with increasing number of the peptides assigned to the proteins suggesting variation of protein quantification is a function of the number of peptides used for quantification. Even for proteins identified by just one peptide in biological replicates, quantification CVs is rarely larger than 20%.

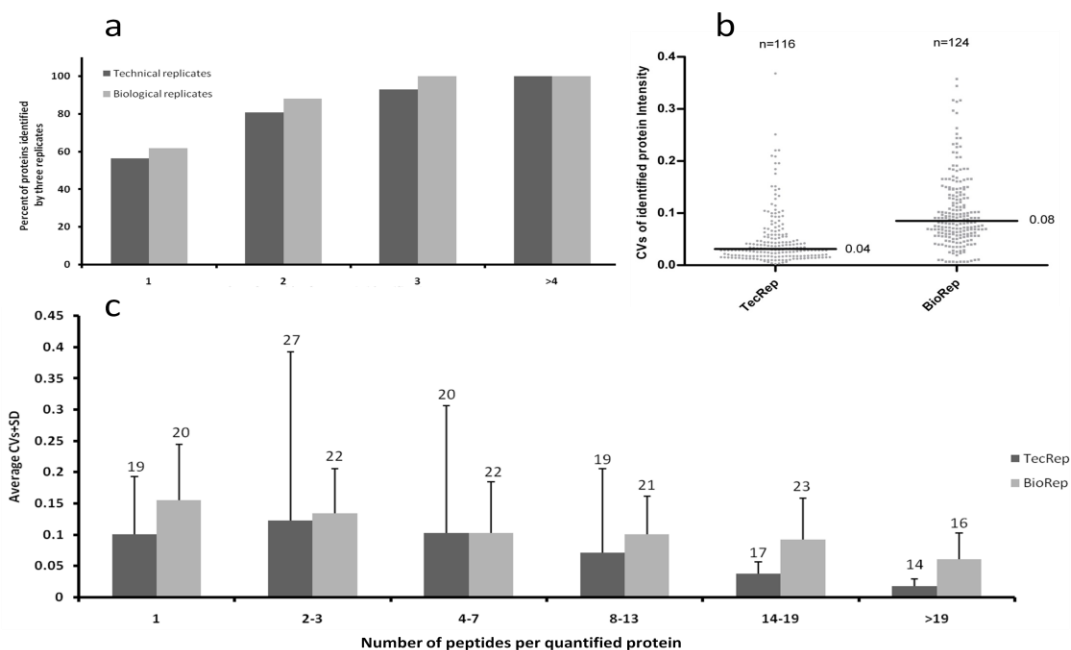


Figure 19. Robustness of protein identification and quantification. a) Reproducibility of protein identification in biological (grey) and technical (black) triplicates using CAL27 cells. Here, no alignment of  $m/z$  and retention times was used for protein identification. All proteins with four or more peptides are identified in all replicates and 80% of all proteins reproduce in all replicates if two peptides are found for a protein. b) Variation in protein quantification. CVs of the quantified proteins from the three replicates were plot here; the black line defines the mean values of all the protein CVs. c) the average CVs with their standard deviation for the groups with different number of peptides used for quantification The number above the error bars indicate the number of proteins that are present in each bin.

## 2.4 Conclusion

Kinobead technology in conjunction with quantitative mass spectrometry is a kinase-centric proteomics strategy enabling to profile the expression of many kinases in parallel. In this study, Kinobead technology was established based on the previous work (75). A serial of parameters has

been established for the following experiments, including the compounds coupling density, detergent, Kinobead capacity, as well as the concentration of phosphatase inhibitors. The results suggested that super high phosphatase inhibitor concentrations are indispensable to retain the activity of some kinases and enable them to bind to the Kinobead. In addition comparison of three label free quantification approaches suggested that Intensity-based quantification clearly outperformed spectral counting and emPAI with the least missing values and best overall linearity. Besides that, the normally distributed dataset allowed using the F-statistics for the statistic analysis. Comparison of the intensity based label free quantification and isotopomeric dimethyl revealed that the former approach showed more protein identifications and comparable quantification with the latter one. Finally, the systematic evaluation of the technical merits of the overall approach suggested a robust quantitative result can be achieved by mass spectrometry based kinase-centric quantitative chemical proteomics with variance of 4% in technical and 8% in biological.

### **3. Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer**

#### **3.1 Introduction**

Head and neck cancer refers to a heterogeneous group of cancers that arise in the upper aerodigestive tract, including the lip, oral cavity (mouth), nasal cavity (inside the nose), paranasal sinuses, pharynx, and larynx. Over 90% of these cancers originate from the squamous cell epithelium of mucosal lining and are therefore referred as squamous cell carcinoma (HNSCC). It accounts for 3% of all cancer cases and is the sixth most common form of cancer with approximately 600,000 new cases worldwide every year (108). The survival rate is 40-50% over 5 years, and despite decades of effort, it has not markedly improved (109). So far the major risk factors identified for HNSCC are tobacco use and alcohol consumption, which contribute to 75% of all the HNSCC patients. Both independent risk factors can generate a synergistic effect when combined (110-112). Besides, infection with the sexually transmitted human papillomavirus (HPV) is another well defined risk factor, particularly for the HNSCCs of the oropharynx (109). Besides, some other risk factors have been recognized to associate with HNSCC, such as the habitual risk factors (e.g., chewing of areca nut) (113), environmental carcinogens, and certain inherited disorders (e.g., Fanconianaemia) (114). While, consumption of fruits and vegetable has been shown to be able to reduce the risk of oral cancer (115).

Previous research revealed various distinguishable subgroups of HNSCC at the histological level. However further RNA and DNA profiling results suggested it to be a very heterogeneous disease at the molecular level which can roughly be divided into HPV positive (20%) and HPV negative cases (80%) (109). In the HPV-negative group, high chromosome instability is frequently observed, which consequently results in genetic alterations. Two most recently exome sequencing studies further confirmed that much more mutations are found in the HPV-negative tumor compared to the positive group and several new mutations were identified in both tumor suppressors (e.g., NOTCH1) and oncogenes (e.g., IRF6) in addition to the known HNSCC genes (TP53, CDKN2A, PTEN, PIK3CA, and HRAS) (116, 117). However, still in around 15% of all the HNSCC cases the genome appears to be rather normal and HPV-negative (109). To date, the molecular carcinogenesis information is limited in the HPV-negative group with high chromosome instability, contributing



## *Chemical proteomics reveals new targets in HNSCC*

---

to 65% of the total HNSCCs. A dozen of candidate cancer genes and several established cancer genes in HNSCC have been shown to be involved in most of the cancer-related phenotypes summarized by Hanahan and Weinberg (see the general introduction), including limitless replicative potential (TP53, CCND1 and CDKN2A), self-sufficiency in growth signals (EGFR and TGF-beta), ability to evade apoptosis (PIK3CA and PTEN), invasion and metastasis (CSMD1), and angiogenesis (VEGFR) (109). However, relatively few signaling pathways have so far been established for the pathogenesis of HNSCC. Among these the TGF-beta/SMAD (118-120) and EGFR/PI3K/AKT are the two most researched pathways (121). The latter offers a number of possible therapeutic intervention points particularly at the level of EGFR itself (122), which is amplified and/or overexpressed in many HNSCC cases (123). Quantitative evaluation of EGFR is a considerable marker for clinical outcome correlation (124, 125). Anti-EGFR monoclonal antibody (126) or tyrosine kinase inhibitor (TKI) (124) therapies have shown some clinical benefits; notably for the combination of antibody and radiation therapy (127). In comparison, TKIs have shown rather low response rates, the reasons for which are currently not clear (124). The success of targeting EGFR for HNSCC therapy demonstrates that the improved understanding of the underlying molecular pathways enables to provide for new strategies for HNSCC therapies. However, given the fact that only a subgroup of HNSCC can benefit from the EGFR intervention, there is a great need to identify new molecular targets whose activity may drive this cancer in individuals in whom EGFR does not play a major role. Recently, several other kinase-centric molecular mechanisms have been under investigation. These include aurora kinase A (AURKA) (128), polo-like kinase 1 (PLK1) (129) and c-MET (130) indicating that the observed molecular heterogeneity of the disease may be rooted in multiple kinase signaling pathways and underscoring the need for potential biomarkers and/or therapeutic targets for an individualized approach to the management of HNSCC.

Given the fact that various molecular mechanisms could result in the aberrant kinase function in human cancer, the kinases based signaling pathways are best studied at the protein level and quantitative proteomics methods are increasingly used to address signaling in a systematic fashion (64, 131). The Kuster laboratory has recently developed a chemical proteomics screening method that allows the interrogation of many signaling pathways in parallel (75). It is an affinity purification matrix termed Kinobead consisting of seven immobilized nonselective kinase inhibitors which allows the purification and identification of several hundred kinases and other

ATP binding proteins from cell lines or tissues (75, 132). Initially, the Kinobead approach was developed to profile the selectivity of small molecule kinase inhibitors; it also lends itself to profiling the expression of kinases in cells or tissues (75, 86). Besides that, the intensity-based label-free quantitative mass spectrometry enables the identification and relative quantification of the purified proteins across many biological samples (133). In this study, the integration of Kinobead based chemical proteomics and intensity based label-free quantitative mass spectrometry was utilized to systematically investigate kinases that might represent novel candidate targets for individualized therapeutic intervention and/or candidate biomarkers describing the underlying individual tumor pathology. Quantitative profiling and statistical analysis of 147 protein kinases across 34 tumor derived HNSCC cell lines revealed that 42 kinases showed highly significant differential protein expression. These included cancer associated kinases such as EGFR, c-MET and AURKA but also novel candidates. Small molecular inhibitor and loss of function experiments using siRNA in high- and low- expressing cell lines further identified kinases including EGFR, EPHA2, LYN, JAK1, WEE1 and NEK9 involved in cell survival and proliferation. Among these, EGFR is confirmed as a drug target and EPHA2 is indicated to be novel drug target. Both contribute to around 20% and 15% of the HNSCC cell lines respectively, which were further confirmed with the clinical sample by immunohistochemical analyses showing that high EGFR and EPHA2 expression is detected in a subset of HNSCC tissues. In addition the several significant candidates are the potential targets of the approved potent pan-SRC family kinase inhibitor dasatinib, which significantly reduces some but not all of HNSCC cell lines as well.

### **3.2 Material and methods**

#### **3.2.1 HNSCC cell lines**

**(See chapter 2.2.1)**

#### **3.2.2 Cell culture and harvesting**

Cell culture was performed as described in chapter 2.2.1 and cells were harvested according the protocol described in 2.2.3 except using 5 times phosphatase inhibitor I & II instead of 2 times of each.

### **3.2.3 Affinity purification and protein digestion**

(See chapter 2.2.4)

### **3.2.4 LC-MS/MS analysis**

(See chapter 2.2.9)

### **3.2.5 Peptide and protein quantification and identification**

(See chapter 2.2.10.2)

### **3.2.6 Statistical analysis**

This data analysis was performed in collaboration with Dr. Amin Moghaddas Gholami, Kuster Lab, TUM. To investigate the data distribution and ensure the appropriate application of statistical tools, frequency histograms and quantile-quantile (Q-Q) plots were created for all non-normalized kinase intensities provided by the Progenesis analysis. Data were then normalized by Variance Stabilization Normalization (VSN) (134) which addresses the error structure in the data and stabilizes the variance across the entire intensity range. VSN has previously been shown to be applicable and useful for MS based quantification (135). The power of the VSN methodology is greatest in situations where hypothesis tests are used to detect changes in protein expression. The underlying assumptions of these hypothesis tests are data normality and homogeneity of variance, which tend to be more appropriately represented by VSN-transformed rather than non-normalized data. Briefly, a complete pairwise comparison was performed as follows: Let  $y_kc$  be the expression values for kinases  $k = 1, \dots, K$  and cell lines  $c = 1, \dots, C$ , pre-processed and normalized, then the systematic effect for each kinase can be described by a linear model  $E(y_k) = X\beta_k$ , where  $y_k = (y_{k1}, \dots, y_{kC})^T$  is the vector of expression values for kinase  $k$ ,  $X$  is a known design matrix with full column rank  $R$  and  $\beta_k = (\beta_{k1}, \dots, \beta_{kR})^T$  is a kinase-specific vector of regression coefficients. Regression coefficients represent comparisons of interest between cell lines in the experiment. These coefficients were estimated with the least squares linear model fitting procedure and tested for differential expression (testing any particular  $\beta_{kR}$  equal to 0) by moderated Student's t-statistics via the empirical Bayesian statistics described in the limma package (103). The null hypothesis was accepted or rejected on the basis of P-values computed for the omnibus F-statistic via limma as described above, at a specified significance level. P-values were adjusted for

## ***Chemical proteomics reveals new targets in HNSCC***

---

multiple testing to control the False Discovery Rate (FDR) at 5% using the algorithm of Benjamini and Hochberg (106). Adjusted p-values allowed identifying differentially expressed kinases.

### **3.2.7 Western blot analysis**

For immunoblot analysis, cells were washed with pre-cooled PBS and lysed in RIPA buffer. Protein concentration was determined by the Bradford assay. Fifty µg of lysate from control and treated cells were mixed with equal volumes of 2x NuPAGE® LDS sample buffer containing 10 mM DTT and heated at 95 °C for 5 min. Proteins were subsequently separated by 4–12% NuPAGE gel and transferred onto PVDF membranes (Invitrogen, Darmstadt, Germany). Membranes were blocked for 1 h in blocking solution (2% BSA in 1x Tris Buffered Saline, TBS, 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 0.1% Tween-20) at room temperature and probed overnight at 4 °C with the respective primary antibody. Immuno reactivity was detected using IgG conjugated peroxidase (GE, Little Chalfont, UK) and visualized by an Image Quant LAS 4000 mini (GE, Little Chalfont, UK). For phosphorylation analysis, membranes were first probed with the respective phospho-specific antibody and then stripped with 62.5 mM Tris-HCl, pH 6.8, 100 mM β-mercaptoethanol, 2% SDS for 30 min at room temperature to allow probing the same blot for the respective total protein.

### **3.2.8 Immunohistochemistry analysis**

This experiment was performed in collaboration with Prof. Dr. Med. Axel Walch, Helmholtz centrum München Institut für Pathologie. The levels of EGFR and EPHA2 were analyzed by immunohistochemistry in tissue microarrays (TMAs). All tissue samples were obtained from patients who were diagnosed with primary HNSCC and underwent a surgical resection. In total, 92 primary formalin-fixed and paraffin-embedded samples from different anatomic subsites were obtained from the archives of the Institute of Pathology, Technische Universität München. All tumor tissue specimens were procured from patients giving written informed consent according to the requirements of the ethics committee of the Technische Universität München. TMAs were generated from these samples in house by sampling one tumor tissue cores (1.0 mm in diameter) from each paraffin-embedded tissue block using the technique pioneered by Kononen *et al.* (136). The TMAs were constructed solely for the purpose of profiling protein expression in HNSCC patients. Therefore, no further clinical data (such as HPV status, smoking habit, alcohol abuse) is available for this patient material. Immunohistochemical staining for EGFR and EPHA2 were

carried out using an automated stainer (Ventana Discovery, Tuscon, AZ, USA) and the DAB Map kit (Ventana) as described elsewhere (137). The tissues were incubated with either an EGFR antibody (Dako EGFR pharmDX TM Kit) or an EPHA2 polyclonal antibody (Santa Cruz, sc-924) in a dilution of 1:250. The EGFR antibody used for TMAs was validated by the vendor. The EPHA2 antibody used for TMAs was validated by single band detection on western blots against a number of human tissues in the Human Protein Atlas project (<http://www.proteinatlas.org/>) and by single band detection on western blots against complete cell lysates of HNSCC cell lines in Kuster laboratory.

### **3.2.9 Sample preparation for DNA sequencing**

This experiment was in collaboration with Dr. Kramer, Kuster lab, TUM. Total mRNA was prepared from cell lines using the magnetic mRNA isolation kit (New England Biolabs Inc., Ipswich, MA, USA). The purified mRNA served as template for cDNA synthesis with anchored polydT primers employing the ProtoScript M-MuLV Taq RT-PCR kit (New England Biolabs Inc., Ipswich, MA). The encoding transcripts were selectively amplified by touchdown PCR. Due to the limited reading length of the sequencing reaction, eight consecutive, overlapping segments were generated by designing eight corresponding PCR primer pairs to cover the full-length EGFR encoding sequence. Individual PCR reactions were prepared in 50  $\mu$ l aliquots by mixing 2  $\mu$ l cDNA (at varying dilutions), 0.2  $\mu$ M primers (biomers.net, Ulm, Germany), 1 U Phusion High-Fidelity DNA Polymerase (New England Biolabs Inc.), 10  $\mu$ l 5x reaction buffer and 200  $\mu$ M dNTP mix. Touchdown PCR was initiated in a Primus25 thermocycler (PEQlab, Erlangen, Germany) at 98 °C for 3 min, followed by 10 touchdown cycles with 30 sec denaturation at 98 °C, 60 °C annealing decreased by 1 °C per cycle for 15 s and 72 °C synthesis for 45 s. After 25 additional cycles at a constant annealing temperature of 55 °C, the reaction was completed by a final extension step at 72 °C for 5 min. The PCR products were purified by agarose gel electrophoresis employing the QIAquick Gel Extraction kit (Qiagen). The DNA sequence of individual PCR amplicons was determined using the ABI Prism 3730 automatic sequencer (Applied Biosystems) and the primers employed for transcript isolation by PCR.

### **3.2.10 Kinase inhibitor treatment**

The EGFR inhibitors gefetinib, erlotinib, the dual EGFR/HER2 inhibitor lapatinib and the pan SRC family inhibitor dasatinib were purchased from LC Laboratories (Woburn, MA, USA) and the c-MET inhibitor PHA665752 was purchased from Tocris Bioscience (Bristol, UK). All inhibitors were dissolved as 10mM stock solutions in dimethyl sulfoxide (DMSO) and kept at -20 °C. HNSCC cells were seeded in 96-well plates at  $3 \times 10^3$  cells/well and grown in DMEM supplemented with 10% (v/v) FBS and 1x NEAA for 24 h prior to experimental treatments. Next, 100  $\mu$ l fresh medium supplemented with different concentrations of kinase inhibitors (range of 40 nM to 10  $\mu$ M in 0.1% DMSO depending on the experiment; control cells were treated with 0.1% DMSO as a vehicle control) were added to the cells. Cell viability was monitored after 96 h of treatment using the XTT Cell Proliferation Kit II (Roche Applied Science, Mannheim, Germany).

### **3.2.11 siRNA treatment**

Small interfering RNA (siRNA) reagents were purchased from Qiagen (Hilden, Germany). The concentrations used in the described experiments were twice the values recommended by the manufacturer (see appendix table S3). Cells serving as negative controls were treated with transfection reagent only; positive controls were transfected with 1 nM AllStarsHs Cell Death Control siRNA (Qiagen Hilden, Germany). 7500 cells were seeded in 96 well plates and incubated in DMEM containing 10% (v/v) FBS and 1x NEAA at 37 °C 1day before transfection. siRNA duplexes were diluted in 50  $\mu$ l DMEM medium without serum, followed by addition of 1  $\mu$ l INTERFERin™ (PolyplusTransfection, Illkirch, FRANCE) and immediate homogenization for 10 s. The mixture was then incubated for 10min at room temperature to allow formation of the transfection complex between siRNA duplexes and INTERFERin™. After exchanging the medium with 125  $\mu$ l fresh pre-warmed complete medium, the transfection mixture was added to the cells and mixed gently by swirling. Cell viability was assessed by the XTT assay after 4 days as described above.

### **3.2.12 Ephrin A1 Fc treatment**

For cell viability assays  $3 \times 10^3$  cells/well were seeded in 96 well plate with 100  $\mu$ l DMEM containing 10 % (v/v) FBS and 1x NEAA and allowed to attach for 24 h. Then, medium with the respective concentrations of EphrinA1-Fc (or 1  $\mu$ g/ml Fc for control cells, R&D Systems,

Minneapolis, MN) was added to the cells and cell viability was assessed by the XTT assay after 4 days as described above. For Western blots:  $10^5$  cells were seeded in 6 well plates containing 1ml DMEM with 10 % (v/v) FBS and 1x NEAA. After 24 h, cells were starved by culturing in serum-free DMEM for another 24 h. Subsequently, cells were stimulated either with  $1\mu\text{g/ml}$  EphrinA1-Fc or  $1\mu\text{g/ml}$  Fc as a control for 30 min. Cells were subsequently washed with PBS and lysed for Western blot analysis as described above.

### 3.3 Results and Discussion

#### 3.3.1 Experimental strategy of identifying differential kinase protein expression in HNSCC cell lines

After the systematic evaluation of the whole Kinobead procedure (see chapter 2), it was applied in this study to discover the therapeutic targets for the head and neck cancer. Established cell lines were used here to avoid considering uncertain factors of the clinical tissue, such as the complicated sample preparation, definition of the cancer tissue and the sample grades, etc. as well as significant issues with candidate validation options for the next step. The strategy employed is illustrated in Fig.20. Briefly, cell lines originally derived from primary HNSCC tumors (30 of 34) or metastases (4 of 34) were grown *in vitro* and subsequently lysed with the lysis buffer containing high concentrations of phosphatase inhibitors to keep the phosphorylation status, which is required for the binding of several kinases as described in the last chapter. Subsequently, kinases were partially purified using Kinobead, which has been shown to allow the purification of 60% of all human protein kinases from cell lines and tissues (75), followed by subjecting to

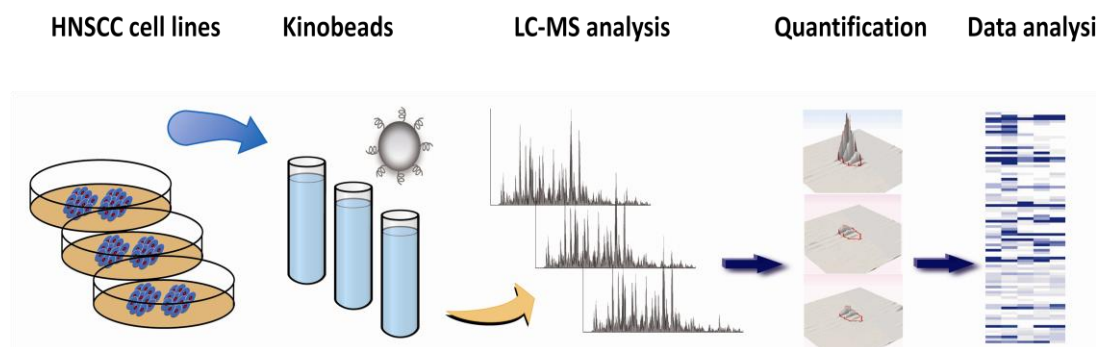


Figure 20. Quantitative chemical proteomics strategy for the identification of kinase drug targets or signaling pathway biomarkers.

## Chemical proteomics reveals new targets in HNSCC

trypsin digestion and nanoscale liquid chromatography tandem mass spectrometry (nLC-MS/MS). Protein identification was achieved by database searching of the tandem mass spectra and label-free quantification was performed using the  $m/z$  and retention time aligned precursor ion intensities of the identified peptides integrated across the chromatographic peak. Moderated F-statistics was employed to identify kinases that are significantly differentially expressed across the 34 cell lines. After highlighting the significant candidates, some additional experiments including small molecular inhibitor and siRNA treatment, were used to validate part of the candidates and clinical data of the head and neck cancer patients was also acquired to connect with the result achieved here.

### 3.3.2 Targets profiling on HNSCC by chemical proteomics combined with quantitative mass spectrometry

In this study, 34 head and neck cancer cell line were profiled in a biological duplicate, generating 68 samples each of which was subjected to a 4 h LC-MS/MS experiment for label-free protein quantification and identification. In total, 17,873 precursor ions and retention time features led to successful peptide identifications. This corresponds to 665 unique proteins including 146 kinases with the latter contributing ~50% of the entire signal intensity. As discussed in the last chapter, consistent sample processing and stable LC-MS system are indispensable for the

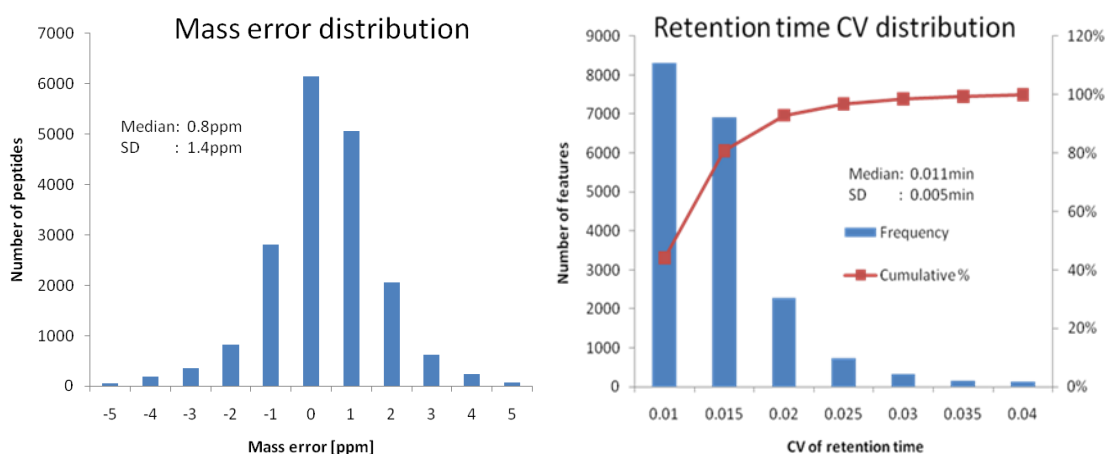
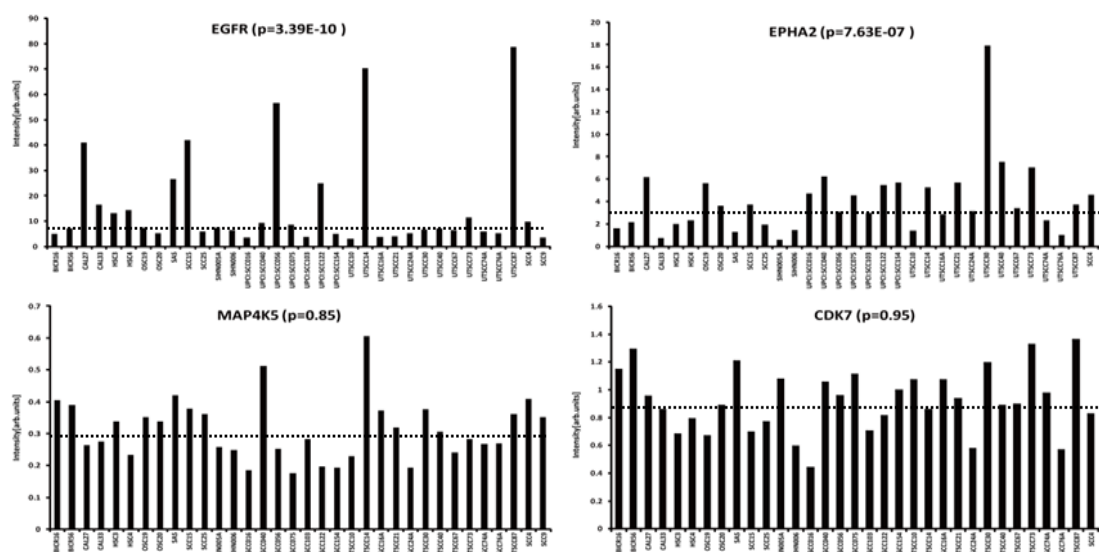


Figure 21. Mass and retention time stability of the employed LC-MS/MS approach. a) Peptide precursor mass error distribution of identified proteins. The median mass error is 0.81 ppm with a standard deviation of 1.40 ppm. b) Retention time stability across 34 experiments. The median CV of all MS retention time features is 0.011+/- 0.005. 42% of all features have CVs of <1% and 93% of all features are within a CV of <2%



successful MS-intensity based label-free quantification. In the analytical set-up here, the reproducibility of protein identification is further improved by the alignment of the data. As long as a protein is robustly identified in at least one of the samples, the alignment of the retention times together with the accurate mass of the peptide precursor will lead to the identification of all peptides in all samples (as long as they are present above noise) (138). Therefore, before performing further data analysis, the systematic variance between the samples and replicates was evaluated. Fig. 21 shows that the retention time distribution in the data is very narrow and the median CV of all identified LC-MS features (i.e. peptides) is 1.1% (42% of all features show retention time CVs <1% and 93% show CVs <2%). The median precursor mass error of identified peptides was 0.8 ppm and 95% of all identified peptides were measured within 1.4 ppm mass error. This demonstrates that the employed LC-MS conditions were very stable and thus well-suited for the intensity based label free quantification purposes. Together with the data from the last chapter showing that there is less than 10% variance in the biological replicates suggested that the intensity based label-free quantification approach used in this study delivers a comparable data quality as stable isotope labeling (133).

For the screening of 34 cell lines (in duplicate) the respective numbers are 5.6% missing values



## Chemical proteomics reveals new targets in HNSCC

for MS intensity and 76.6% for spectral counting, which can have a significant negative impact on the statistical analysis, confirmed the distinct advantage of MS intensity-based quantification over spectral counting described in the last chapter. With the quantification approach applied for the Kinobead experiments here, it is impossible to know the absolute quantities of any of the purified proteins. It should therefore be noted that all the quantification data are relative (i.e. relatively comparing the quantities of one protein across the different experiments is possible but the quantities of different proteins cannot be compared to each other). All the results discussed above suggested that the experimental approach generated robust qualitative and quantitative data for subsequent data analysis. In total, 146 kinases were identified and quantified across all 34 cell lines and as shown in Fig. 22, most kinases were similarly expressed across the panel of cell lines (e.g. MAP4K5 and CDK7, Fig. 2a bottom panel) but some kinases showed marked differences in expression (e.g. EGFR and EPHA2, Fig. 2a upper panel).

### 3.3.3 Statistic analysis

This data analysis was in collaboration with Dr. Amin Moghaddas Gholami, Kuster Lab, TUM. To investigate the data distribution and to ensure the appropriate application of statistical tools, frequency histograms and quantile-quantile (Q-Q) plots were created for all non-normalized kinase intensities provided by the Progenesis analysis. Fig. 23a shows a frequency histogram of

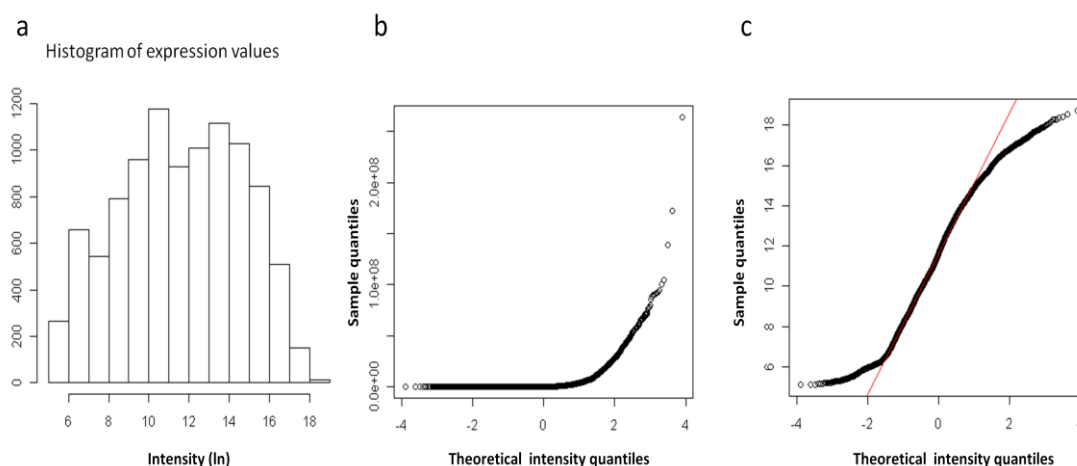


Figure 23. Variance stabilization normalization (VSN) of label-free quantification data. a) Intensity distribution of all identified kinases. b) Q-Q plot of non-normalized sample and theoretical intensities shows that the data are not normally distributed across the entire intensity range. c) Q-Q plot of VSN processed data showing much improved normality (Figure generated by Dr. Amin Moghaddas Gholami, Kuster Lab, TUM).



**3.3.4 Validation of the MS data by Western blot**

To validate the MS profiling result, EGFR, the most differentially expressed kinase in the cell line panel ( $p=3.39E-10$ ) from the top of the statistic analysis list, was selected for the validation experiment. EGFR has been shown to be overexpressed in the majority of HNSCC patients by other groups (14). Western blot analysis, a classic method currently widely used by many groups, was applied here to address how consistent these orthogonal methods are. The same amount of lysate from both batches of 34 cell lines used for targets profiling were separated by SDS-PAGE gel and probed with anti-EGFR antibody after transferring to the PVDF membrane. The band intensities were extracted and normalized across the whole samples according to the reference sample (Fig. 25a). Concomitantly, the MS intensities for all the samples were also extracted by Progenesis LCMS. Both intensities from Western blot and MS were normalized again to the same scale and generated the displayed histogram. Basically the results from these two methods were

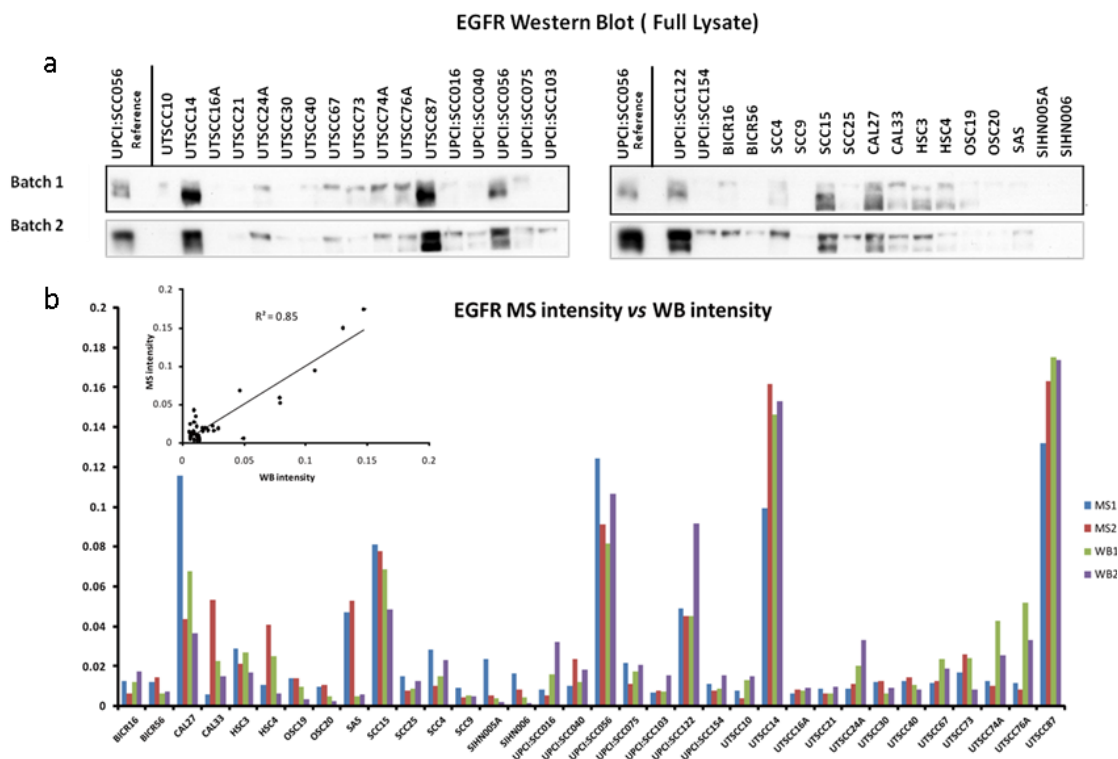


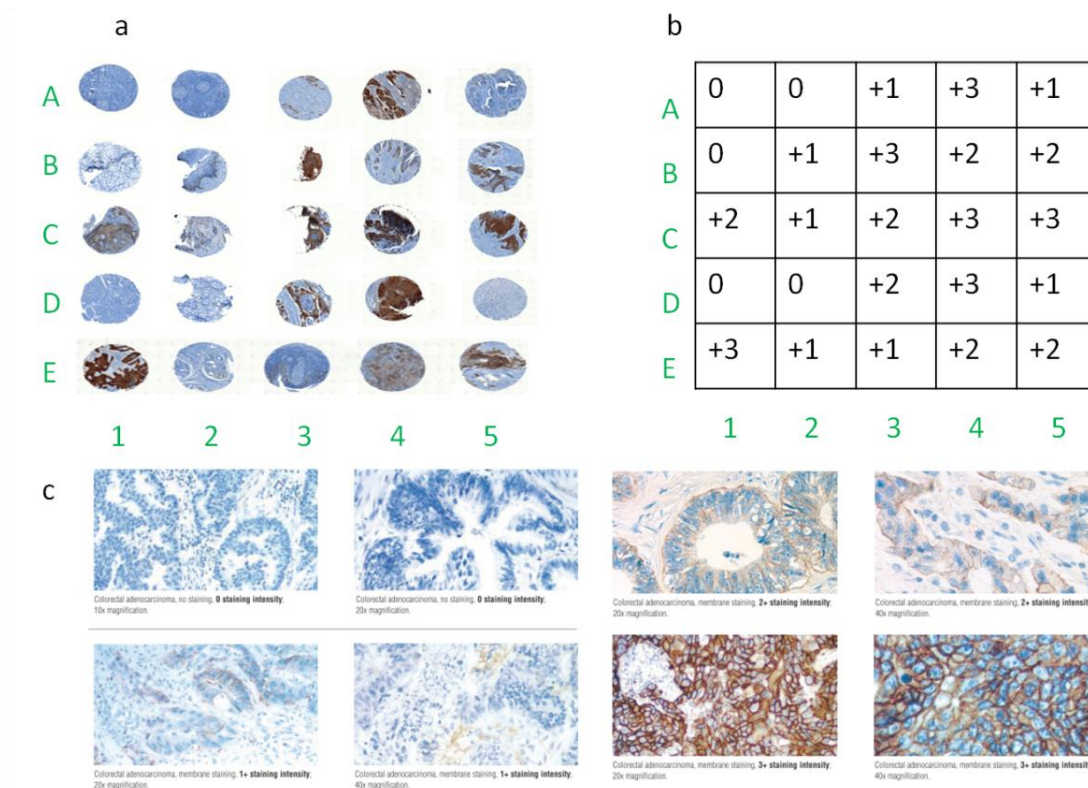
Figure 25. Correlation of protein quantification by Western blotting and MS intensity. a) Two batches of cells were profiled by EGFR Western blotting on lysate. b) Comparison between Western blotting and MS quantification data (after Kinobead purification). The inset shows that both measurements correlate well indicating that the EGFR levels identified in lysates or on Kinobeads are similar.

quite comparable with each other, in addition the correlation of the two quantification methods was found to be very good with  $R^2=0.85$  for EGFR (Fig. 25b).

As a classic method, western blot has been developed for more than 20 years. Despite its obvious success and convenience for many biologists, problems exist with this approach, as it is semi-quantitative, has a limited dynamic range and requires highly specific antibodies. As one more fundamental limitation of this method, using specific antibodies only allows the researchers to focus on their targets. Consequently, unexpected effects in different cellular compartments or in different cellular processes than the one studied will be neglected, which might be the main effects of a given factor (139). In contrast, mass spectrometry-based quantitative proteomics could solve these problems, since it enables researchers to examine global differences in protein expression between for instance healthy versus the diseased states, with a comparable or even better quantification result than that from western blot (140-142), which was also confirmed in this study. All these data together suggests that the mass spectrometry-based quantification strategy applied in this study was a valid quantification approach and subsequently deserved further data analysis on the results.

### **3.3.5 Immunohistochemical analysis of EGFR on clinical sample**

This experiment was in collaboration with Prof. Dr. Med. Axel Walch, Helmholtz Zentrum München Institut für Pathologie. Immunohistochemical staining (IHC) is widely used in the diagnosis of abnormal cells such as those found in cancerous tumors by targeting a specific molecular marker. The other utility of IHC is to understand the distribution and localization of biomarkers and differentially expressed proteins in different parts of a biological tissue (143, 144). Here, 25 clinical formalin-fixed paraffin-embedded HNSCC samples were analyzed by IHC. As shown in Fig.26a, EGFR is differentially expressed in the clinical samples. The expression status was estimated (Fig.26b) according to the instruction supplied by the company (Fig. 26c). In total, 6 out of 25 samples (24%) were found strongly stained representing high expression of EGFR, and there was no obvious signal observed in 5 of these samples. The remaining 14 samples were stained intermediate. Interestingly, the profiling results of the 34 investigated cell lines showed that EGFR expression of 7 cell lines (21%) to be elevated by  $> 3$  fold of the median value, which is quite congruent with the IHC data. Previous research suggested the EGFR to be overexpressed in around 80% of the HNSCC (145). However the overexpression was defined as EGFR mRNA

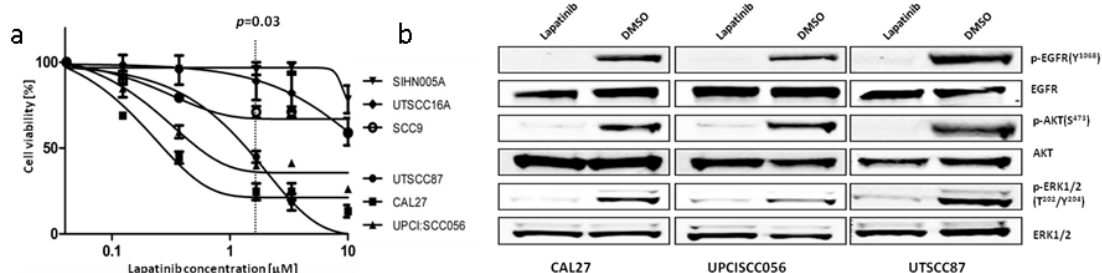


*Figure 26. HNSCC tissue array analysis for the EGFR by immunohistochemistry. a and b) EGFR is highly expressed (+3 on a scale of 0-3) in a subset of 25 HNSCC patient samples. The classification was performed by a trained pathologist. c) EGFR pharmDx™ Staining Intensity guideline for grading. Left upper panels are the examples of negative staining; left lower panels are 1+ staining intensity; right upper panels are 2+ staining intensity and right lower panels indicate 3+ staining intensity.*

median level in the samples to be higher than the control with 2 times the standard deviation (SD), which may not be exactly correlated to the protein level (63). Besides, the EGFR mRNA level from only 7 out of 24 tumors elevated more than 3 fold compared to the corresponding controls (145). Together, these data suggested that only around 20~30% of the HNSCC show a high overexpression of EGFR despite 80% overexpression in total of HNSCC samples (145).

### 3.3.6 Confirmation of EGFR contributing to the proliferation of some but not all HNSCC

EGFR has been shown to be overexpressed in the majority of HNSCC patients (126, 145) and it is the most differentially expressed kinase on the significant list ( $p= 3.39E-10$ ), which was also orthogonally confirmed by the western blot and IHC analysis as described above. To verify the hypothesis that the approach described here would enable to highlight abnormally expressed



*Figure 27. Some but not all HNSCC cell lines are growth dependent on EGFR overexpression and activity. a) High EGFR-expressing cells respond more strongly to lapatinib treatment than low EGFR-expressing cells. b) High EGFR-expressing cells respond more strongly to siRNA mediated EGFR knockdown than low EGFR-expressing cells.*

kinases as drivers for the proliferation of a subpopulation of HNSCC cells, three high EGFR expressing cell lines (CAL27, UPCI: SCC056 and UTSCC87) with three low EGFR-expressing cell lines (SIHN005A, SCC9 and UTSCC16A) according to the profiling result were selected to compare their viabilities in response to EGFR small molecule inhibitors and siRNA interference. Submicromolar concentrations of lapatinib, a highly selective EGFR/HER inhibitor, kills half of the high but not low EGFR-expressing cells (Fig. 27a). The same differential drug response was observed for the EGFR inhibitors, gefitinib and erlotinib (Fig. 28). Next, to investigate whether the observed growth inhibition by Lapatinib correlated with the activity of EGFR and its downstream pathways, the three high EGFR expressing cells were treated with 10 $\mu$ M Lapatinib for 30min prior to cell lysis. The level of EGFR, AKT and ERK activation was determined by Western blot analysis. Consistent with the drug dose-response curves, Lapatinib treatment on the high EGFR expressing cells dramatically reduced the phosphorylation of EGFR. Furthermore, the data suggested that EGFR induced reduction of cell survival and proliferation by downregulation of AKT and ERK, two main components of the survival and proliferation signaling pathways (Fig. 27b). Furthermore, the hypothesis was validated from the genetic aspect by knocking down EGFR protein expression in cells via siRNA. Similar dose-response curves were observed as in the drug treatment experiments: a strong reduction in cell viability occurred in the high but not low EGFR-expressing cells (Fig. 29a). The siRNA knockdown efficiency was examined by Western blot (Fig. 29b). Small molecular inhibitor and small interference RNA represent two fundamental different but complementary approaches against the target proteins. The former approach can specifically bind to a certain domain of the target proteins and consequently inhibit the protein function or

## Chemical proteomics reveals new targets in HNSCC

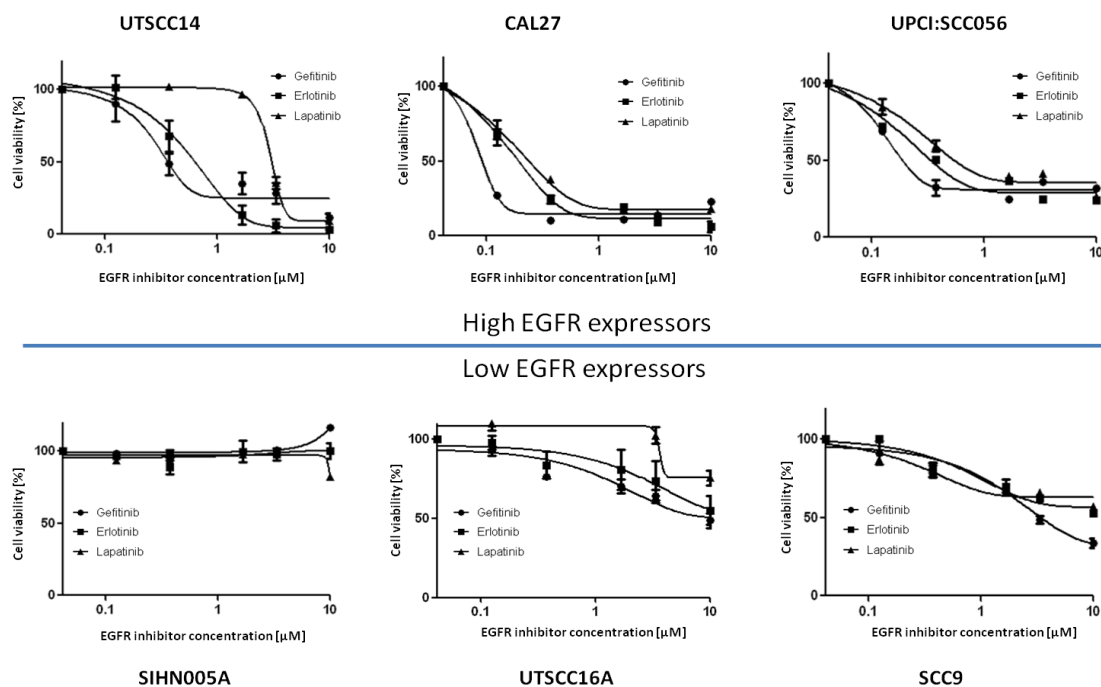


Figure 28. Differential killing of high/low EGFR expression cell lines by EGFR inhibitors. Upper panel: response of high EGFR expressing cell lines to the EGFR inhibitors gefitinib, erlotinib and lapatinib. Lower panel: response of low EGFR expressing cell lines to the EGFR inhibitors gefitinib, erlotinib and lapatinib.

disrupt the protein-protein interaction without altering the expression of the target proteins, while the latter one can recognize the mRNA of the corresponding target proteins and results in sequence specific inhibition of gene expression. Both are promising therapeutic approaches for different diseases including cancer. Numerous of inhibitors and RNAi are under different stages of clinical trials (146, 147). However, both approaches have the potential possibility to target undesired targets (so called off-targets) and cross-reaction with targets of limited sequence similarity by siRNA is observed (75, 148). Therefore, the cross-validation of the cell phenotypes using these two complementary methods provides high confidence conclusions drawn about the biological function and tractability as a drug target of the studied protein. In this study, convergent response of high but not low EGFR-expressing cells to specific EGFR inhibitors and siRNA treatment was observed. These data suggests that the continued very high expression of EGFR is required for maintaining cell survival and the malignant phenotype and furthermore offers an ideal therapeutic target for some but not all the head and neck cancer patients.



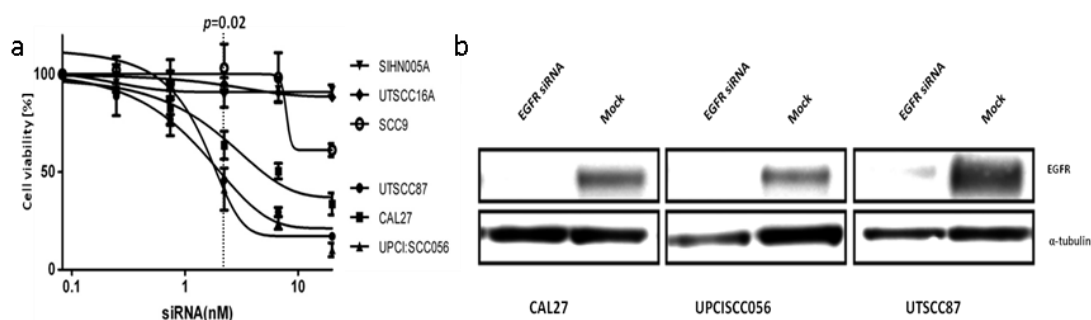


Figure 29. a) Signaling pathway analysis using phospho-specific antibodies reveals that lapatinib treatment inhibits the EGFR/AKT and EGFR/ERK pathways. d) Western blot analysis documenting EGFR knockdown efficiency in CAL27, UPCISCC056 and UTSCC87 cells.

### 3.3.7 Synonymous substitutions in EGFR encoding sequence

This experiment was performed in collaboration with Dr. Kramer, Kuster lab, TUM. Point mutations resulting from single nucleotide substitutions change the property of proteins if it's a non-synonymous (i.e. the encoded amino acid changes), such as EGFR L858R, which consequently causes constitutive activation of the EGFR often found in lung cancer. To address whether this also contributes to the HNSCC samples investigated here, EGFR encoding sequences from five cell lines including one low EGFR expression cell line were analyzed (UTSCC14, UTSCC87, UPCI: SCC056, CAL27 and UTSCC40). In total, four point mutations were found: two of them were in the kinase domain (2361 bp, cag  $\rightarrow$  caa and 2709 bp, act  $\rightarrow$  acc), one in the extracellular domain (1839 bp: gcc  $\rightarrow$  gcX) and the last one was in the C terminal domain (2982 bp: gac  $\rightarrow$  gat). However all of them were synonymous, meaning that no change was induced on the protein sequence level. The facts that pharmacological inhibition and genetic knockdown resulted in the same viability phenotypes and no non-synonymous mutations in EGFR was detected by cDNA sequencing illustrated that EGFR overexpression rather than mutation was a major driver of survival in some, (20-25%) but not all HNSCC cell lines.

Numerous types of cancers including head and neck, lung, oesophagus, stomach, rectum and anus have been revealed to be driven by highly activated EGFR, either via overexpression or mutation (149). In discrepancy with 40% of never-smokers lung cancer patients in Caucasians carrying the mutated EGFR (150), aberrant EGFR signaling as observed in HNSCCs are often due to overexpression rather than due to a activating mutation in EGFR, the later which are reported

in only 1% of the Caucasians (109). In this study, EGFR encoding sequences of five cell lines were sequenced resulting in 4 synonymous mutations, which confirmed the previous research (109).

### **3.3.8 Loss of function screening by siRNA highlights new candidate drug targets**

Due to the huge biological variance, the impact of differential kinases expression on cell viability simply cannot be deduced from proteomic (and transcriptomic) data alone. Further experimental data are required to reveal the consequences of the dysregulated kinases. As discussed above, small molecular inhibitor and small interference RNA are two common tools to investigate the proteins biological function and its tractability as a drug target. Ideally, for the drug discovery, the cellular phenotype can be validated by both methods (151). But, the fact that not so many small molecular inhibitors with high specificity are available, makes small interference RNA (siRNA) the only method of choice to validate and prioritize the candidate list. In particular, these proteins, knocking down of which would differentially affect cell viability in high versus low expressing cells, would represent possible targets for individualized therapeutic intervention. To elucidate if other differentially expressed kinases are also involved in the proliferation of some of the cell lines, a further 22 proteins from the list of differentially expressed kinases ( $p < 0.05$ , Fig. 24a), and 6 kinases for which no strong evidence for differential behavior was obtained from the statistical analysis, serving as a negative control, were subjected to siRNA knockdown in the respective high and low expressing cell lines (see appendix table S3). The data show that for many proteins, loss of expression did not significantly alter cell viability. However, nine kinases (AURKA, EPHA2, EPHB2, EPHB4, JAK1, LYN, NEK9, RIPK2 and WEE1) showed strong (>40%) reduction of cell viability 96h post siRNA transfection suggesting that these kinases play an important role in promoting and maintaining the survival of cells. Interestingly, while EPHA2, NEK9, LYN, WEE1 and JAK1 knockdown killed cells with high but not low expression, AURKA and RIPK2 knockdown killed both high- and low-expressing cells (Fig. 30a and b). These data suggest that at least EPHA2, NEK9, WEE1, LYN and JAK1 might constitute novel targets with therapeutic potential in HNSCC while the other proteins are more generally important for cell survival and/or proliferation. It has to be noted that the result described here only investigated the cell survival and/or proliferation by the XTT assay; other effects of these candidate proteins on other biological functions were not investigated, such as migration and/or invasion. For example, the DDR1 protein has been

## Chemical proteomics reveals new targets in HNSCC

reported to relate to cell migration (152) but the data from this study suggested no effect on the cell viability at all, so it could be still developed as a target for inhibiting cancer metastasis.

Receptor tyrosine kinases (RTKs) refer to a subclass of cell-surface growth-factor receptors with an intrinsic, ligand-controlled tyrosine-kinase activity. Several characteristics of the RTKs, such as being upstream of signaling pathways, their molecular architectures and the important molecular functions of these proteins and their ligands in tumorigenesis, lead these RTKs to be the attractive targets for cancer therapy (153). Together with the scope and the limited time of this project, EPHA2 was selected for the following validation experiments. There was no follow up for the other potential targets highlighted by the analysis (JAK1, LYN, NEK9 and WEE1) because of the scope in this study. However, JAK1 and LYN are established drug targets in a number of cancers (154, 155) and may thus also represent genuine targets in HNSCC. JAK1 (but not LYN) mRNA levels have been found to be upregulated in some HNSCC cases (156). Despite the absence of the systematic verification, NEK9 and WEE1 are very plausible therapeutic candidates for HNSCCs.

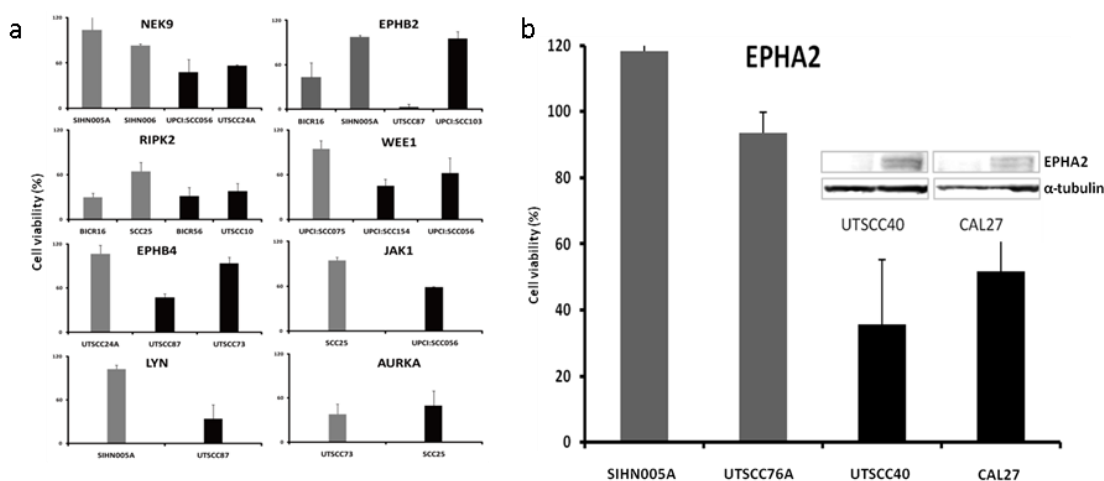


Figure 30. siRNA screen. a) Example results from a siRNA screen of 29 differentially expressed kinases in the respective low (grey) and high (black) expressing cell lines. b) Cell viability of low (grey) and high (black) EPHA2-expressing cell lines following EPHA2 protein knockdown by siRNA. EPHA2 protein levels are shown in the inset.

NEK9 is important in the regulation of mitosis and gene copy number and associated mRNA changes have been detected in some HNSCC cell lines (157). WEE1 is an important regulator of the G2/M cell cycle checkpoint (158) and again, the evidence of increased mRNA levels has been

found in some cases of HNSCC (156). It has recently been shown that WEE1 inhibition leads to the increased sensitivity of osteosarcoma cells to radiation (159). This finding is particularly relevant in the context of the work in this study on the molecular basis for HNSCC as surgery and radiation are the main therapies in this disease. Therefore, radiation sensitization of cells via WEE1 inhibition might further increase the effectiveness of this therapeutic regimen (160).

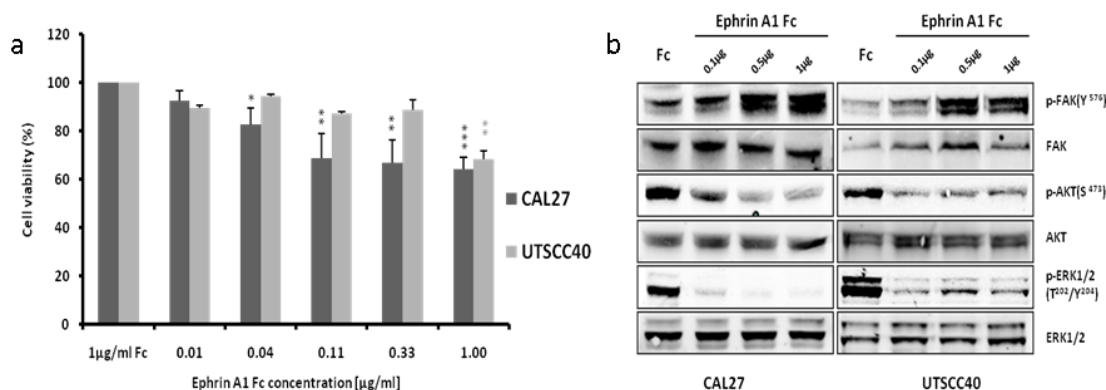
AURKA and RIPK2 showed a strong effect on cell viability, regardless of their relative expression level. Except the possibility that they may be generally important for the cell proliferation, there might be some other interpretations. For example, given that the samples used here are established cancer cell lines, there is no real control (normal tissue), so it's likely possible that these two proteins actually are highly overexpressed in all the selected cells, and the expression difference is the biological variance only. Therefore, applying this strategy on the real clinical samples containing diseased and healthy tissue from the same patient may result in more valuable information and at least provide some complementary data.

### **3.3.9 Target validation of EPHA2**

As one member of the largest subfamily of receptor tyrosine kinases, EPHA2 appears to play important roles in cancer cells by affecting the tumor growth, invasiveness, angiogenesis and metastasis (161). In this study, differential EPHA2 expression profiled by MS was confirmed by probing the samples with an EPHA2 antibody and siRNA of EPHA2, which killed high (CAL27, UTSCC40) but not low (SIHN005A, UTSCC76A) expressing cells (Fig. 30b). All these results indicate EPHA2 to be a novel potential therapy target for head and neck cancer. Compelling further evidence would come from the use of selective inhibitors of EPHA2 but, unfortunately, such molecules have not yet been reported in the literature. Alternatively, Ephrin A1 was reported as a negative ligand of EPHA2 in some other cancers, which could function as an inhibitor although the full signaling pathway still remains unclear (162, 163). A modest (30-40%) but statistically significant and dose-dependent reduction in cell viability relative to control Fc-treated (The fragment crystallizable region is the tail region of an antibody) after incubating high EPHA2 expressing cells (CAL27, UTSCC40) with Fc-conjugated Ephrin A1 (Fig. 31a), which is in line with similar data reports for other cell lines in the literature (164). Western blot analysis showed that downstream AKT and ERK phosphorylation were reduced and FAK phosphorylation (an EPHA2

downstream but proximal signaling molecule) was increased by EPHA1-Fc treatment (Fig. 31b), suggesting that EPHA2 signaling is functional in the cell lines studied.

Recently, EPHA2 has been proved to be clinical significant when overexpressed in HNSCC patients (165, 166). In this study, siRNA and Ephrin A1 ligand experiments propose EPHA2 to be a novel therapy target for HNSCC samples and further data confirmed partially the existing hypothesis of



*Figure 31. EPHA2 target validation a) Cell viability decreases in high EPHA2 expressing cell lines following treatment with Fc-conjugated Ephrin A1. b) Signaling pathway analysis using phospho-specific antibodies reveals that Ephrin A1 treatment inhibits ERK and AKT and induces FAK phosphorylation.*

the biological mechanism. However EPHA2 plays a complex function in tumorigenesis and tumor progression and relates to tumor type and microenvironment. Some contradicting results were reported by different groups: both ephrinA1-mediated suppression and stimulation of extracellular signal-regulated kinase phosphorylation has been observed in the same type of cancer (161). Therefore, further investigation of the biological function of EPHA2 is absolutely required for developing EPHA2 as a therapeutic target for HNSCC patients.

### 3.3.10 EPHA2 was differentially expressed in the clinical samples

This experiment performed in collaboration with Prof. Dr. Med. Axel Walch, Helmholtz Zentrum München Institut für Pathologie. Target profiling results suggested that EPHA2 is significantly differentially expressed in HNSCC and further siRNA screening and Ephrin A1 treatment indicated that this receptor tyrosine kinase has a significant impact on the cell viability. To connect the profiling data to clinical patients, immunohistochemistry (IHC) was applied here to determine the expression status of EPHA2 in the clinical HNSCC sample. For an IHC experiment, selection of the primary antibody is the most important factor and the critical feature of a primary antibody is

## Chemical proteomics reveals new targets in HNSCC

specificity for the epitope. Sc-924 anti-EPHA2 antibody, which was used for the Human Protein Atlas project (<http://www.proteinatlas.org/>), was evaluated for its specificity in the Kuster laboratory since the reliability of the antibody was indicated to be uncertain for IHC by the reporter. The single band detection on western blots against a complete cell lysates of HNSCC cell lines showed that this antibody gave a nice signal to noise and only few obvious unspecific bands were detected, which was highly comparable with the primary antibody against EGFR, a well established antibody for IHC as well (Fig. 32).

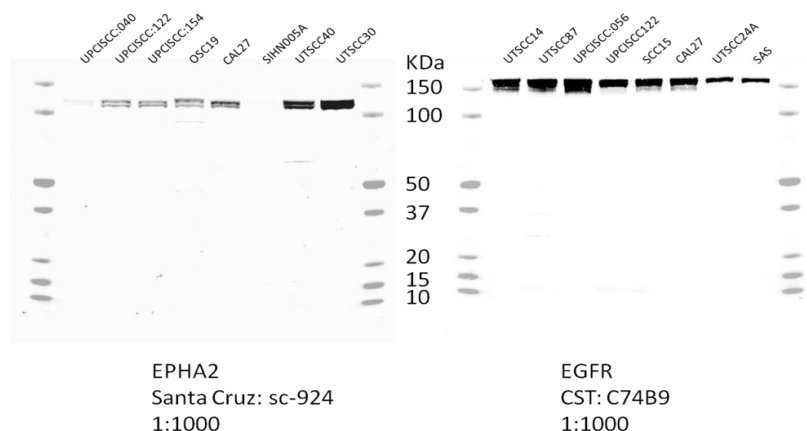
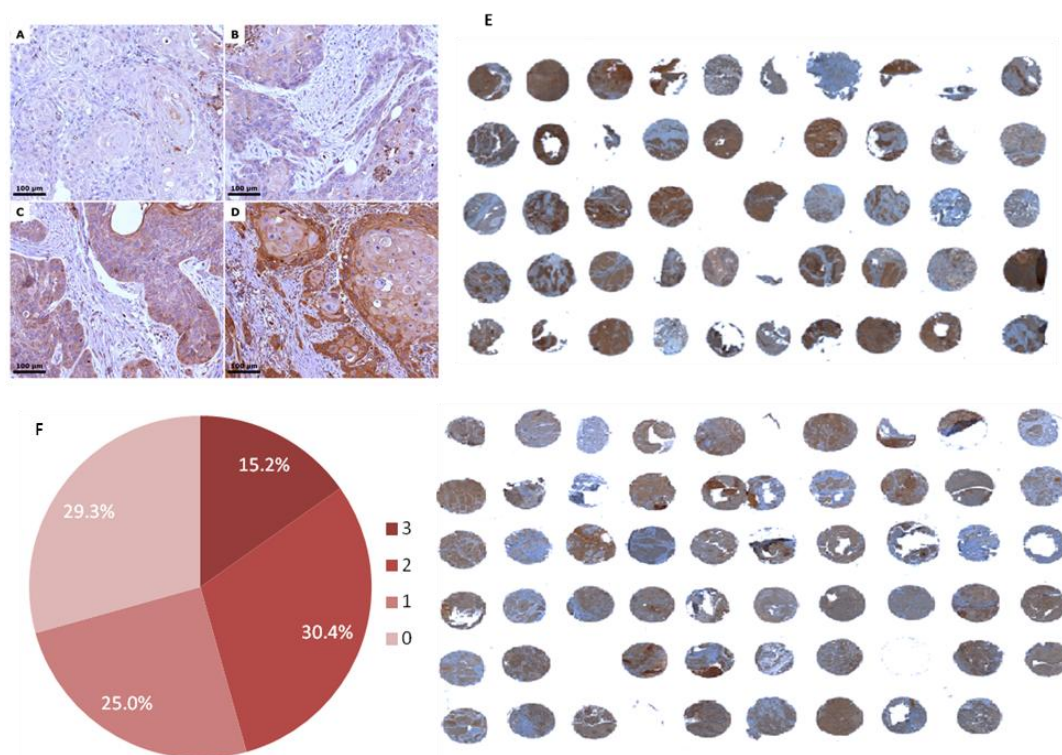


Figure 32. Antibody quality control. Western blot analysis of the EGFR and EPHA2 antibodies used in this study show very clean signals in the background of complete cell lysates of HNSCC cell lines.

In total, the EPHA2 abundance of 92 primary formalin-fixed and paraffin-embedded HNSCC samples from different anatomic subsites were determined by IHC. As shown in Fig. 33 a-d, EPHA2 was also differentially expressed in different clinical samples. Overall, there were 14 out of 92 (15%) HNSCC samples expressing a very high level of EPHA2, and another 30 samples, contributing to around 30% of the total samples, were showing weak or absent signals (Fig. 33e). The rest of the samples showed an in between expression. Overall, the immunohistochemistry data for EPHA2 protein expression was in a good agreement with the MS targets profiling results and earlier studies (76, 77) Together with the siRNA treatment, natural ligand stimulation and kinase inhibition data presented here, it suggests that EPHA2 may well be a new drug target in addition to being a potential molecular biomarker for a subpopulation (10-20%) of HNSCC patients (Fig. 33f).



*Figure 33. HNSCC tissue array for EPHA2 by immunohistochemistry. EPHA2 is highly expressed in a subset of 92 HNSCC patient samples. a) Example picture for a tumor without EPHA2 expression. b) Example picture for a tumor with low EPHA2 expression. c) Example picture for a tumor with medium EPHA2 expression. d) Example picture for a tumor with high EPHA2 expression. e) Tissue array pictures of all 92 HNSCC patient samples. The classification was performed by a trained pathologist. f) Pie chart showing a % of different grade of the investigated 92 clinical samples.*

### **3.3.11 Combination treatment by targeting EGFR and EPHA2**

Combination chemotherapy often results in enhanced effects compared to giving the single drug and in addition decreases the likelihood of developing resistant cancer cells (167). With these rationales, combination treatment becomes a popular concept of cancer therapy (168). For example, the significance of combining Rituximab Fludarabine and Cyclophosphamide for Chronic Lymphocytic Leukemia has been approved by FDA recently (169). As an advantage of the strategy described here, profiling large numbers of targets simultaneously may offer a global view of the expression status of numerous oncogenic proteins, which can be used for the combination treatment to enhance the treatment efficiency and reduce the toxicity.

Of course, multiple targets selection also should be based on a thorough understanding of the biology behind. As an example, CAL27 cells had been found to express high levels of EPHA2 as

## Chemical proteomics reveals new targets in HNSCC

well as EGFR according to the profiling result and both have been validated as the therapeutic targets for HNSCC by the data from this study or other groups (170, 171). Given the fact that currently no specific EPHA2 inhibitor is available, alternatively siRNA was utilized to investigate the co-effect of knocking down both proteins. CAL27 cells were transfected with either the siRNA alone or both together. As shown in Fig. 34, knockdown of EGFR alone reduced cell viability by 30% ( $p=0.01$ ). Knockdown of EPHA2 alone reduced cell viability by 58% ( $p=0.0006$ ). Combined knockdown reduced cell viability by 75% ( $p=3E-5$ ). The results suggested that the effect of the double knockdown on cell viability was additive (but not the same or synergistic) indicating that

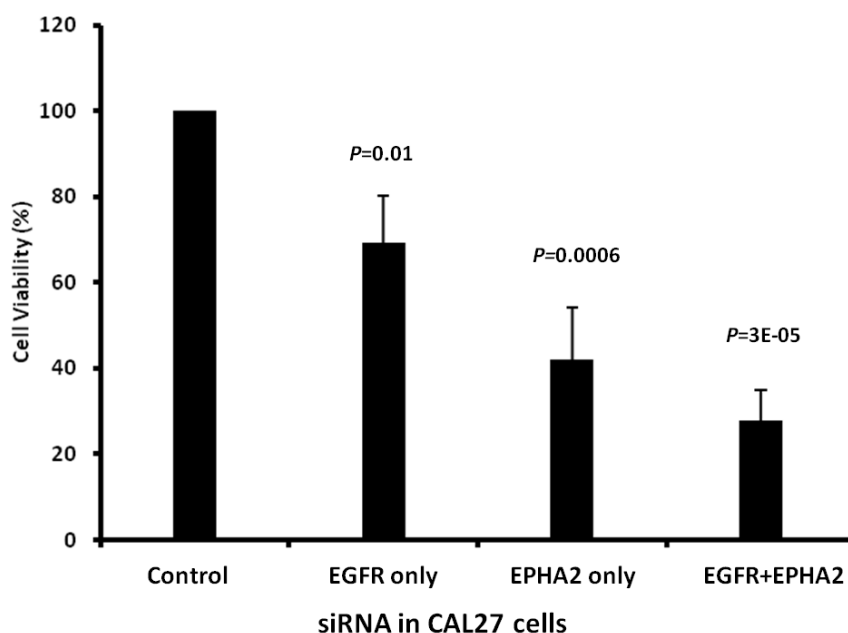


Figure 34. Combination treatment of targeting both EGFR and EPHA2 in CAL27 cells. CAL27 was transfected with 5nM siRNA against EGFR, or 25nM siRNA against EPHA2, or 5nM and 25nM siRNAs against EGFR and EPHA2 respectively. Only transfection reagent was used as control. Cell viability was measured after 96hs with XTT. The experiments were repeated three times, the histogram shown here represents the average value with SD.

both proteins trigger reduced cell viability independent of each other. As mentioned earlier, the strategy applied in this study allows profiling numerous targets in parallel, thereby providing a rich resource for selecting multiple targets for combination treatment.

### 3.3.12 Pathway analysis

Abrogation of ERK and AKT phosphorylation by either targeting EGFR or EPHA2 was observed, suggesting that these two pathways intersect or converge at some point. EPHA2 signaling is not



well studied in general and much less so in HNSCC. However, from mining the literature of the Ephrin receptor family and EGFR using the Ingenuity Pathway Analysis software, it is possible to find intersections of the pathways occurring at the kinase FAK and/or the adaptor protein SHC. Both FAK and SHC are in fact interaction partners of EPHA2 and SHC is also a described interaction partner of EGFR (172), which could explain the observed reduced ERK and AKT phosphorylation and increased FAK phosphorylation and the overall reduced cell viability following EGFR drug and EPHA1-Fc ligand stimulation (see Fig. 35 for a possible pathway models). However, future work is required to address whether these axis are also present in HNSCC cells. One indicative experiment that could be performed to test this this putative pathway is to knockout the adaptor protein SHC and in combination with the inhibitor/ligand treatment to study the effect on the cellular phenotype.

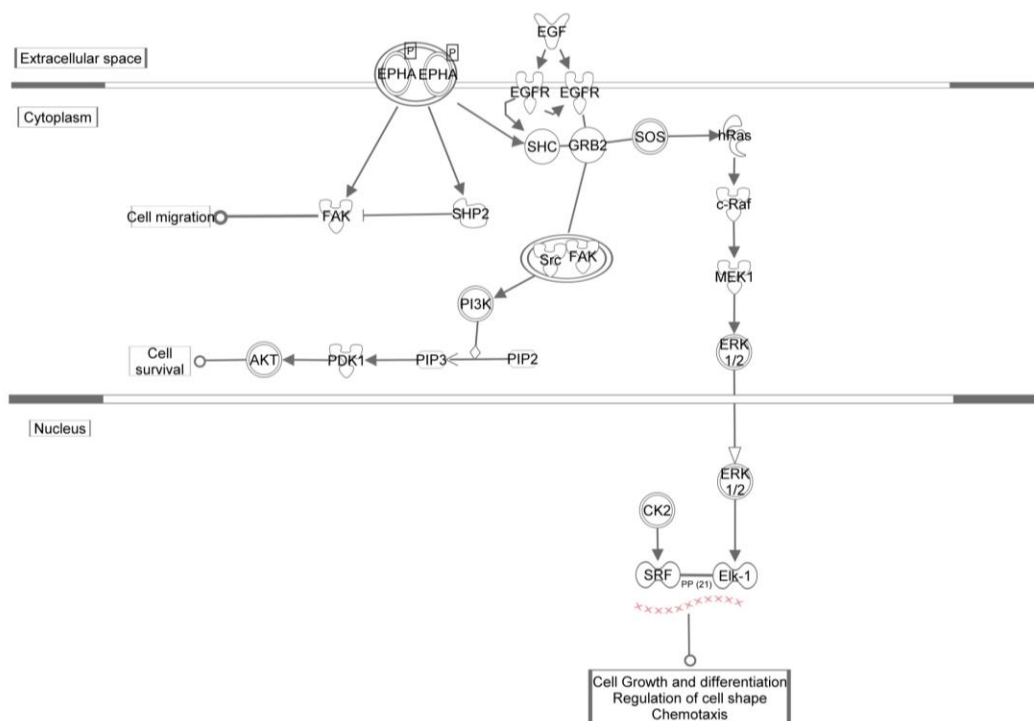


Figure 35. Putative signaling pathways of EGFR and EPHA2 in HNSCC cells (derived from the Ingenuity pathway analysis software).

### 3.3.13 Investigation of Src family kinases in HNSCC

Src family kinases (SFKs) play a vital role in many aspects of cancer development, including cell adhesion, proliferation, survival and angiogenesis, which is not hard to imagine because they are

## Chemical proteomics reveals new targets in HNSCC

central mediators in multiple oncogenic signaling pathways, for example, EGFR signaling pathway. SFKs have been reported to be involved in the GRP-induced EGFR activation signaling pathway, and consequently mediate EGFR ligand cleavage, proliferation, and invasion of HNSCC (173). In this study, the role of SFKs in the HNSCC samples was also evaluated not only because of its importance described above but also because members of the SFKs were on the list with a significant p-value. The same samples used for EGFR estimation (SIHN005A, UTSCC16A, SCC9 representing low EGFR expression; UTSCC87, CAL27 and UTSCC14 with high EGFR expression) were also used here and treated with different concentrations of dasatinib (a potent pan-SRC family kinase inhibitor) (174). Interestingly, not only the EGFR inhibitors used earlier but also dasatinib could discriminate between high and low EGFR-expressing cells ( $p=0.007$  at 100 nM dasatinib dose, Fig. 36a). This effect cannot be attributed to EGFR inhibition as dasatinib has no

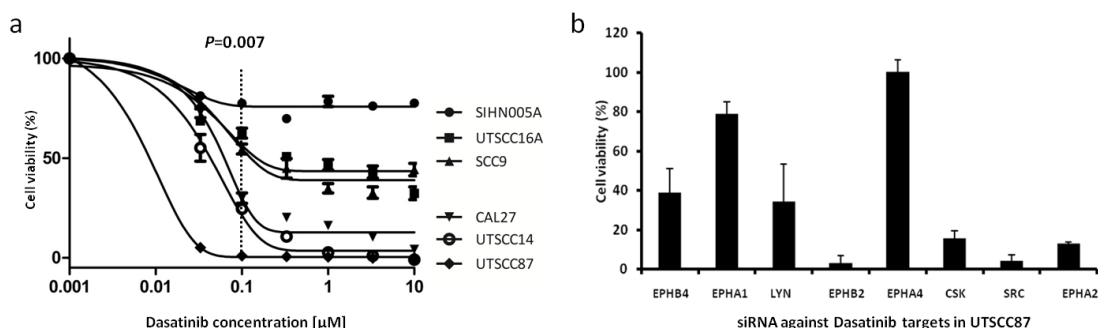


Figure 36. Role of SRC family kinases a) High EGFR-expressing cells respond more strongly to Dasatinib treatment than low-expressing cells. b) siRNA mediated protein knockdown of known Dasatinib targets in the highly Dasatinib sensitive cell line UTSCC87 shows that differential killing observed in a) can be attributed to inhibition of several proteins including EPHA2.

activity against EGFR at the concentration employed (175). Furthermore, comparing three high EGFR expressing cell lines, UTSCC87 showed a super high sensitivity to dasatinib ( $EC_{50}$  for dasatinib  $<10$  nM, Fig. 36b), indicating that likely not only the SFKs-EGFR pathway might be involved. In order to investigate which proteins might be responsible for the observed effect, several known dasatinib target proteins were knocked out via siRNA in this cell line. Interestingly, loss of SRC, CSK, EPHA2 or EPHB2 expression each led to  $>80\%$  reduction in cell viability which may be taken as circumstantial evidence that a part of the cell killing induced by dasatinib might be mediated by EPHA2 or EPHB2 besides the EGFR factor via SFKs. Nevertheless, as the biggest non-receptor tyrosine kinases, SFKs can interact with a diverse set of cell surface receptors, such

as EGFR, VEGFR and G-protein-coupled receptors as well, to mediate the cell signaling to downstream proteins, such as Ras/ERK/MAPK (176) and consequently promotes the cell proliferation, migration and invasion. Therefore, this may offer another therapy avenue to block SFKs in HNSCC patients driven by high expressed or over activated receptors, or combine treatment with the receptor inhibitor for better therapeutic effects (177).

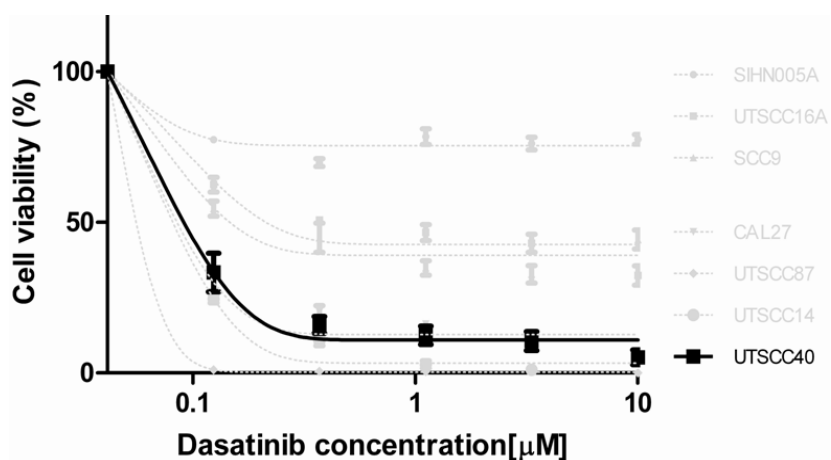


Figure 37. Dasatinib treatment on UTSCC40 cell line. UTSCC40 cells were treated with the indicated concentrations of dasatinib and the viability was estimated by XTT after 96h. Gray lines were extracted from Figure 36 as a reference for the comparison.

In addition, the dose-response curve of UTSCC40 for Dasatinib was noted to be similar to the group with high EGFR expression (Fig. 37). It cannot be attributed to the highly overexpressed EGFR pathways since this is a low EGFR expression cell line. Hence, the interpretation mentioned above cannot be applied on this cell line. However, profiling and western blot results revealed that UTSCC40 showed high EPHA2 expression, which was also proved to contribute to the proliferation of UTSCC40 cells in this study. Together with the fact that EPHA2 is a potent target of dasatinib (4 nM Kd) (175), it is most likely that reduction of UTSCC40 viability induced by dasatinib is mediated via EPHA2. However, formal proof of this hypothesis would require some highly selective inhibitors to target specific pathways, again which are not available yet. As an alternative approach, phosphoproteomics profiling might enable to comprehend the signaling pathways disrupted by Dasatinib to get deep insight of the mechanism (78).

### 3.3.14 Evaluation of c-MET as a target in HNSCC

Earlier research suggested that c-MET was a target in HNSCC (130). In this study, intensity based label-free MS protein expression profiling and western blot validation data also confirmed that c-

## Chemical proteomics reveals new targets in HNSCC

MET is moderately differentially expressed across the panel of 34 cell lines with  $p=1.80E-3$  as shown in Fig. 38. To test whether this correlates to the cell viability in response to the c-MET inhibition, four relatively high (UPI: SCC016, UTSCC74A, UTSCC10, SAS) and two relatively low (SCC25, SIHN005A) c-MET expressing cell lines were treated with the potent and selective c-MET

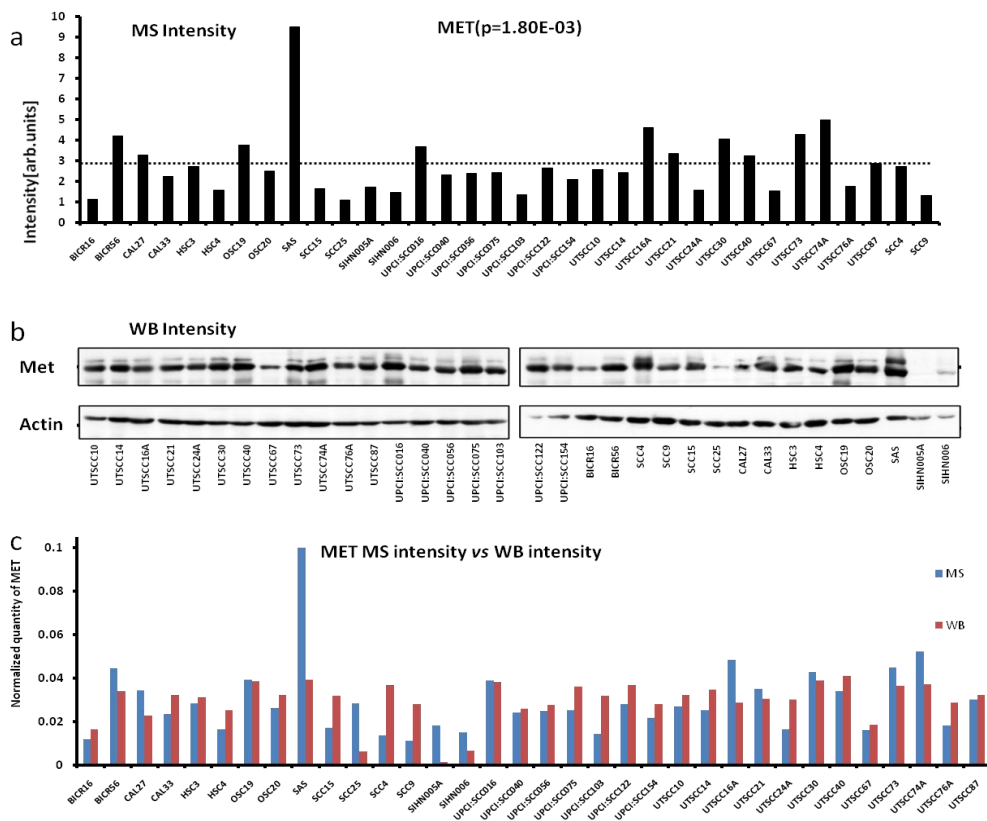


Figure 38. Correlation of c-MET protein quantification by Western Blotting and MS intensity. All 34 HNSCC cell lines were profiled for c-MET by MS intensity (Kinobeadpulldowns, panel a; the dotted line represents the median expression level) and western blotting (lysates, panel b). Actin levels by western blotting are shown as loading control. Panel c: Both quantification methods show comparable quantification results.

inhibitor PHA665752 (biochemical kinase assay IC<sub>50</sub> of 9 nM, Fig. 39a). Unexpectedly, the drug response result showed that none of the cell lines responded to the drug below the cytotoxic dose of 2  $\mu$ M. Before making a conflicting conclusion, some of the cell lines were treated by siRNA-mediated c-MET knockdown, which again did not show any significant effect on the cell viability in any of the four cell lines tested (Fig. 39b). From this data, neither pharmacological inhibition with appropriate doses of a selective c-MET inhibitor nor siRNA knockdown could significantly reduce the viability of c-MET overexpressing cells (Fig. 39). However, the earlier

studies also suggested that very high doses (2-5  $\mu\text{M}$ ) of the c-MET inhibitor SU11274 were required to kill HNSCC cells. This was more than ten times the reported effective dose for inhibition of cellular c-MET phosphorylation (130) and  $\sim 200$  higher than the reported *in vitro* dose for inhibition of the kinase activity. The selective c-MET inhibitor (PHA665752) used here has a similar potency and more than 50-fold higher selectivity across a panel of diverse tyrosine and serine-threonine kinases (178). Therefore, the interpretation of the observed reduction in cell viability is more likely due to off-target effects and nonspecific cytotoxicity rather than the result of selective c-MET inhibition. Thus, the potential therapeutic value of c-MET inhibition in HNSCC definitely requires further investigation. Clearly, c-MET may still play important roles in regulating other aspects such as motility and invasiveness of tumor cells in many HNSCCs (179-181), but this was not investigated in this study.

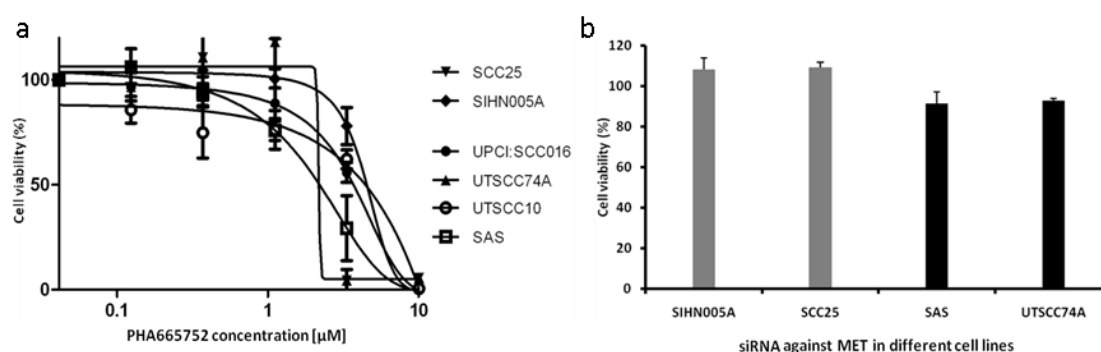


Figure 39. Role of c-MET in the survival of HNSCC cell lines. a) High and low c-MET-expressing cell lines respond similarly to the selective c-MET inhibitor PHA665752. b) siRNA mediated c-MET knockdown of low (grey) and high (black) c-MET-expressing cell lines shows no significant effect on cell viability.

### 3.4 Conclusion

In this study, a kinase-focused chemical proteomics approach in conjunction with intensity-based label-free quantitative mass spectrometry was applied here to profile the expression of 146 kinases across 34 HNSCC cell lines of the tongue, in order to identify novel potential drug targets and/or signaling pathway biomarker candidates. The analysis revealed a great molecular heterogeneity within this group of cell lines. A number of proteins previously implicated in the disease (EGFR, c-MET, etc) as well as several novel protein kinases were identified to be important for the survival and proliferation of HNSCC cells, including ephrin A2 receptor. These kinases may represent targets for therapeutic intervention or act as biomarkers indicative of

## ***Chemical proteomics reveals new targets in HNSCC***

---

active signaling pathways. Both genetic and pharmacological inhibition of previously implicated therapeutic target c-MET suggests that further investigation was required to establish c-MET for HNSCC therapy.

## **4. Systematic identification of the HSP90 regulated proteome**

### **4.1 Introduction**

Heat shock protein 90 (HSP90) is an abundant, ubiquitously expressed chaperone protein, comprising 1~2% of all cytosolic proteins under non-stress conditions (11). By cycling between two major conformations, i.e. an “open” and a “closed” state, HSP90 regulates many aspects of post-translational protein homeostasis and consequently modulates multiple cellular processes, including assembly of protein complexes, trafficking, protein function, and quality control and so on. During the “open” state, ATP and the nascent protein are loaded to the HSP90 protein to form a transiently hetero-complex (10, 11). Subsequently, ATP hydrolysis drives the HSP90 to twists to the “closed” status and ADP is released to finish the cycle. To accomplish these functions, HSP90 needs the assistance of many co-chaperones, such as HSP70, HOP and CDC37 which is specially required for kinases(182). These HSP90 dependent proteins are known as clients, and the majority of them are signal transducers, representing a large number of key nodes of many biological networks (183, 184). Thereby, HSP90 exerts marked effect on normal biology and diseases. Particularly for cancer, genetic instability allows the common occurrence of overexpressed and mutated oncogenic proteins, such as EGFR, AKT, KIT, Bcr-Abl, most of which have been proved to be HSP90 dependent clients. With decades of effort, approximately 200 proteins have been identified as HSP90 client proteins. Genome-wide screens in yeast illustrated an unexpectedly large number of potential clients of HSP90 (approximately 20% of yeast proteome) (185, 186), which suggests that, most likely, the complete set of HSP90 dependent proteins is unknown (187, 188).

HSP90 has been considered to be a noncogeneic protein. Despite the fact that no mutation is found, overexpression and hyperactivated states of HSP90 are common features in many solid tumors and hematological malignancies, which allow malignant cells to maintain correct protein folding under the hostile hypoxic and acidotic stress conditions of tumor growth (183, 189, 190). In addition, mutated or overexpressed oncogenic proteins have shown a particular requirement of HSP90 for stability and function (191, 192). Therefore, the disruption of HSP90 function by small molecule drugs has become an attractive therapeutic strategy and about a dozen of HSP90 inhibitors are currently undergoing clinical trials in a number of tumor entities and indications (10, 183, 193). The highly conserved sequences of HSP90 consists of three domains: an amino

## ***HSP90 regulated proteome***

---

terminal domain (NTD) containing the motifs for co-chaperones and ATP binding; a middle domain (MD) with the docking sites for substrate proteins and co-chaperones and a carboxy terminal domain (CTD) responsible for the dimerization of two HSP90 protomers. As many kinase specific inhibitors, most of the HSP90 inhibitors are currently designed to competitively bind to the ATP pocket, which subsequently interrupts the chaperon cycling. Another type of inhibitor, designed to disrupt the protein-protein interaction between HSP90 and its clients, is also under research. The latter has been shown to be of lower toxicity, such as novobiocin targeting the interaction between HSP90 and the chaperon protein CDC37 (10, 11, 183). However development of this type of small molecular inhibitor remains rather challenge due to issues such as the lack of well-defined binding pocket (194). Geldanamycin (GA) belongs to the former group of HSP90 inhibitors that target the ATP binding pocket of HSP90 and block the chaperone cycle which on the one hand leads to proteasome mediated degradation of HSP90 substrates via ubiquitin pathways (182, 183) and, on the other hand, to cell responses such as transcription factor activation and subsequent gene expression changes (e.g. HSF1) (195, 196). For cancer therapy, these HSP90 specific inhibitors indirectly inhibit these client proteins (inhibiting their maturation) and interrupt the oncogenic signaling pathways at multiple nodes. However, like an unselective kinases inhibitor, an avoidless consequence of targeting multiple targets is the serious toxicity which also has been observed in clinical trials. In addition it has been observed that the efficacy and toxicity of HSP90 targeted therapy varies greatly between tumors. Therefore it is most likely that the current repertoire of client proteins and our understanding of drug mechanism of action are far from being complete (197). To predict an individual patient's responsiveness, it would thus be highly desirable to identify the complete set of HSP90 regulated proteins. Because HSP90 directly (e.g. by degradation) and indirectly (e.g. by induction of gene/protein expression) affects proteostasis, proteomic approaches are particularly attractive for studying e.g. the HSP90 interactome and the global effects of HSP90 inhibition on cellular systems. A number of proteomic approaches have been introduced to explore the HSP90 dependent proteome which are functionally interacted and/or physically interacted with HSP90. Global proteome profiling using 2D gels and mass spectrometry (198) as well as focused proteomic experiments utilizing immunoprecipitation of HSP90 complexes and chemical precipitation using immobilized HSP90 inhibitors (188) have identified numerous important HSP90 clients. But these methods generally fail to provide a global view of the HSP90 regulated



proteome because the attained proteomic depth was very limited and many HSP90 interactions are too transient or have too weak affinity to be purified by these methods (199). Recently, advanced mass spectrometry based approaches enables to investigate the proteome wide regulation which allowed investigation of some oncogenic pathways involved in the HSP90 inhibition (200). However, kinases, the main components of the signaling pathways, are still a challenge to investigate because of the low abundance of these proteins and the limited capacity of mass spectrometry. Model based and bioinformatics assistant methods enable researchers to predict the physical interactions according to the common sequence of the physical interaction domain (187), however, the conclusions normally still requires validation *in vitro/vivo* and part of the conclusions drawn are still under debate , such as the conflicting result obtained for ERK5. Compiling information suggested that ERK5 has no common surface interacting with HSP90, but it was proved to be a *bona fide* client of HSP90 by biological experiments (187, 201).

Recently, a chemical compounds based affinity purification method (Kinobead) has been developed to enable the enrichment of hundreds of protein kinases by immobilizing the pan-kinase inhibitors onto the resins which can selectively capture the kinases from the protein mixture (75). Combining this method with SILAC-based quantitative mass spectrometry (65), it enables to profile the changes in kinase expression induced by HSP90 inhibition in a kinome wide manner since the HSP90 clients are thereby enriched. In this study, together with 1D gel separation, this strategy was utilized to profile the global response of the proteome and kinome of the four cancer cell lines K562, Colo205, Cal27 and MDAMB231 to the HSP90 inhibitor GA. ~1,600 out of total >6200 identified proteins showed common as well as cell type specific regulation upon drug treatment. Bioinformatic analysis enabled a functional organization of this data into protein pathways, networks and complexes, highlighting many known and novel aspects of HSP90 function. Protein turn-over measurements using pulsed SILAC (63, 202) showed that protein half life is a main gradient for HSP90 inhibition induced hierarchical destabilization. In addition it was shown that protein kinases have significantly shorter half lives than other proteins, having implications for HSP90 inhibition in cancer therapy. A comparison of the effects of GA and the purine-based phase I drug PU-H71 (203) suggests that both molecules work by similar molecular mechanisms. Differential reduction of cell viability induced by specific kinase inhibitors revealed an altered AKT independent EGFR-PI3K-mTOR pathway to be involved in CAL27 cell line. Using HSP90 immunoprecipitation and pulldowns with immobilized GA, several

## ***HSP90 regulated proteome***

---

kinases (ARAF, DDR1, and TRIO) and other signaling proteins (IKBKG, BIRC6) were validated as novel *bona fide* clients of HSP90. This demonstrates the value of this global profiling approach and provides a rich resource for future investigation.

### **4.2 Material and methods**

#### **4.2.1 SILAC labeling and cell culture**

CAL27 and MDAMB231 cells were cultured in DMEM (4.5 g/liter glucose) medium, K562 and COLON205 cells were cultured in RPMI1640 medium. To enable quantitative mass spectrometry, cells were converted with SILAC medium, i.e. normal medium deficient in arginine and lysine (PAA, Pasching, Austria) were supplemented with stable isotope-encoded arginine 6 and lysine 4 (Euriso-top) for heavy and arginine 0 and lysine 0 for light as described previously (204). For both SILAC conditions, medium was supplemented with 10% dialyzed fetal bovine serum (Gibco®, Invitrogen, Darmstadt) and 200 mM L-proline (Sigma-Aldrich, Germany) to inhibit the arginine to proline conversion. General cell culture conditions were 37 °C in humidified air with 5% CO<sub>2</sub>. Cells were seeded at a density of 2x10<sup>5</sup> and cultured 24 h before further adding GA or DMSO. SF268 cells carrying the established DDR1 isoforms were kindly supplied by Dr. Kramer and cells were cultured in DMEM medium supplemented with 10% FBS (PAA, Pasching, Austria) and 150 µg/ml of the antibiotic Hygromycin B (PAA, Pasching, Austria).

#### **4.2.2 Pulsed SILAC**

Pulsed SILAC experiments for half-life determination were performed as described (63). Briefly, cells were growing in normal DMEM (CAL27 and MDAMB231) or RPMI (COLO205 and K562) under the conditions as described above until around 40% confluence followed by switching to the corresponding SILAC labeled medium. Cells were harvested at three time points (6, 12 and 24 h).

#### **4.2.3 Drug treatment and harvesting**

20mM stock solution of GA (LC Laboratories, Woburn, MA) was prepared by dissolving in DMSO and used within 2 weeks. Cells were treated 24 h with the IC<sub>50</sub> concentrations of GA according the drug-response curve, briefly, 5 µM final concentration for K562, CAL27 and MDMBA231 and 10 µM final concentration for COLO205. The corresponding control groups were incubated with same concentration of DMSO (0.1%) only. After washing with pre-cooled PBS 3 times,

approximately an equal number of cells from the GA treated and control were lysed in 50 mM Tris/HCl pH 7.5, 5% glycerol, 0.8% NP-40, with freshly added protease (SIGMAFAST, Sigma-Aldrich) and phosphatase inhibitors (5x phosphatase inhibitor cocktail1, Sigma-Aldrich, Munich, Germany; 5x phosphatase inhibitor cocktail 2, Sigma-Aldrich, Munich, Germany). Homogenates were centrifuged at 6000x g at 4°C for 10 min followed by centrifugation at 4 °C for 1 h at 145,000x g. Supernatants collected, aliquoted, frozen in liquid nitrogen and stored at -80 °C until further use. The protein concentration in lysates was determined by the Bradford assay.

### **4.2.4 Gel based sample preparation for full proteome**

100 µg mixture of GA and DMSO treated lysate was reduced and alkylated by incubating with 10 mM DTT at 56 °C for 1 h and subsequently with 55 mM IAA in dark at room temperature for half an hour before denaturing at 95 °C with NuPAGE® LDS Sample Buffer (Invitrogen, Darmstadt, Germany) for 5min. Samples were then separated into 16 slices with a 4–12% NuPAGE gel (Invitrogen, Darmstadt, Germany) prior to in-gel trypsin digestion. In-gel trypsin digestion was performed according to standard procedures.

### **4.2.5 Affinity purification**

(See chapter 2.2.4)

### **4.2.6 Immunoprecipitation and GA-NHS affinity purification**

MDAMB231 cells were cultured in SILAC DMEM with either Heavy (Arg6, Lys4) (Euriso-top) or Light DMEM (Arg0, Lys0) (PAA, Pasching, Austria) supplemented with 10% (v/v) dialyzed fetal bovine serum (FBS, Gibco®, Invitrogen, Darmstadt). Cells were lysed in 50 mM Tris/HCl pH 7.5, 5% glycerol, 0.8% NP-40, 1.5 mM MgCl<sub>2</sub>, 150 mM NaCl, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 25 mM NaF, and protease inhibitors (SIGMAFAST, Sigma-Aldrich). Homogenates were centrifuged at 6000x g at 4°C for 10 min to remove cell debris. Cleared lysates were either incubated overnight at 4 °C with anti HSP90 antibody (C20, Santa Cruz) or the same source of normal IgG (mouse, Santa Cruz) and both were followed by incubating with protein A/G beads (Santa Cruz) for another 4 h at 4 °C. After extensive washing, immunoprecipitates from heavy and light cells were boiled in LDS buffer, combined and then separated on a NuPAGE® 4-12% Bis-Tris Mini Gel. Gel lanes containing separated immunocomplexes were cut into 12 slices and in-gel trypsin digestion was performed according to standard protocols.

## ***HSP90 regulated proteome***

---

For GA-NHS affinity purification, the same cell lysate as for the immunoprecipitate was pre-incubated with either 25  $\mu$ M GA or 0.1% DMSO for 1 h at 4 °C. Sepharose beads with the immobilized GA were subsequently added for another 1 hour at 4 °C. The following steps were the same as for the immunoprecipitation experiment mentioned above.

### **4.2.7 LC-MS/MS analysis**

The mass spectrometric analysis was performed as described in chapter 2.2.9 except using 2h gradient for each slice of the full proteome samples and Top 15 most intensity of the precursor ions were picked up for tandem mass spectrometer.

### **4.2.8 Peptide and protein quantification and identification**

Raw MS spectra were processed by Maxquant (version 1.1.1.25)(54, 62). The maximum false discovery rate (FDR) for proteins and peptides was 0.01 and a minimum peptide length of 6 amino acids was required. MS/MS spectra were searched against the IPI human database human (v. 3.68, 87,061 sequences) by the Andromeda search engine (54, 62) enabling the contaminants and the reversed versions of all sequences with the following search parameters: fixed modification of carbamidomethylation of cysteine residues, Acetyl (Protein N-term), Gln\_pyro-Glu (N-Term Q), Glu\_pyro-Glu(N-Term E), Oxidation(M), Phospho(ST), Phospho(Y) as variable modifications, Trypsin as proteolytic enzyme with up to 2 miss cleavages, fragment ion mass tolerance of 0.6 Da. The mass accuracy of the precursor ions was decided by the time-dependent recalibration algorithm of Maxquant. After filtering the contaminants and reverse hits by the Perseus Software, normalized intensities were exported from Maxquant.

### **4.2.9 Statistic analysis**

This analysis was in collaboration with Dr. Amin Moghaddas Gholami, Kuster Lab, TUM. Statistical analysis of the quantified proteins was performed using R (v 2.12.1) (205). Raw protein abundance values were first normalized using Variance Stabilization Normalization (206). VSN is able to stabilize the variance across the entire intensity range and addresses the error structure in the data. The application of VSN has previously been shown to be beneficial for MS-based quantification (135, 207). To investigate the data distribution and ensure the appropriate application of statistical tools, normal quantile-quantile plots were created for all protein

intensities in each cell line. Variance stabilization as well as variance-mean dependencies were visually verified for each cell line.

Differential expression of paired samples was assessed with a moderated linear model using the limma package (103) in Bioconductor (104). Differences in protein expression between 'treated' and 'control' samples were estimated with the least squares linear model fitting procedure and tested for differential expression with moderated Student's t-statistic via the empirical Bayesian statistics described in the limma package (103). The null hypothesis were accepted or rejected on the basis of P-values computed for the omnibus B-statistic via limma at a specified significance level. P-values were adjusted for multiple testing to control the False Discovery Rate (FDR) at 5%. For multiple testing adjustments, the FDR was calculated using the algorithm of Benjamini and Hochberg (208). p-values, with appropriate multiple testing adjustment to control the FDR at 5% allowed us to identify differentially expressed proteins.

#### **4.2.10 GO enrichment/ Pathway analysis/complex analysis**

This analysis was in collaboration with Dr. Amin Moghaddas Gholami, Kuster Lab, TUM. Classification and functional enrichment analysis of the identified proteins were performed using DAVID Bioinformatics Database (209, 210) for the biological process (BP), molecular function (MF) and cellular component (CC). In order to make the functional categories more understandable, terms were clustered according to their functional similarity using REVIGO (211).

Pathway membership of the identified proteins was analyzed by the Ingenuity Pathway Analysis (IPA) tool (Ingenuity Systems, Redwood City, CA, USA) for their functional significance and in the context of biological association networks.

To investigate HSP90 targeting of many macromolecular complexes, the list of significantly expressed proteins (adjusted  $p < 0.05$ ) were analyzed using Comprehensive Resource of Mammalian protein complexes (CORUM) (212), a database of manually curated and validated mammalian protein complexes.

#### **4.2.11 Immunoblot analysis**

Anti EGFR, AKT, ERK, STAT3 and phospho-STAT3 were purchased from Cell Signaling Technology (Frankfurt, Germany) and the rest was purchased from Santa Cruz (CA, USA). For immunoblot analysis, cells were washed with pre-cooled PBS and lysed in RIPA buffer. Protein concentration was determined by the Bradford assay. Fifty  $\mu\text{g}$  of lysate from control and treated cells were

## ***HSP90 regulated proteome***

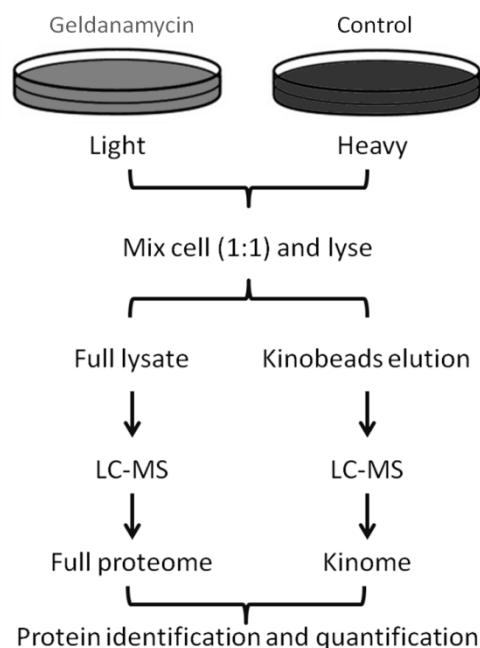
---

mixed with equal volumes of 2x NuPAGE<sup>®</sup> LDS sample buffer containing 10 mM DTT and heated at 95 °C for 5 min. Proteins were subsequently separated by 4–12% NuPAGE<sup>®</sup> gel and transferred onto to PVDF membranes (Invitrogen, Darmstadt, Germany). Membranes were blocked for 1 h in blocking solution (2% BSA in 1xTris Buffered Saline, TBS, 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 0.1% Tween-20) at room temperature and probed overnight at 4 °C with the respective primary antibody. Immunoreactivity was detected using IRDye<sup>®</sup> conjugated secondary antibody (LI-COR<sup>®</sup>, Nebraska, USA) and visualized by Odyssey imagin system (LI-COR<sup>®</sup>, Nebraska, USA).

### **4.3 Results and discussion**

#### **4.3.1 Strategy of kinome and proteome wide investigation of HSP90 regulation**

HSP90 inhibition triggers the degradation of the abandoned clients (182) and subsequent cell responses. With this fact, a mass spectrometry based quantitative proteomics strategy is described here to compare the proteomes of cancer cell lines in response to HSP90 inhibition by the small molecule drug GA. The approach allows the investigation of regulated full proteomes and kinomes. Briefly, as shown in Fig. 40, cell lines were metabolically labeled with stable isotope labeling amino acids in cell culture (SILAC) to enable the MS based quantification. The same cell line with different mass color was treated with GA, a specific HSP90 inhibitor, or DMSO as control.



*Figure 40. Quantitative proteomic strategy for proteome and kinome wide investigation of HSP90 dependent regulation.*

Approximately an equal number of cells from treated and untreated samples were mixed and lysed. Kinases were partially purified using Kinobead which comprise a mixed sepharose matrix of eight immobilized kinase inhibitor mimics while the corresponding full proteome was separated with 1D-SDS-PAGE in parallel. Subsequently, both kinases and full proteome were subjected to trypsin digestion and nano-scale liquid chromatography tandem mass spectrometry (nano LC-MS/MS). Protein identification and quantification was achieved by Maxquant integrated Andromeda search engine (54, 62) and the m/z and retention time aligned precursor ion intensities of the identified peptides were integrated across the chromatographic peak respectively. After performing statistical analysis, proteins that were significantly down regulated upon drug treatment should represent HSP90 clients, interactors and more complex regulatory events while proteins up-regulated upon drug treatment should represent only the latter category. Furthermore, co-immunoprecipitation in conjunction with quantitative mass spectrometry or classic western blot analysis was applied to primarily validate part of the potential clients and interactors.

### **4.3.2 GA reduces cell viability**

In this study, 4 cell lines derived from different organs, CAL27 (head and neck cancer), K562 (chronic myelogenous leukemia), COLO205 (colon cancer) and MDAMB231 (breast cancer) were used to investigate the proteome and kinome wide protein changes induced by HSP90 inhibition. To determine the IC<sub>50</sub> concentration, cells were treated with a serial concentrations of GA from ranging from 0.12  $\mu$ M to 10  $\mu$ M (3 times dilution) for 24 hrs to establish dose-response curves on four different cell lines. Cell viabilities of CAL27, COLO205 and MDAMB231 were measured by XTT, while cell viability of K562 was determined by trypan blue cell counting after incubating with indicated concentrations of GA for 24 hrs (Fig. 41a). For the COLO205 cell lines, a significant reduction in cell viability (determined by XTT assay) occurred at 0.37  $\mu$ M GA, but no further significant reduction was observed at higher concentration of GA. COLO205 is a mixture of adherent and suspension cells, the low concentration of GA may only have an impact on the attachment of the cells to the flask resulting in higher proportion of suspension cells which will be discriminated with the inviable cells by the XTT assay. To be on the safe side, 10  $\mu$ M GA was used for COLO205 cell line in the following experiments, and the GA concentrations used for the other three cell lines are listed in Fig. 41b.

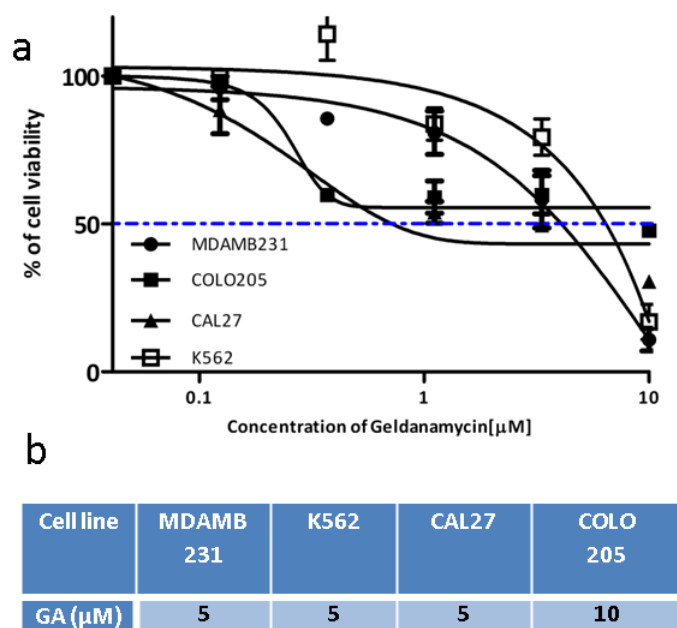
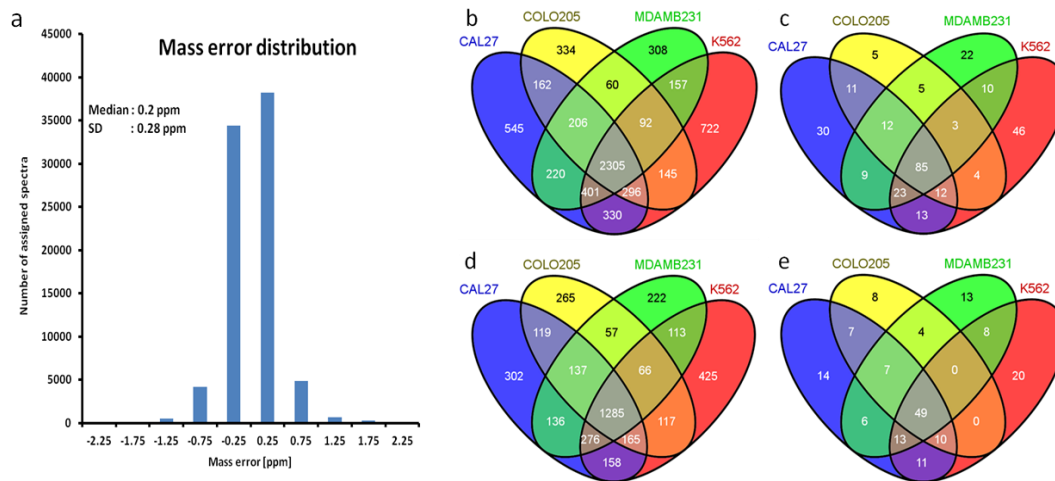


Figure 41. Geldanamycin (GA) treatment of human cancer cells. a) Cells were incubated with the indicated concentrations of GA, and cell viability was determined after 24h. Blue dashed line indicates 50% cell viability compared to the control (0.1% DMSO). b) GA concentration used for the proteome profiling experiments.

#### 4.3.3 Proteome wide mass spectrometric identification and quantification of protein changes induced by GA

For full proteome analysis, 100µg of total protein were first separated into 16 fractions by 1D-SDS-PAGE gel, followed by 2 h gradient nanoLC-MS/MS. The raw data from the same biological experiment were processed together by the Maxquant software and the integrated search engine Andromeda was used for database searching. Peptide masses were recalibrated by Maxquant pre-searching against a small database with 20 ppm mass tolerance, which significantly improved the mass accuracy as shown in Fig. 42a. Using the criteria of protein and peptide identification false positive rate (FDR) less than 1%, 83,387 high resolution spectra corresponding to 42% of the total fragment spectra were assigned to 25,326 peptides with a median precursor mass error of 0.2 ppm and less than 0.3 ppm standard deviation (data illustrated from the first replicate of CAL27). 3600 to 4465 proteins were identified from each of the four cell lines, and in total of the four cell lines, 3843 proteins out of 6283 identified proteins were quantified in all three biological replicates. Four different organ derived cell lines shared 2305 identified and 1285 quantified proteins representing 36.7% and 33.4% of total identified





*Figure 42. Peptide Mass accuracy and protein identification. a) Peptide precursor mass error distribution of identified proteins after recalibration by Maxquant. The median mass error is 0.2 ppm with a standard deviation of 0.28 ppm. b) and c) Overlap of all identified proteins and protein kinases from four cell lines. c) Overlap of all the quantified proteins and protein kinases from four cell lines (only proteins quantified in all three replicates were considered here).*

and quantified proteins respectively. The kinobead experiment resulted in a total of 288 identified protein kinases representing 60% of the whole human kinome and 172 protein kinases quantified in all three replicates from the four cell lines. Compared to the overlap of identified and quantified full proteomes, the identified and quantified protein kinases (29.5% and 26.2%) showed less overlap by the four different organ derived cell lines (Fig. 42b-e). As one of the biggest family of proteins, kinases have been shown to participate in the regulation of many functions of a metazoan cell (213). In a multi-cellular organism, different types of cells have their own unique functions which likely require different clusters of kinases. For example, Ephrin receptor tyrosine kinases play vital roles in cell communication systems; they are commonly expressed in nervous cells and many solid tumor cells but only EPHB4 and Ephrin-B2 ligands have been found in leukemia (214). Therefore, it is most likely to be the case that the overlap of expressed kinases between different cell types is smaller than any other groups of proteins (i.e. cancer central proteome (215)) or the general full proteome as observed in this study.

#### 4.3.4 Data evaluation

This analysis was performed in collaboration with Dr. Amin Moghaddas Gholami, Kuster Lab, TUM. For the robust analysis, all statistical analysis was only performed on the proteins quantified in all three biological replicates. First, in order to investigate the data distribution to

## HSP90 regulated proteome

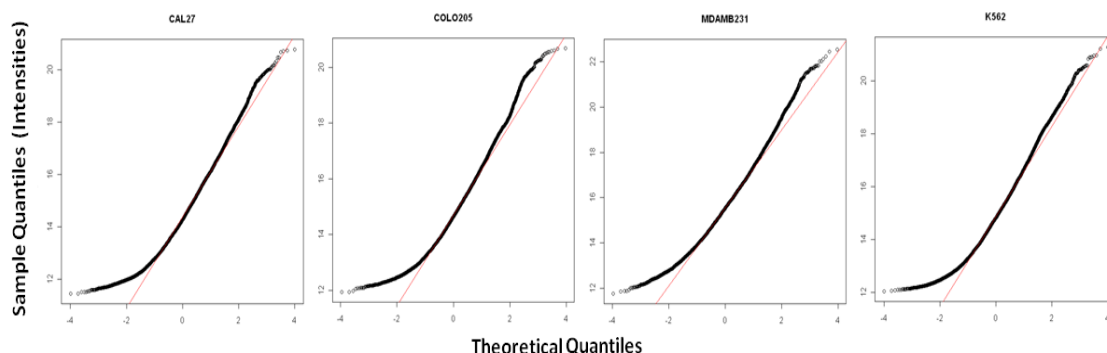


Figure 43. Normal quantile-quantile plot. Q-Q plots showing the MS intensity data distributions of every cell line after VSN normalization on the vertical axes vs the standard normal distribution on the horizontal axes. (Figure generated by Dr. Amin Moghaddas Gholami, Kuster Lab, TUM)

ensure the appropriate application of statistical, frequency histograms and quantile-quantile (Q-Q) plots were created for the quantified proteins intensities after VSN normalization. The linearity of the data points suggests that the data have a normal (or close to normal) distribution (Fig. 43).

To evaluate the variance stabilization by VSN, estimated standard deviations of each identified protein expression (y-axis) were plotted against the rank of the average (x-axis). Ideally a stable variance is a horizontal line although under the actual circumstances there may be some random fluctuations, but it should not show an overall trend (206). As shown in Fig. 44, the raw data from this study show a significantly increasing variance with the increase of intensity ranking.

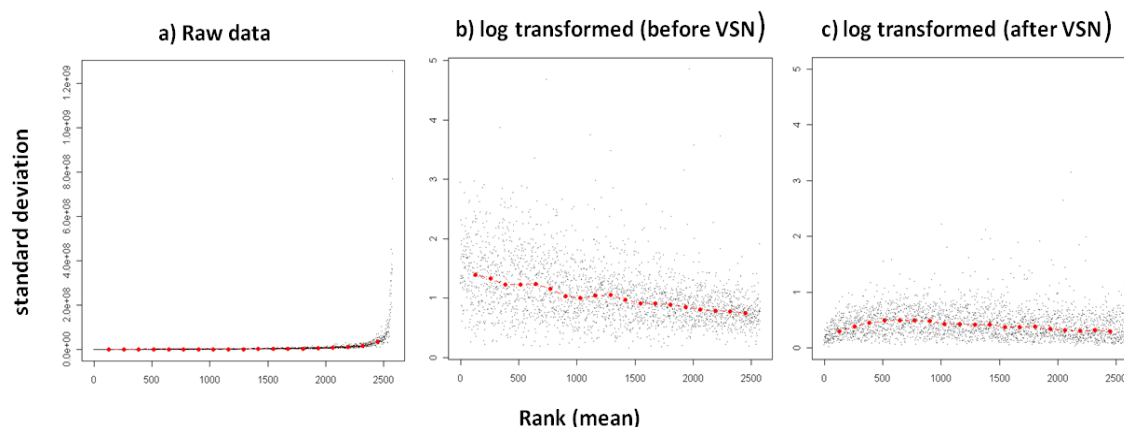


Figure 44. Rank of mean and standard deviation relations of the CAL27 data before and after VSN. Comparison of the variance-stabilization effects of CAL27 data are shown before and after normalization. Each data point represents a protein. The red dots depict the running median estimator (window-width 10%). a) Raw data without transformation. b) Log transformed before VSN. c) After VSN. (Figure generated by Dr. Amin Moghaddas Gholami, Kuster Lab, TUM)

Comparing the ranking of the intensity mean and standard deviation relations of the transformed data (b and c), the VSN method outperforms the log transform which does not perform well for the proteins with low ranking (low-intensity signals). Two conclusions could be made from the results described above: first the VSN normalization method applied here should be applicable and useful for MS based quantification, which has been discussed previously(135). Second, the normally distributed data allows further performing with the moderated Student's t-statistic to highlight the significantly regulated proteins.

**4.3.5 High quality map of HSP90 regulated proteins**

For all subsequent data analysis and bioinformatic analysis, only those proteins that were quantified in all three biological replicates of a particular cell line were considered. Using these rigorous criteria and moderated Student's t-statistical analysis resulted in 480 to 800 significantly regulated proteins from each of the four cell lines contributing 21% to 30% of the quantified proteins. In total, more than 1600 proteins were significantly ( $p < 0.05$ ) up- or downregulated following GA treatment (Fig.45) including HSP70, a co-chaperone, whose upregulation has been proved to be a marker of HSP90 inhibition. The mass spectra and extracted ion chromatograms shown in Fig. 46 give examples of three proteins that are downregulated (left, DDR1), not regulated (middle, CK2A2) or upregulated (right, AURKA). Depending on the cell line, between 20% (MDAMB231 cells) and 30% (K562 cells) of all identified proteins were significantly regulated and, in all cell lines, more proteins were downregulated (52-64%) than upregulated (36-48%). Previous work in yeast has suggested that up to 10-20% of all proteins may be influenced by

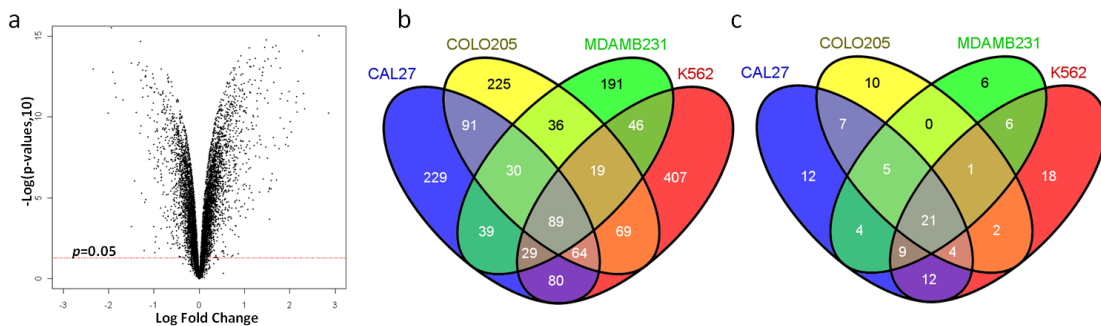


Figure 45. Significantly regulated proteome. a) Volcano plots show the -Log (p-values, 10) value versus the log fold change of the quantified proteins from four cell lines. The dashed red-line shows where  $p = 0.05$  with points above the line having  $p < 0.05$  and points below the line having  $p > 0.05$ . b) and c) Overlan of significantly regulated proteome and kinome from four cell lines.

## HSP90 regulated proteome

HSP90 function (11, 216); and together with the data shown here it is suggested that this might be also broadly the case in human cancer cells.

Interestingly, more than 65% of the significantly regulated proteins were exclusively presented in one of the cell lines. For example, there were 407 proteins showing significant regulation only in K562 cells and similar results were observed in all the four other cell lines. This might suggest that distinct HSP90 dependent pathways are specifically involved in different types of cancer and also may confirm another opinion that many of the clients are conditional: a particular constellation of proteins is regulated in one sample but not the other. For example, the protein kinase ZAP-70, which has been reported to be a HSP90 inhibitor sensitive client in chronic lymphocytic leukemia cells, but not in normal peripheral blood lymphocytes (199). Likewise, it is most likely that the whole regulated proteome is also conditional depending on the individual cell.

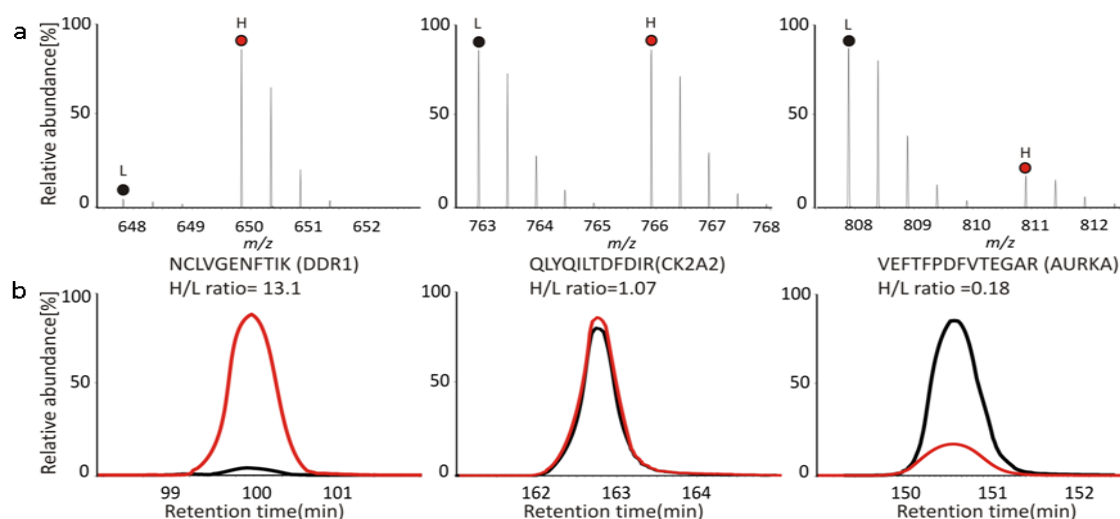


Figure 46. a) Example mass spectra and b) corresponding extracted ion chromatograms of drug induced down-regulated (left), not regulated (middle) and up-regulated (right) proteins. L and H denote peptide species containing either light or heavy amino acids.

### 4.3.6 Kinome-wide quantitative profiling reveals new potential clients of HSP90

Coordinating with CDC37, HSP90 facilitates the maturation of many kinases. For these kinase clients, functional HSP90 is required to prevent degradation (11, 182). Thereby, the degradation of a given kinase induced by HSP90 inhibition highly implicates this kinase to be a client of HSP90. Utilization of the kinase-centric strategy conjugated with the SILAC based quantitative MS enables to profile the kinases changes induced by HSP90 inhibition in kinome wide manner (75).

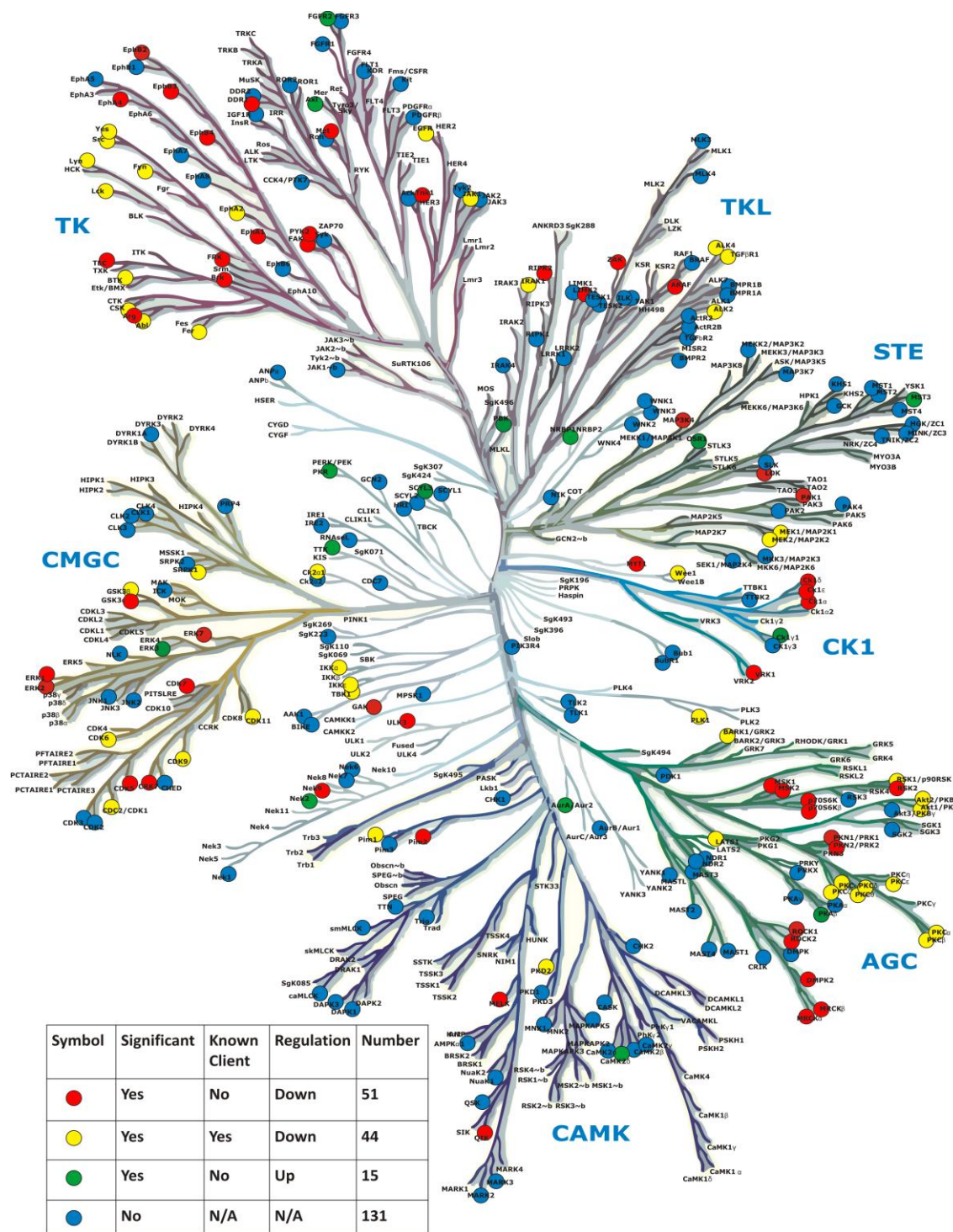


Figure 47. Phylogenetic tree of human kinases. All identified protein kinases were mapped onto the tree. Kinases unaffected by HSP90 inhibition are marked in blue, up-regulated kinases in green, down-regulated known HSP90 clients in yellow and down-regulated novel kinases in red. The kinome tree illustration was adapted from Cell Signaling Technology, Inc. (<http://www.cellsignal.com/>).

## ***HSP90 regulated proteome***

---

After processing the data as described above, between 137 to 195 protein kinases were identified from each cell line. In total of four cell lines, 170 kinases (in triplicates) were quantified out of total 288 identified proteins kinases, 117 showed significant regulation including 100 down-regulated (around 90% of the significantly regulated kinases) and 17 up-regulated proteins kinases. Among these, 44 protein kinases have been reported as HSP90 client kinases (i.e. EGFR, EPHA2 and AKT) (217, 218), whereas the rest has not been reported but constitutes potential novel clients of HSP90. Comparing with the summarized list from Piscard's lab, all identified protein kinases were mapped onto the phylogenetic kinome tree indicating the known clients (yellow marks) and unknown but highly possible novel clients (red marks) based on the data from this study. As shown in Fig. 47, most kinases were actually unaffected (blue marks) and the significantly regulated kinases were found in all the major kinase classes (red, yellow and green marks) but some branches of the TK and AGC classes are particularly rich in regulated kinases. Interestingly, there were 17 kinases showing a significant up-regulation (green marks) including AURKA, FGFR2 and AXL. Among these are a significant number of kinases involved in the cell cycle (AURKA, PBK, PKA $\beta$ , NEK2, ERK3, TTK, and PKR). One effect of HSP90 inhibition, that has been reported previously, is arresting cells in the cell cycle (G2/M), which may either be mediated by these kinases or result in their accumulation during the drug treatment (219, 220), such as AURKA for which the mRNA and protein levels progressively rise as the cells enter G2/M phases and subsequently degrade when the cells exit mitosis (220, 221). Therefore, it might be not surprising to identify the significantly upregulated AURKA after incubating the cells with GA. Two of the kinases are involved in stress responses (OSR1, MST3) whose upregulation could be a response to the cellular stress.

### **4.3.7 GO analysis reveals subgroups of proteins with diverse response to the HSP90 inhibition.**

To further evaluate the strategy, GO analysis was performed for over-represented molecular functions. Considering the fact that affinity-based kinases enrichment approach implemented in this strategy could result in serious bias of the significantly enriched category of protein kinase when performing GO analysis, to estimate this before further analysis, the same GO analysis was tested on the separated full proteome and entire data (full proteome and kinome data). As expected, most enriched clusters remained on the list except the protein kinase activity group suggesting that combining the full proteome and kinome dataset for further analysis was doable (Appendix Fig. S1). Subsequently, the significantly regulated proteins including full proteome data

and Kinobead profiling data from each cell line were loaded for molecular functional GO analysis. The assessment revealed two considerably enriched super clusters common to all four cell lines: protein kinase activity proteins and unfolded protein binding proteins (Fig 48, data from CAL27 cell line). To investigate the regulation behavior for different subgroups of protein, the same GO analysis on down- or up- regulated proteins induced by HSP90 inhibition were performed separately. As shown in Fig. 49 (and Appendix Fig. S2-4), the group of significantly down-regulated proteins was mainly over-represented by the GO term of protein kinase activity, while the unfolded protein binding group was mainly due to the significantly up-regulated proteins containing many up-regulated co-chaperone proteins, such as HSP70, HSP27 and CDC37.

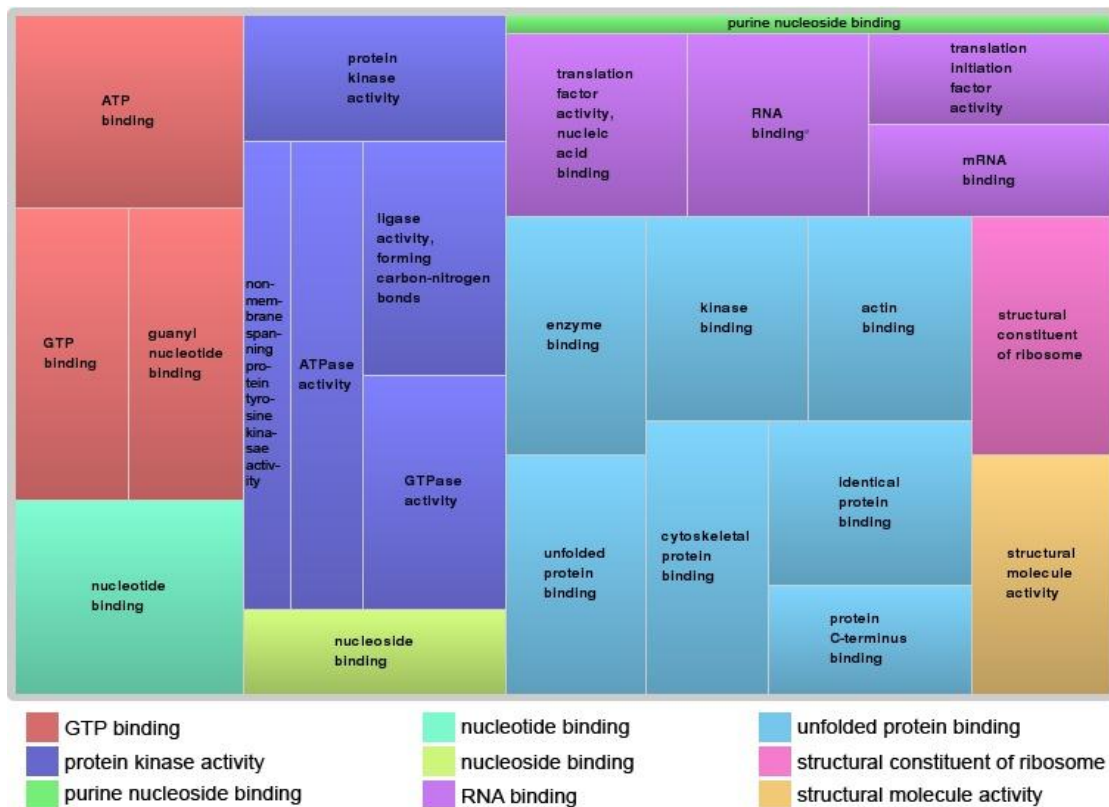


Figure 48. Molecular function GO classification of whole regulated protein on CAL27 cell line. Each rectangle is a single cluster representative. The 'superclusters' of loosely related term sare visualized with different colors. Size of the rectangles reflects the frequency of the GO term in the underlying GOA database. The terms are clustered and the tree map is generated using REVIGO (Supek et al., 2011)

It has been well known that the HSP90 clients are often signal transducers, particular protein kinases. So it is perhaps not surprising to find the over-representation of protein kinase activity in the down regulated protein group (11). Interestingly, the data analysis also suggested that

## HSP90 regulated proteome



Figure 49. Go analysis with down regulated protein (upper panel) and up-regulated proteins (lower panel). The layout follows figure 48.

another group of proteins, with biological function of RNA binding, functionally rely on the HSP90, including the translation initiation factors, ribosomal subunits and some other RNA binding proteins, which are the key players of protein translation (222, 223). So far, it has been



implicated that many RNA binding proteins functionally depend on or physically interact with HSP90 (150, 172, 173). Together with the results described here, it is mostly likely that more RNA binding proteins will be revealed as HSP90 clients and HSP90 may also play more important roles in regulating the protein translation than ever thought before.

### **4.3.8 Cancer specific pathways and networks involved in different cancers**

Next, to investigate the signaling pathways interrupted by HSP90 inhibition, which may implicate the HSP90 dependent aberrant network for promoting the tumorigenesis of a cancer, Ingenuity Pathway Analysis (IPA) was utilized to navigate the biological networks from the significantly regulated proteins. As expected, HSP90 inhibition affected multiple pathways but with strong differences between individual cell lines (Fig. 50a). The head & neck line CAL27 and the leukemia cell line K562 showed a much stronger down regulation of mTOR signaling than the colon cancer line COLO205 and the breast cancer line MDAMB231. Ephrin receptor signaling was strongly regulated in all three solid tumor lines but not in the leukemia cells. Similarities and differences in the general signaling capacities of the different cell lines can also be observed when placing regulated proteins into molecular networks derived from the literature (Fig. 50b and c). Regulated networked proteins are often involved in the cell cycle, cell death, cellular growth and proliferation, DNA replication, recombination and repair (exemplified by networks around the tumor suppressor p53 and the transcription factor complex NFkB (Fig. 50b)), but also include ordinary cellular tasks such as protein synthesis. The protein ubiquitination pathway was overrepresented by the upregulated proteins in all four cell lines. It might be not surprising if considering that HSP90 inhibition triggers the degradation of abandoned clients via the compensatory upregulation of the ubiquitin machinery. Recent research on the significantly up-regulated HSP70 suggested that reduction of HSP90 function can induce an increase in HSP70 binding to the clients to form a complex containing Sti1/Hop, and subsequently initiate the protein degradation processes via the ubiquitin/proteasome pathway (224, 225). Together, in this study, a large number of regulatory proteins were found in many networks, it may not be hard to imagine the pleiotropic effects of HSP90 inhibition will exert on cells.

## HSP90 regulated proteome

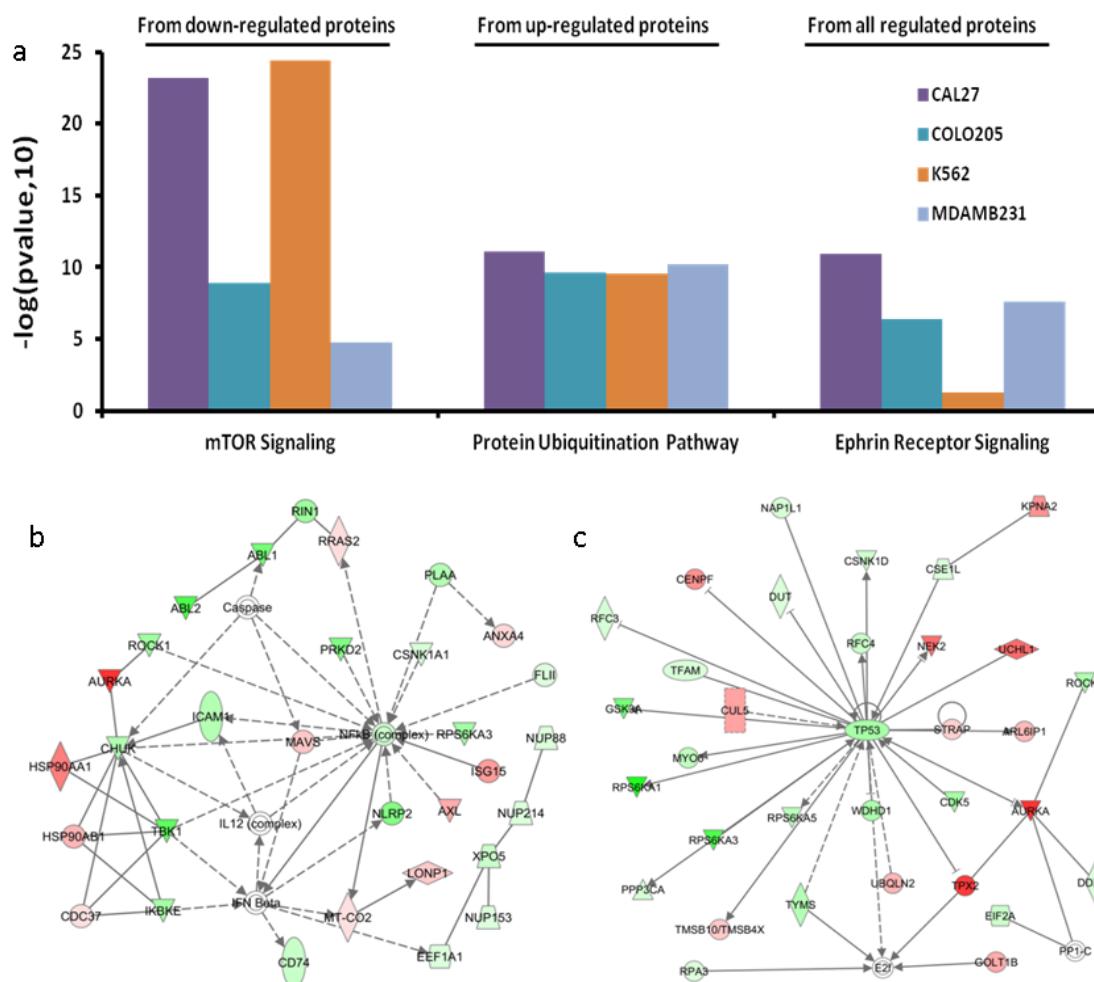


Figure 50. a) Example results of Ingenuity Pathway Analysis (IPA) across the GA regulated proteomes of the four cell lines analyzed. The height of the histogram bars indicates the degree (significance) of overrepresentation of a particular pathway in a particular cell line. Lower panel: Example protein networks generated by Ingenuity Pathway Analysis from significantly GA-regulated proteins of CAL27 (b) and MDAMB231 (c)

### 4.3.9 Co-regulation in protein complex induced by GA

Proteins interacting with each other as a complex play an important role in many cellular processes and functions. As a chaperon protein, HSP90 itself engages in many protein complexes and the data in this study recapitulate a lot of the known binary interactions including many (mostly downregulated) kinases (Fig. 51, data from Cal27 cells). Given the comprehensive analysis of the HSP90 regulated proteome here, one would wonder if proteins within a complex perform similar responses to GA. Interestingly, by querying the CORUM protein complex

database from mammalian organisms (212) with the significantly regulated proteins from each cell line, several members of protein complexes sometimes appeared to be co-regulated by HSP90 inhibition. For example, core members of the kinase maturation complex (HSP90a and b, CDC37, HSP70) were upregulated upon drug treatment; this might be a simple compensatory upregulation because of the inhibition of HSP90 function. In addition, six out of the seven members of the Arp2/3 complex (a cellular complex responsible for actin filament nucleation and branching) (226) were detected and all were upregulated upon GA treatment. This observation

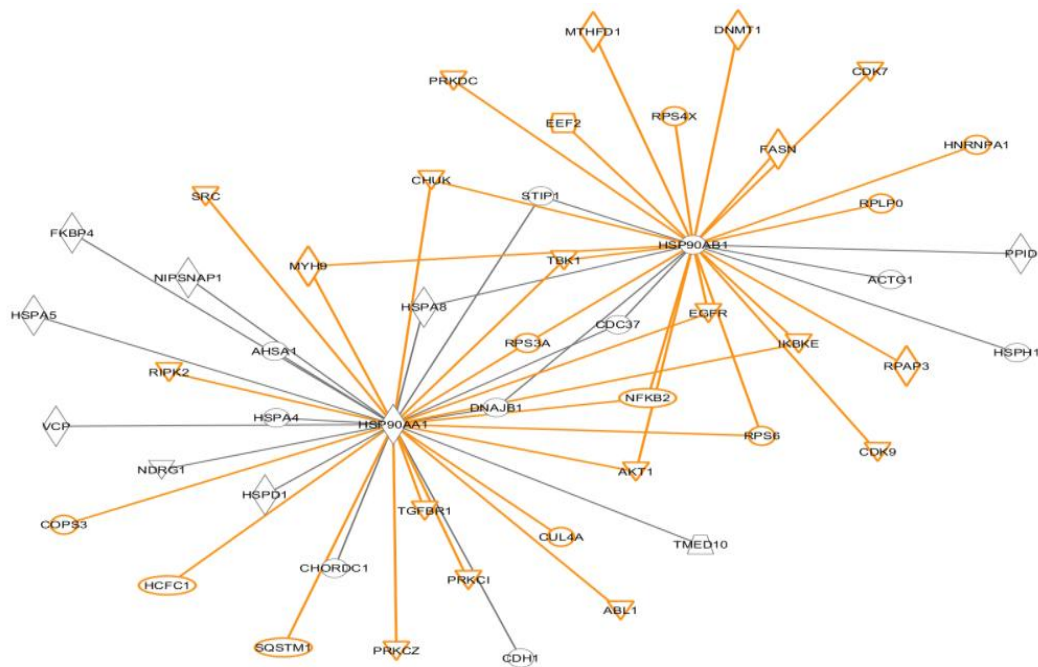


Figure 51. Binary protein-protein interaction map of HSP90alpha and HSP90beta extracted by IPA from the list of geldanamycin regulated proteins. Orange edges and nodes indicate down-regulation; black ones indicate up-regulation upon drug treatment.

could be directly connected to the morphological change after incubating the cells with GA. Cell morphological changes induced by GA might be a comprehensive result of many aspects, hence more work is absolutely required to prove this hypothesis. Down-regulated protein complexes were also observed in this study, such as the DNA-PK-Ku-eIF2-NF90-NF45 complex (DNA repair), the Nop56p-associated pre-ribosomal ribonucleoprotein complex (ribosome biogenesis), the spliceosome (mRNA splicing) and the MNK1-eIF4F complex (protein translation initiation). HSP90 has been found to mediate the assembly by stabilizing the unstable subunit and facilitating its incorporation into the complex, which starts to attract more attention recently (227). Many proteins exert their function in the form of a complex and most, if not all, of the members in the

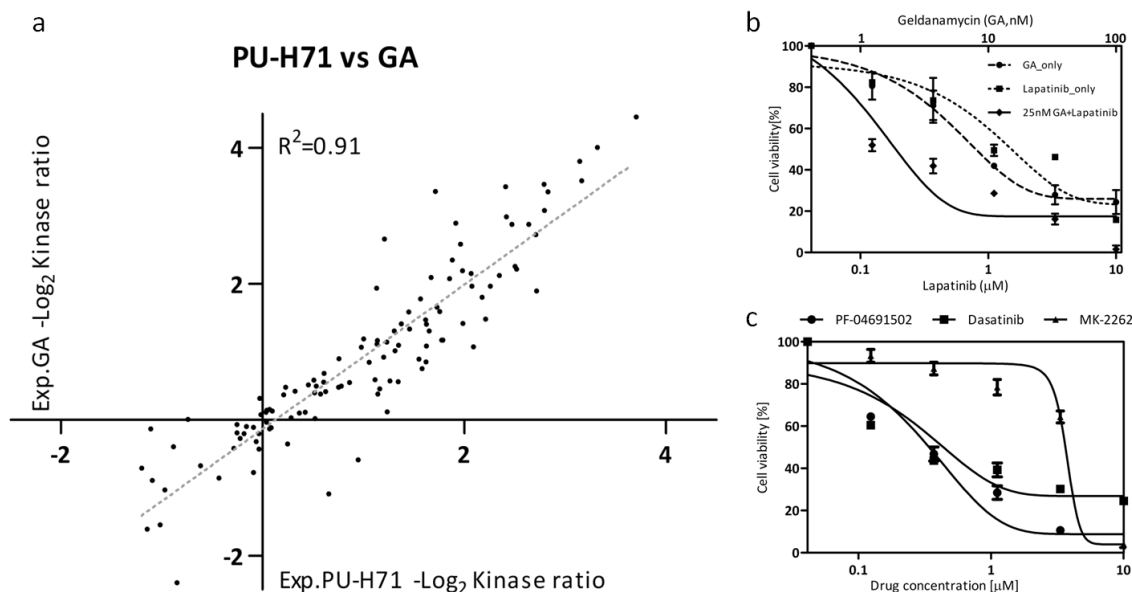
complex are indispensable for the proper function. Therefore most likely, the multiple protein complexes with one HSP90 dependent member will lose its normal function upon the HSP90 inhibition. Several of such cases have been found, for example the Nop56p protein for the Nop56p-associated pre-ribosomal ribonucleoprotein complex (227). The fact of large regulatory proteins of HSP90 inhibition may further emphasize the biological importance and complexity of HSP90.

### **4.3.10 Pharmacological intervention with HSP90 regulated pathways**

Clearly, attenuation of multiple cancer-promoting signaling pathways makes HSP90 become a unique and attractive target for the cancer therapy, particularly the cancers characterized with multiple disordered signaling pathways (228). GA has been proved to be a very effective HSP90 inhibitor but along with many unexpected side effects, which may be not hard to understand if considering the observation that a large number of human proteins with diverse molecular functions appear to be affected by GA. Therefore, other HSP90 inhibitors with better drug properties have been developed, such as the purine-based inhibitor PU-H71, which targets the N-terminal ATP pocket just as GA but with different structure (203), so one would wonder the differences in specificity between GA and PU-H71 which is currently in phase I clinical trials. To compare this, the Kinobead profiling experiment was repeated using both drugs (Fig. 52a). The highly comparable ( $R^2 > 0.9$ ) effect on the kinome suggested the mechanism by which PU-H71 inhibits HSP90 is very similar to GA. However, once these molecules become more widely available, further proteome wide profiling experiments such as the ones described here will be necessary to investigate their global effect and define their specificities.

Combining HSP90 inhibitors with other drugs can improve the therapy efficiency and suppress resistance formation by 'oncogene switching'. This combination can extend the effective treatment time of effective targeted drugs such as kinase inhibitors (10, 229, 230). Prior work from the last chapter has shown that CAL27 cells are growth dependent on EGFR (231), a known client of HSP90. To test the combination effects, CAL27 cells were treated with either lapatinib or GA or both. As expected, either lapatinib or GA alone can significantly reduce the cell viability. Combination of the lapatinib with a fixed dose of GA (25 nM) led to an increased reduction of cell viability albeit the effect was only additive rather than synergistic (Fig. 52b). This preliminary data confirm earlier work that the drug combination is more effective than either drug alone but more work is required to establish if this concept can be generalized (230, 232).

The pathway analysis presented above revealed that mTOR signaling pathway in CAL27 cells was overrepresented by the significantly downregulated proteins indicating GA has a strong negative impact on the mTOR pathway. Considering the fact that constitutively activated mTOR pathway has been found in numerous type of cancers including head and neck cancer (233, 234), which would potentially represent yet another avenue for treating these cancer cells. Indeed, the potent dual PI3K/mTOR inhibitor PF-04691502 also effectively killed CAL27 cells (and better than



*Figure 52. Comparison of different HSP90 inhibitors and kinase inhibitor response of CAL27 cancer cells. a) Correlation plot of kinases regulated by the HSP90 inhibitors geldanamycin (y-axis) and PU-H71 (x-axis). The strong correlation of the two data sets indicates that both drugs work by similar mechanism. b) CAL27 cells are growth dependent on the receptor tyrosine kinase EGFR. Combined treatment of CAL27 cells with geldanamycin and the EGFR inhibitor lapatinib kills cells more effectively than either drug alone. c) The PI3K/mTOR pathway is highly active in CAL27 cells (see Fig. 3b). The dual PI3K/mTOR inhibitor PF-04691502 kills CAL27 cells in an AKT independent fashion (indicated by the inability of the AKT inhibitor MK-2262 to kill the cells) and with similar efficiency as the unspecific kinase inhibitor dasatinib.*

the rather unspecific pan-SRC, pan-EPHR inhibitor dasatinib). Given that the cells are growth dependent on EGFR and respond to lapatinib, this data suggests that EGFR signals into PI3K/mTOR. However, it is not like the previous proposed frequently altered EGFR-PI3K-AKT-mTOR pathway (121), the data in this study suggested that it is in an AKT independent fashion in CAL27 as shown by the inability of the potent AKT inhibitor MK-2262 to kill the same cells (Fig. 52c). EGFR signaling to mTOR through PKC and independently of AKT has been reported in the EGFR driven glioma. Despite that the decrease of phosphorylated AKT abundance in response to

## ***HSP90 regulated proteome***

---

EGFR inhibition was observed (235), which is in correlation with decreased AKT phosphorylation in response to Lapatinib in CAL27 cell (see last chapter). Therefore, it is most likely that the AKT dependent EGFR to mTOR signaling axis is also involved as some of the HNSCC patients.

### **4.3.11 Protein turnover time is a main gradient of HSP90 inhibition induced hierarchical destabilization**

From proteome and kinome wide investigation of the HSP90 based regulation, an interesting observation was serendipitously found that some kinases were removed from cells much faster than others upon drug treatment, which was further validated by Western blot staining (Fig.53a). After incubating with GA for 12 hours, DDR1 and BTK almost completely disappeared while only half of the well known HSP90 client EGFR was significantly degraded after 24 hours. Given the fact that for the client proteins HSP90 is indispensable for their stabilization after the synthesis, it naturally leads to associate this observation to the different turnover rate of individual proteins. To ascertain the hypothesis, the labeling with method with metabolic labeled amino acids (pulsed-SILAC) was used here to determine the kinases turnover time based on the fact that the incorporation rate of the heavy amino acids into the proteins relates to the protein turnover. High turnover proteins incorporate the heavy amino acids faster than the lower turnover proteins after exposing cells to the medium containing the isotopic labeled the amino acids (63, 236). Briefly, kinases were affinity purified by Kinobead from the cells collected at different time points, measured by LC-MS/MS, identified and quantified by Maxquant (Fig. 53b). Three time point (6, 12 and 24 hrs) located in a population of exponentially growing phase were selected, where the incorporation rate should be constant and thereby, heavy to light (H/L) peptides ratio from these time points should perform a linear increase with the time (Fig. 53c). 100 to 120 protein kinase half-lives were determined out of 120 to 150 protein kinases for each of the four cell lines by using linear regression modeling (63). As expect, protein kinases with different turnover times were observed, for example, DDR1 has a half-life time of 7 hours while it takes more than 30 hours to replace half of the existing EGFR (Fig. 53d). Beside, EGFR and AKT, two *bona fide* HSP90 clients, showed a good correlation between the HSP90 induced degradation rate (determined by the western blot) and their turnover ( $R^2=0.96$ ) (Fig. 54a), which suggested the hypothesis was valid to some extent. To further assess the correlation comprehensively in a kinome wide manner, the protein turnover of kinases determined by pulsed SILAC was plotted against the level of kinase down-regulation by HSP90 inhibition (24 h of drug treatment). As

shown in Fig. 54b, a significant positive correlation between HSP90 inhibition and protein turnover was observed. The degree of correlation may reflect the HSP90 dependent regulation: for these strong correlated proteins (e.g. DDR1, EGFR and AKT in the blue zone), which are likely

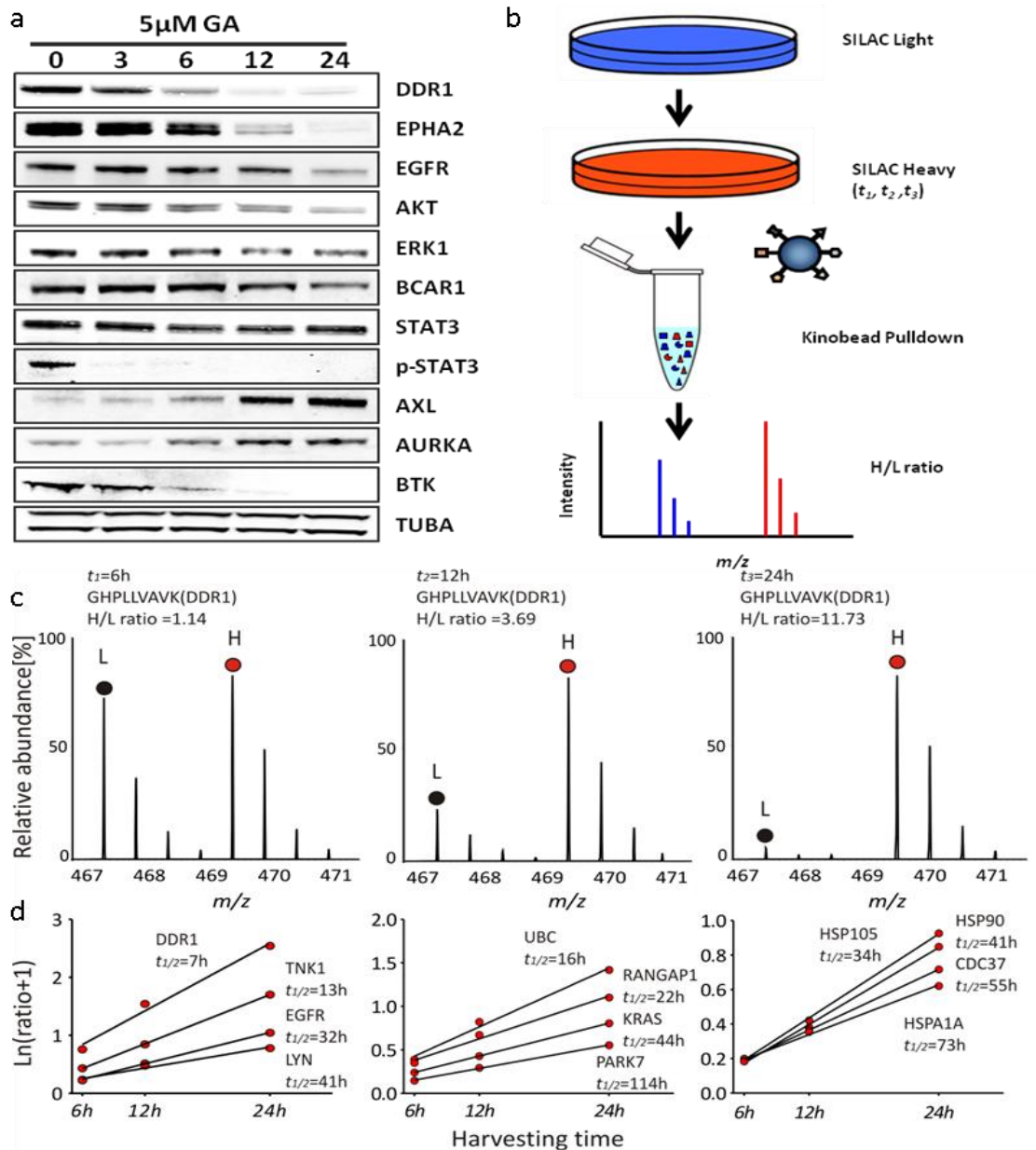


Figure 53. Kinetics of geldanamycin induced proteins degradation and protein turnover. a) Western blot analysis of cells treated with GA for up to 24 hours. Proteins are removed from cells at different rates and consistent with the profiling results. b) schematic representation of a pulsed SILAC experiment. Cells were labeled with heavy amino acids for three different times, light and heavy sample for each time point were combined and analysed by MS. Newly synthesised proteins are detected as heavy peptides while old proteins are detected as light peptides. c) Example mass spectra used for protein half life determinations of the receptor tyrosine kinase DDR1 and d) additional examples for other proteins.

## HSP90 regulated proteome

to be the HSP90 clients, functional HSP90 is indispensable to stabilize these proteins, or else they will be removed from cells at the rate of their normal turnover; while, for proteins with poorer correlation coefficients, more complex regulations must be involved, which requires further work to investigate the details, particular the proteins which remove fast than the turnover rate, such as BTK.

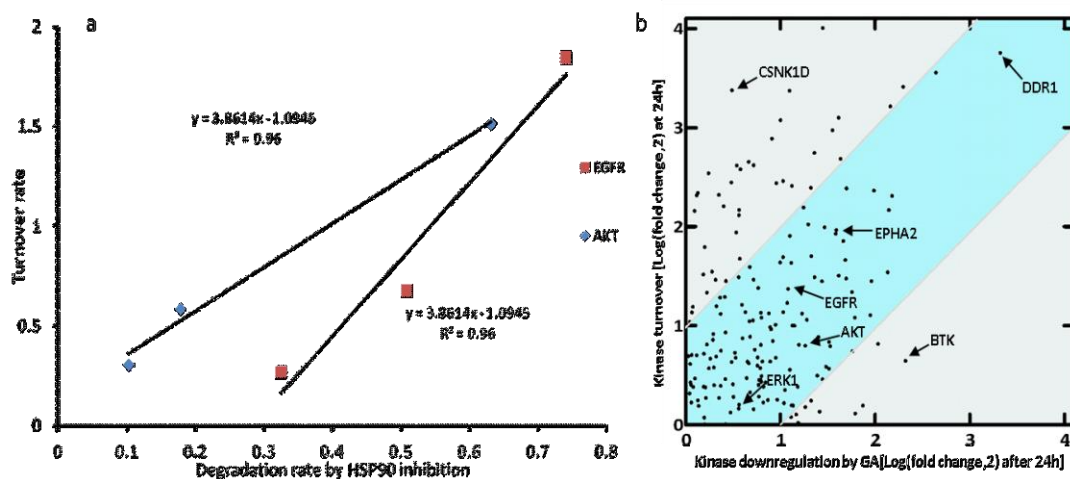


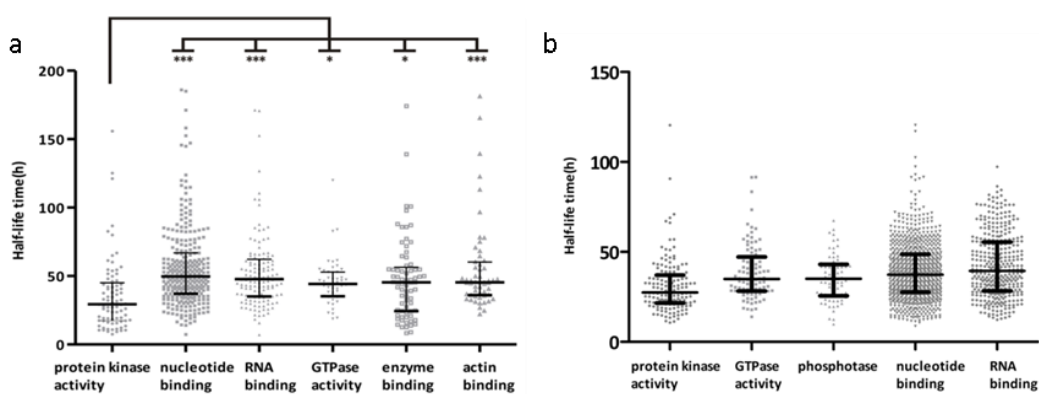
Figure 54. Kinetics of geldanamycin induced proteins degradation is correlated to protein turn over. a) Degradation rates of EGFR and AKT induced by HSP90 inhibition (determined by western blot from Fig. 53a 6, 12 and 24h data points) were plotted against the turnover rate of these two proteins showing a well correlation with both  $R^2 > 0.96$ . b) Correlation plot of quantified kinase protein half lives and levels of corresponding kinase removed by geldanamycin. For strongly correlated kinases (blue area), the protein is removed from cells at the rate of normal turn over.

Given that kinase downregulation is a major consequence of HSP90 inhibition, these differences in pharmacokinetics on individual targets may have several clinically relevant implications. First, it is likely that blocking kinase signaling cascades is the first and possibly a major contributor to drug efficiency. Second, tumors that are driven by kinase of rapid turnover rates (say AKT, EPHA2 or DDR1, Fig. 54a) are likely to respond more quickly to the drug than tumors driven by longer lived proteins (e.g. EGFR). Third, rapid (or prolonged) removal of certain kinases may be undesirable. For example, BTK which is an indispensable kinase for B lymphocyte development, differentiation and signaling will be very rapidly removed upon HSP90 directed therapy and may therefore lead to a loss of function of immune cells of the innate as well as adaptive immune system (237, 238).



**4.3.12 Proteins with protein kinase activity have shorter half-life**

Another incidentally found interesting observation was that the half-lives of kinases were shorter than that of the proteins with other molecular functions. To comprehensively compare the half-life time of protein kinases, the corresponding full proteome from the CAL27 cell lines was further analyzed. In total, the half-lives of 1182 proteins from the full proteome were successfully determined. Before further comparison, proteins with established half-life time were first grouped according to their molecular function on GO terms and the half-life time was compared between the groups with different molecular functions (Fig. 55). The median half-life time of protein kinases was 26 hrs, significantly shorter than the half-life time of nucleotide binding group (50 hrs) and other groups. This finding is furthermore in good agreement with the data from another recently published paper despite using different equations for protein half-life time determination (236), which may artificially enlarge or curtail the values, but should not change the tendency (Fig. 55b).



*Figure 55. GO analysis of protein turnover. a) Kinases have significantly shorter half lives than other protein classes categorized by different molecular function. b) Similar results were observed from the another recently published dataset (Cambridge, et al., J Proteome Res, 2011)*

Previous protein metabolism research in yeast and mammals suggested that proteins can be clustered into two main regimes: production and regulation (63, 239). The former cluster consists of the proteins that are essential for the cell maintenance, which are rather stable and highly abundant. The regulation group represents the regulatory proteins which are rapidly degrading with a short half-life, for example the proteins with kinase activity. These short half-lives of kinases may allow the cancer cells to make quick responses to any stress, such as hypoxia, acidosis, and nutrient deprivation, which frequently occur in tumors.

### **4.3.13 Identification of HSP90 physical interactions by immune precipitation and affinity purification**

As another indispensable evidence to be a client of HSP90, physical interaction was also investigated by co-immunoprecipitation (IP) combined with quantitative mass spectrometry in this study. MDAMB231 and CAL27 cell lines were cultured either in normal DMEM (Light) or DMEM with isotopic labeled Arginine (R6) and Lysine (K4) (Heavy) to enable us to quantitatively distinguish specific interactors of HSP90 from background binding proteins. HSP90 and its interactors were immunoprecipitated from MDAMB231 cells using a HSP90 antibody against its C terminus. Biological triplicates of this experiment led to 766 identified proteins including 744 quantified proteins. In total, 418 proteins were quantified in three replicates contributing to 54.5% of the total identified proteins. All the quantified proteins from the three replicates were subject to Perseus software (62) to determine the significant binding to HSP90. As expected, HSP90 and its co-chaperone proteins were identified with a high ratio, intensity and number of peptides from the immunoprecipitates with HSP90 antibody, suggesting the high quality of the co-immunoprecipitate and indicating a high confidence of the new interactors. In total from two cell lines, around 50 proteins were found to be significantly enriched in HSP90 immunoprecipitates, including several previously proved *bona fide* HSP90 clients: EGFR, CDC2, CDK9, AURKB (240-243) and a number of novel kinases (AXL, AMPK1, EIF2AK2, TRIO, ARAF), but also GTPases and activators (ARHGAP29, SMAP1, XPOT, SPAG1), metabolite transporters (ASCT2, SLC3A2, GLUT1), a further member of the mitochondrial protein import machinery (TOMM34) as well as other (signaling) proteins such as the apoptosis related proteins BIRC6, many of which have not been reported before.

In addition to the chaperon function to stabilize the proteins, HSP90 regulates protein quality by coordinating with many HSP90 co-chaperones, which confer the substrate specificity, to promote the proteins degradation via the ubiquitin/proteasome pathway (225). Current research suggests that the ubiquitin ligases forming part of HSP90 complexes have begun to expand from the established E3-ligase CHIP (77). In this study, a new HSP90 interacting protein called BIRC6 was found, which contains a BIR (baculoviral inhibition of apoptosis protein repeat) domain and a UBCC (ubiquitin-conjugating enzyme E2, catalytic) domain and functions as an apoptosis inhibitor by promoting the degradation of apoptotic proteins by ubiquitination (244). It is therefore tempting to speculate that BIRC6 may be part of the HSP90 ubiquitination system inducing

*Table 2. List of significant HSP90 interactors identified from triplicate HSP90 co-immunoprecipitation experiments and quantitative MS analysis from SILAC labeled CAL27 and MDAMB231 cells.*

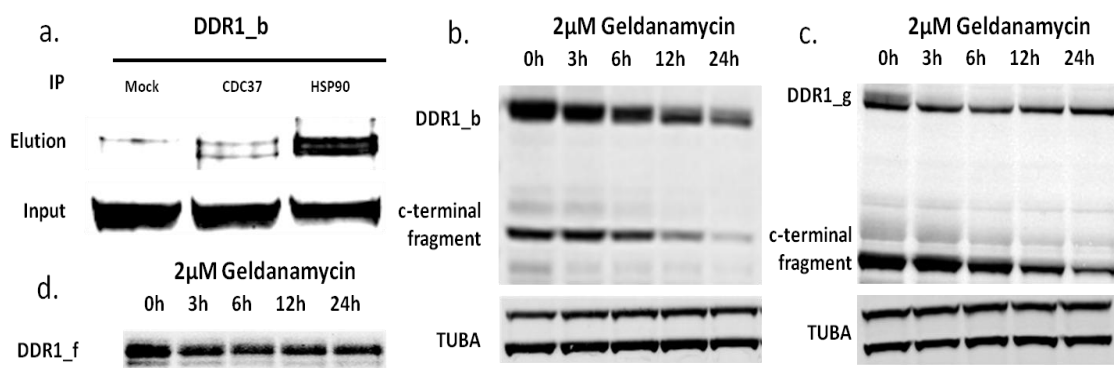
Protein ID	Gene Name	Log2 Ratio	p-value	Known interactor	Protein ID	Gene Name	Log2 Ratio	p-value	Known interactos
IPI00828150	SUGT1	5.0	2.8E-13	Yes	IPI00009946	TOMM34	2.6	2.0E-04	Yes
IPI00152011	ARHGAP29	5.0	3.3E-12	No	IPI00020578	ARAF	3.1	2.1E-04	Yes
IPI00382470	HSP90A	4.8	1.1E-11	Yes	IPI00003865	HSC70	2.5	2.8E-04	Yes
IPI00550746	NUDC	5.0	1.4E-11	Yes	IPI00015602	TOMM70	3.1	3.9E-04	Yes
IPI00455599	HSP90AB2P	4.5	3.4E-11	Yes	IPI00021008	AHR	2.2	5.9E-04	Yes
IPI00735181	SMAP1	3.7	4.0E-11	No	IPI00789101	PTGES3	2.2	1.1E-03	Yes
IPI00796914	AURKB	3.4	9.0E-11	Yes	IPI00071189	CDK11	1.8	1.5E-03	Yes
IPI00414676	HSP90AB1	4.5	1.0E-09	Yes	IPI00296992	AXL	2.3	2.2E-03	No
IPI00555957	HSP90AA4P	3.6	4.6E-09	Yes	IPI00218775	FKBP5	2.0	2.5E-03	Yes
IPI00013894	STIP1	4.4	4.8E-08	Yes	IPI00299635	BIRC6	2.4	4.3E-03	No
IPI00375631	ISG15	3.0	8.5E-08	No	IPI00025156	CHIP	1.8	5.3E-03	Yes
IPI00013122	CDC37	3.2	8.6E-08	Yes	IPI00791333	C20orf4	1.2	5.4E-03	Yes
IPI00072534	SMAP1	3.2	5.8E-07	No	IPI00307572	TMEM165	2.2	7.0E-03	No
IPI00925804	AIP	2.4	1.7E-06	Yes	IPI00019463	EIF2AK2	1.7	9.3E-03	Yes
IPI00019812	PPP5	2.8	1.8E-06	Yes	IPI00657953	TRIO	2.3	1.0E-02	No
IPI00515119	HSP90AB1	3.6	3.2E-06	Yes	IPI00797612	WDR6	2.0	1.1E-02	No
IPI00007811	CDK4	2.7	4.8E-06	Yes	IPI00030706	AHSA1	2.5	1.1E-02	Yes
IPI00005104	IKKA	3.2	5.1E-06	Yes	IPI00019472	ASCT2	1.9	1.8E-02	No
IPI00798127	UBA52	1.7	7.5E-06	No	IPI00021267	EPHA2	1.6	2.2E-02	Yes
IPI00552413	CDK9	2.5	8.3E-06	Yes	IPI00031023	FLII	2.1	2.4E-02	No
IPI00641117	IKBKG	2.8	1.4E-05	Yes	IPI00306290	XPOT	2.1	2.5E-02	No
IPI00556393	FGFR3	2.6	2.1E-05	Yes	IPI00550998	KLHL26	2.1	2.6E-02	No
IPI00410287	AMPK1	2.5	2.3E-05	Yes	IPI00554481	SLC3A2	1.7	2.9E-02	No
IPI00761051	EDRF1	3.0	2.4E-05	No	IPI00018274	EGFR	1.9	3.7E-02	Yes
IPI00026689	CDK1	2.9	3.4E-05	Yes	IPI00014053	TOMM40	1.6	3.8E-02	Yes
IPI00000643	BAG2	2.8	3.9E-05	No	IPI00793155	RAB34	1.7	4.1E-02	No
IPI00292326	SPAG1	2.2	1.9E-04	Yes	IPI00872375	GLUT1	1.8	4.3E-02	No

degradation of specific substrates which remains unknown. Recent literature also suggests that BIRC6 is a potential therapeutic target in colon cancer stem cells which warrants further research on this protein (245).

In addition, affinity based purification via the immobilized GA as a complementary approach to HSP90 IPs was applied to purify HSP90 and its partners from SILAC labeled MDAMB231 cells (in biological triplicate). The quantitative MS analysis identified about 40 significantly drug-enriched proteins including several known HSP90 interactors, such as SUGT1 and STIP1. Finally, the obtained list of physical HSP90 interactors was compared with the 1,600 GA regulated proteins in order to examine which of these proteins may qualify as a HSP90 client. About 50 such proteins were identified that are currently the best validated set of novel HSP90 client proteins in this study as displayed in Fig. 56.



induced dramatic reduction of DDR1 in CAL27 cells (approximately 13 fold), which was consistent with the time dependent reduction determined by western blot (Fig. 57). Furthermore, the degradation rate of DDR1 was quite close to the turnover rate suggesting that DDR1 is a bona fide client of HSP90: the reduction of DDR1 by HSP90 inhibition was mainly due to the loss of function of HSP90, which consequently induces the degradation of the nascent DDR1. To validate this hypothesis, the glioblastoma SF268 cell line stably transfected with the established DDR1b (provided by Dr. Kramer), the most common isoform in nature, was used to investigate the physical interaction between the HSP90 and DDR1. As shown in Fig.57a, DDR1 interacts both with HSP90 and CDC37, which is a HSP90 associated co-chaperone protein, especially for kinases. In addition two novel DDR1 isoforms, DDR1f and DDR1g (provided by Dr. Kramer), were also cloned into the SF268 cell line to investigate their dependence of HSP90. Compared to DDR1b, DDR1f is a partial isoform without the C-terminal domain while DDR1g is missing the juxtamembrane domain which may keep this isoform consecutively active. Our data illustrated that DDR1 isoform g but not isoform f is HSP90 dependent (Fig. 57) suggesting that C-terminal domain of DDR1 is indispensable for the physical interaction between DDR1 and the HSP90 chaperone protein. An interesting observation that only the C-terminal fragment of DDR1 isoform g shows a significant degradation, and not the entire DDR1 isoform g protein, supports the hypothesis that juxtamembrane domain works auto inhibitory on DDR1 activity. Without the juxtamembrane domain the new matured protein is immediately activated and cleaved, therefore only the C-terminal fragment part is showing significant degradation, while the intact protein might represent the part of matured forms which no longer need the functional HSP90. However to



*Figure 57. DDR1 is a bona fide client of HSP90. a) co-immunoprecipitation and western blot analysis reveals that DDR1 is an interactor of the HSP90/CDC37 complex. b-f) time course of 2 μM Geldanamycin in whole cell lysates from SF268 cells carrying different DDR1 isoforms.*

## ***HSP90 regulated proteome***

---

fully understand the biological roles of DDR1 and the different isoforms, additional research is required in order to reveal the impact of HSP90 on the cells via the DDR1 pathways, which currently is not clear yet.

### **4.4 Conclusion**

In this study, a combination of stable isotope labeling, core proteome profiling, chemical precipitation of kinases and quantitative mass spectrometry has been used to investigate the regulated proteome in four different tissue derived cell lines, which resulted in >6,300 identified proteins including around 300 protein kinases. Statistical analysis of the quantitative data revealed that the expression levels of a rather unexpectedly high number of ~1,600 proteins were affected by the drug treatment. Among these, 117 protein kinases showed significant regulation including 100 down-regulated ones. Pathways analysis on these significantly regulated proteins revealed specific pathways and networks affected by HSP90 inhibition in different types of cancers. Protein half-lives determined by pulsed-SILAC suggested that protein turnover was a main gradient of the HSP90 induced proteins hierarchical destabilization. HSP90 IP and chemical compound based affinity purification approaches confirmed part of the known client proteins and also revealed numerous novel partners. Combination with the regulated proteome data, part of the novel partners was validated to be novel client proteins of HSP90. In addition, DDR1 was validated as a novel *bona fide* client of HSP90 by the classic method. Taken together, this study broadly defines the cellular proteome response to HSP90 inhibition and provides a rich resource for further investigation relevant for the treatment of cancer.

### Summary

Protein kinases are among the most intensively studied cellular signaling molecules in normal biology as well as many pathophysiological events, such as cancer. So far, numerous aberrant kinases have been revealed to be involved in most if not all the stages of tumorigenesis. Therefore, characterization of the cancer kinome may offer new insights of therapeutic intervention. In this study, a kinase-centric chemical proteomics (Kinobead) in conjunction with quantitative mass spectrometry is established and optimized, which enable quantitative profiling around 150 kinases in parallel within 4% and 8% overall variance between technical and biological replicates respectively. First, it is employed to investigate 34 head and neck squamous cell carcinoma (HNSCC) cell lines resulting in 146 quantified kinases, of which 42 kinases showed statistically significant ( $p < 0.05$ ) expression inter-cell lines. Additional loss of function experiment using siRNA in high- and low-expressing cell lines further identified kinases including EGFR, EPHA2, LYN, JAK1, WEE1 and NEK9 involved in cell survival and proliferation. Among these, EGFR is confirmed as a drug target and EPHA2 is revealed to be a novel drug target. Both contribute to around 20% and 15% of the HNSCC cell lines respectively. This notion is underscored by immunohistochemical analyses showing that high EGFR and EPHA2 expression is detected in a subset of HNSCC tissues. Downstream signaling pathway analysis suggests that EPHA2 promotes the cell proliferation via activating AKT and ERK signaling pathway. In addition, the several significant candidates are the potential targets of the approved potent pan-SRC family kinase inhibitor dasatinib, which significantly reduces some but not all of HNSCC cell lines. These findings may lead to new therapeutic options for HNSCC patients. This may ultimately lead to a more rational approach to individualized cancer diagnosis and therapy. Second, together with a global approach, Kinobead-based profiling is applied to study the HSP90-dependent proteome, which enriches in signal transducer including many kinases, by investigating the protein response upon the HSP90 inhibition. Employing stable isotope labeling with amino acids in cell culture (SILAC) and quantitative mass spectrometry, >6,200 proteins are identified from four different human cell lines and ~1,600 proteins showed significant regulation upon drug treatment. Gene ontology and pathway/network analysis revealed common and cell type-specific regulatory effects with strong connections to unfolded protein binding and protein kinase activity. Of the 288 identified protein kinases, 98 were downregulated (e.g. EGFR, BTK) and 17 up-regulated (e.g.

## ***Summary***

---

AURA, AXL), in response to GA treatment, almost half of which are formerly unknown HSP90 client kinases. Furthermore pulsed-SILAC results suggested that protein down-regulation by HSP90 inhibition correlates with protein half life in many cases. Protein kinases show significantly shorter half lives than other proteins highlighting both challenges and opportunities for HSP90 inhibition in cancer therapy. The highly similar proteomic responses to the HSP90 drugs GA and PU-H71 suggest that both drugs work by similar molecular mechanisms. Several kinases (AXL, DDR1, TRIO) and other signaling proteins (BIRC6, ISG15, FLII) are validated as novel clients of HSP90 using HSP90 immunoprecipitation and affinity-based purification. Taken together, the strategy employed in this study is generic and therefore also of more general utility for the identification of novel drug targets and molecular pathway markers in tumors and broadly definition of the cellular proteome response to HSP90 inhibition provides a rich resource for further investigation relevant for the treatment of cancer.



## **Zusammenfassung**

Proteinkinasen gehören zu den am besten erforschten Molekülen der zellulären Signaltransduktion und spielen sowohl unter physiologischen als auch unter pathophysiologischen Bedingungen, beispielsweise im Krebsgeschehen, eine entscheidende Rolle. Heutzutage sind zahlreiche aberrante Kinasen in fast allen Stadien der Tumorentstehung beschrieben. Die gezielte, Kinasen-zentrierte Untersuchung des Krebsgeschehens bietet daher das Potential, neue Erkenntnisse mit therapeutischem Potential zu gewinnen. In der vorliegenden Studie wurde ein Proteom-weiter, auf quantitativer Massenspektrometrie basierender Kinase-Assay (kinobead) etabliert und optimiert. Dieser Assay ermöglicht die parallele Identifizierung und Quantifizierung von etwa 150 Kinasen bei 4% technischer und 8% biologischer Varianz.

Mithilfe dieses Assay wurden 34 Zelllinien von Plattenepithelkarzinomen des Kopf-Hals-Bereichs (engl. head and neck squamous cell carcinoma, HNSCC) untersucht. Von den 146 quantifizierten Kinasen wiesen 42 Kinasen statistisch signifikante ( $p < 0.05$ ) Expressionsunterschiede zwischen den Zelllinien auf. Darauf aufbauende RNA-Interferenz-Experimente in niedrig- und hoch-exprimierenden Zelllinien identifizierten eine Reihe von Kinasen, u. a. EGFR, EPHA2, LYN, JAK1, WEE1 und NEK9, die am Überleben bzw. der Proliferation der Krebszellen beteiligt sind. Während EGFR als bekanntes Zielprotein für Arzneimittel (engl. drug target) bestätigt werden konnte, wurde mit EPHA2 ein neues drug target identifiziert. 20% bzw. 15% der HNSCC-Zelllinien werden durch EGFR und EPHA2 angetrieben. Immunhistochemische Untersuchungen konnten dies bestätigen und zeigen, dass EGFR und EPHA2 in einigen der untersuchten HNSCC-Geweben stark exprimiert werden. Die Analyse der nachgeschalteten Signaltransduktionswege deutet an, dass EPHA2 die Zellproliferation über die Aktivierung von AKT und des ERK-Signaltransduktionsweges einleitet. Darüberhinaus sind mehrere der signifikant regulierten Kinasen potentielle Wirkorte des zugelassenen Wirkstoffs Dasatinib, der Kinasen der SRC-Familie inhibiert und das Wachstum einiger, jedoch nicht aller HNSCC-Zelllinien verhindert. Zusammengenommen könnten diese Erkenntnisse neue therapeutischen Optionen für HNSCC-Patienten eröffnen und letztlich zu einer rationaleren und besser begründeten Krebsdiagnose und individualisierten Therapie führen.

## ***Zusammenfassung***

---

In einer zweiten Studie wurde der Kinobeads-Assay zusammen mit der globalen Analyse des Proteoms dazu verwendet, das HSP90-abhängige Proteom zu untersuchen, das bekanntermaßen insbesondere Proteine der Signaltransduktion und eine Vielzahl an Kinasen umfasst. Die Proteom-weiten Auswirkungen der Inhibierung von HSP90 wurden mithilfe des SILAC (engl. stable isotope labeling with amino acids in cell culture)-Ansatzes und quantitativer Massenspektrometrie analysiert. Dabei konnten mehr als 6200 Proteine in vier verschiedenen humanen Zelllinien identifiziert werden, und etwa 1600 davon wiesen signifikante Regulation nach HSP90-Inhibierung durch Geldanamycin (GA) auf. Bioinformatische gene ontology-Analysen und Untersuchungen von Signaltransduktionswegen und -netzwerken konnten sowohl allgemeine als auch Zelltyp-spezifische regulatorische Effekte mit starker Assoziation zu den Proteinfunktionen „Bindung ungefalteter Proteine“ und „Proteinkinase-Aktivität“ aufdecken. Von den 288 identifizierten Proteinkinasen wurden 98 nach Stimulus herunterreguliert (z. B. EGFR, BTK) und 17 hoch-reguliert (z. B. AURA, AXL), wobei fast die Hälfte der Kinasen bis dato nicht als HSP90-Klienten bekannt waren. Mithilfe eines pulsed-SILAC-Experiments konnte darüberhinaus gezeigt werden, dass die Herunterregulierung durch Inhibierung von HSP90 in vielen Fällen mit der Lebensdauer der Proteine korreliert. Dass insbesondere Proteinkinasen im Vergleich zu anderen Proteinen eine signifikant verringerte Lebensdauer aufweisen, hebt die besonderen Schwierigkeiten, aber auch Chancen der HSP90-Inhibierung für die Krebstherapie hervor. Interessanterweise sind die Proteom-weiten Auswirkungen der beiden HSP90-Inhibitoren GA und PU-H71 sehr ähnlich, was auf einen ähnlichen molekularen Mechanismus beider Wirkstoffe hindeutet. Mehrere Kinasen (AXL, DDR1, TRIO) und andere Signalproteine (BIRC6, ISG15, FLII) konnten als neue HSP90-Klienten über Immunopräzipitation und Affinitäts-basierte Aufreinigungen validiert werden.

Insgesamt ist die angewandte Strategie generisch ausgelegt und allgemein nützlich für die Identifikation von neuen drug targets und molekularen Signaltransduktionsmarkern in Tumoren. Darüberhinaus bietet die gewonnenen Daten der Proteom-weiten Antwort auf die Inhibierung von HSP90 eine ergiebige Resource für weitere Untersuchungen für die Behandlung von Tumoren.

---

**Reference**

1. Gerritsen, D. J. M. a. M. E. (2010) Targeting protein kinase for cancer therapy.
2. Siegel, R., Naishadham, D., and Jemal, A. (2012) Cancer statistics, 2012. *CA Cancer J Clin* 62, 10-29.
3. Croce, C. M. (2008) Oncogenes and cancer. *N Engl J Med* 358, 502-511.
4. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *Cell* 144, 646-674.
5. Hanahan, D., and Weinberg, R. A. (2000) The hallmarks of cancer. *Cell* 100, 57-70.
6. Pao, W., and Chmielecki, J. (2010) Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat Rev Cancer* 10, 760-774.
7. Zuo, L., Weger, J., Yang, Q., Goldstein, A. M., Tucker, M. A., Walker, G. J., Hayward, N., and Dracopoli, N. C. (1996) Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* 12, 97-99.
8. Jeffers, M., Rao, M. S., Rulong, S., Reddy, J. K., Subbarao, V., Hudson, E., Vande Woude, G. F., and Resau, J. H. (1996) Hepatocyte growth factor/scatter factor-Met signaling induces proliferation, migration, and morphogenesis of pancreatic oval cells. *Cell Growth Differ* 7, 1805-1813.
9. Quintavalle, M., Elia, L., Price, J. H., Heynen-Genel, S., and Courtneidge, S. A. (2011) A cell-based high-content screening assay reveals activators and inhibitors of cancer cell invasion. *Sci Signal* 4, ra49.
10. Trepel, J., Mollapour, M., Giaccone, G., and Neckers, L. (2010) Targeting the dynamic HSP90 complex in cancer. *Nat Rev Cancer* 10, 537-549.
11. Taipale, M., Jarosz, D. F., and Lindquist, S. (2010) HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat Rev Mol Cell Biol* 11, 515-528.
12. Hirai, H., Sootome, H., Nakatsuru, Y., Miyama, K., Taguchi, S., Tsujioka, K., Ueno, Y., Hatch, H., Majumder, P. K., Pan, B. S., and Kotani, H. (2010) MK-2206, an allosteric Akt inhibitor, enhances antitumor efficacy by standard chemotherapeutic agents or molecular targeted drugs in vitro and in vivo. *Mol Cancer Ther* 9, 1956-1967.
13. Pearson, M. A., and Fabbro, D. (2004) Targeting protein kinases in cancer therapy: a success? *Expert Rev Anticancer Ther* 4, 1113-1124.
14. Rivera, S., and Yuan, F. (2012) Critical Issues In Delivery of RNAi Therapeutics In Vivo. *Curr Pharm Biotechnol*.
15. Brown, T. A. (2002) Genomes, 2nd edition.
16. Gerszten, R. E., and Wang, T. J. (2008) The search for new cardiovascular biomarkers. *Nature* 451, 949-952.
17. Anderson, N. L., and Anderson, N. G. (1998) Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* 19, 1853-1861.
18. Blackstock, W. P., and Weir, M. P. (1999) Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol* 17, 121-127.
19. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207.
20. Han, X., Aslanian, A., and Yates, J. R., 3rd (2008) Mass spectrometry for proteomics. *Curr Opin Chem Biol* 12, 483-490.
21. Chait, B. T. (2006) Chemistry. Mass spectrometry: bottom-up or top-down? *Science* 314, 65-66.

## Reference

---

22. Bogdanov, B., and Smith, R. D. (2005) Proteomics by FTICR mass spectrometry: top down and bottom up. *Mass Spectrom Rev* 24, 168-200.
23. Siuti, N., and Kelleher, N. L. (2007) Decoding protein modifications using top-down mass spectrometry. *Nat Methods* 4, 817-821.
24. Steen, H., and Mann, M. (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* 5, 699-711.
25. Granvogl, B., Ploscher, M., and Eichacker, L. A. (2007) Sample preparation by in-gel digestion for mass spectrometry-based proteomics. *Anal Bioanal Chem* 389, 991-1002.
26. Wysocki, V. H., Resing, K. A., Zhang, Q., and Cheng, G. (2005) Mass spectrometry of peptides and proteins. *Methods* 35, 211-222.
27. Ahmed, F. E. (2009) Sample preparation and fractionation for proteome analysis and cancer biomarker discovery by mass spectrometry. *J Sep Sci* 32, 771-798.
28. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* 11, 49-79.
29. 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.
30. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60, 2299-2301.
31. Taylor, G. (1964) Disintegration of water drops in an electric field. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 383-397.
32. CKMeng, M., M., and Fenn, J. (1988) Of protons or proteins. *Zeitschrift für Physik D Atoms, Molecules and Clusters*, 361-368.
33. Wilm, M. (2011) Principles of electrospray ionization. *Mol Cell Proteomics*.
34. Cole, R. B. (2010) Electrospray and MALDI Mass Spectrometry. Second Edition.
35. Wilm, M., and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source. *Anal Chem* 68, 1-8.
36. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379, 466-469.
37. Beavis, R. C., and Chait, B. T. (1989) Matrix-assisted laser-desorption mass spectrometry using 355 nm radiation. *Rapid Commun Mass Spectrom* 3, 436-439.
38. Michael Karas, D. B., Franz Hillenkamp (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem* 57, 5.
39. McDonnell, L. A., and Heeren, R. M. (2007) Imaging mass spectrometry. *Mass Spectrom Rev* 26, 606-643.
40. Chaurand, P., Norris, J. L., Cornett, D. S., Mobley, J. A., and Caprioli, R. M. (2006) New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *J Proteome Res* 5, 2889-2900.
41. Caldwell, R. L., and Caprioli, R. M. (2005) Tissue profiling by mass spectrometry: a review of methodology and applications. *Mol Cell Proteomics* 4, 394-401.
42. Reyzer, M. L., and Caprioli, R. M. (2007) MALDI-MS-based imaging of small molecules and proteins in tissues. *Curr Opin Chem Biol* 11, 29-35.
43. Schwartz, J. C., Senko, M. W., and Syka, J. E. (2002) A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom* 13, 659-669.
44. Gross, J. H. (2011) Mass Spectrometry, a textbook (second edition).

45. Douglas, D. J., Frank, A. J., and Mao, D. (2005) Linear ion traps in mass spectrometry. *Mass Spectrom Rev* 24, 1-29.
46. 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* 40, 430-443.
47. Makarov, A. (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* 72, 1156-1162.
48. Scigelova, M., and Makarov, A. (2006) Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics* 6 Suppl 2, 16-21.
49. Erickson, B. (2006) Linear ion trap/Orbitrap mass spectrometer. *Anal Chem* 78, 2089.
50. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem* 78, 2113-2120.
51. Roepstorff, P., and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* 11, 601.
52. Syka, J. E., 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.
53. 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.
54. 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.
55. Xu, C., and Ma, B. (2006) Software for computational peptide identification from MS-MS data. *Drug Discov Today* 11, 595-600.
56. Dancik, V., Addona, T. A., Clauser, K. R., Vath, J. E., and Pevzner, P. A. (1999) De novo peptide sequencing via tandem mass spectrometry. *J Comput Biol* 6, 327-342.
57. Taylor, J. A., and Johnson, R. S. (1997) Sequence database searches via de novo peptide sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom* 11, 1067-1075.
58. Johnson, R. S., and Taylor, J. A. (2000) Searching sequence databases via de novo peptide sequencing by tandem mass spectrometry. *Methods Mol Biol* 146, 41-61.
59. Standing, K. G. (2003) Peptide and protein de novo sequencing by mass spectrometry. *Curr Opin Struct Biol* 13, 595-601.
60. Johnson, R. S., and Taylor, J. A. (2002) Searching sequence databases via de novo peptide sequencing by tandem mass spectrometry. *Mol Biotechnol* 22, 301-315.
61. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 1, 252-262.
62. 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.
63. Schwanhauss, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* 473, 337-342.
64. 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.

## Reference

---

65. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376-386.
66. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17, 994-999.
67. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlett-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3, 1154-1169.
68. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A* 100, 6940-6945.
69. Jensen, O. N. (2004) Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 8, 33-41.
70. Uhlen, M., and Ponten, F. (2005) Antibody-based proteomics for human tissue profiling. *Mol Cell Proteomics* 4, 384-393.
71. Corthals, G. L., and Nelson, P. S. (2001) Large-scale proteomics and its future impact on medicine. *Pharmacogenomics J* 1, 15-19.
72. Sem, D. S. (2004) Chemical proteomics from a nuclear magnetic resonance spectroscopy perspective. *Expert Rev Proteomics* 1, 165-178.
73. Rix, U., and Superti-Furga, G. (2009) Target profiling of small molecules by chemical proteomics. *Nat Chem Biol* 5, 616-624.
74. 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.
75. 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.
76. 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., Dumpelfeld, 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.
77. Moulick, K., Ahn, J. H., Zong, H., Rodina, A., Cerchietti, L., Gomes Dagama, E. M., Caldas-Lopes, E., Beebe, K., Perna, F., Hatzi, K., Vu, L. P., Zhao, X., Zatorska, D., Taldone, T., Smith-Jones, P., Alpaugh, M., Gross, S. S., Pillarsetty, N., Ku, T., Lewis, J. S., Larson, S. M., Levine, R., Erdjument-Bromage, H., Guzman, M. L., Nimer, S. D., Melnick, A., Neckers, L., and Chiosis, G. (2011) Affinity-based proteomics reveal cancer-specific networks coordinated by Hsp90. *Nat Chem Biol* 7, 818-826.
78. Li, J., Rix, U., Fang, B., Bai, Y., Edwards, A., Colinge, J., Bennett, K. L., Gao, J., Song, L., Eschrich, S., Superti-Furga, G., Koomen, J., and Haura, E. B. (2010) A chemical and phosphoproteomic characterization of dasatinib action in lung cancer. *Nat Chem Biol* 6, 291-299.
79. Lenz, T., Fischer, J. J., and Dreger, M. (2011) Probing small molecule-protein interactions: A new perspective for functional proteomics. *J Proteomics*.

80. Fleischer, T. C., Murphy, B. R., Flick, J. S., Terry-Lorenzo, R. T., Gao, Z. H., Davis, T., McKinnon, R., Ostanin, K., Willardsen, J. A., and Boniface, J. J. (2010) Chemical proteomics identifies Nampt as the target of CB30865, an orphan cytotoxic compound. *Chem Biol* 17, 659-664.
81. Fadden, P., Huang, K. H., Veal, J. M., Steed, P. M., Barabasz, A. F., Foley, B., Hu, M., Partridge, J. M., Rice, J., Scott, A., Dubois, L. G., Freed, T. A., Silinski, M. A., Barta, T. E., Hughes, P. F., Ommen, A., Ma, W., Smith, E. D., Spangenberg, A. W., Eaves, J., Hanson, G. J., Hinkley, L., Jenks, M., Lewis, M., Otto, J., Pronk, G. J., Verleysen, K., Haystead, T. A., and Hall, S. E. (2010) Application of chemoproteomics to drug discovery: identification of a clinical candidate targeting hsp90. *Chem Biol* 17, 686-694.
82. Barglow, K. T., and Cravatt, B. F. (2007) Activity-based protein profiling for the functional annotation of enzymes. *Nat Methods* 4, 822-827.
83. Uttamchandani, M., Li, J., Sun, H., and Yao, S. Q. (2008) Activity-based protein profiling: new developments and directions in functional proteomics. *Chembiochem* 9, 667-675.
84. Piersma, S. R., Labots, M., Verheul, H. M., and Jimenez, C. R. (2010) Strategies for kinome profiling in cancer and potential clinical applications: chemical proteomics and array-based methods. *Anal Bioanal Chem* 397, 3163-3171.
85. Parikh, K., and Peppelenbosch, M. P. (2010) Kinome profiling of clinical cancer specimens. *Cancer Res* 70, 2575-2578.
86. Kruse, U., Pallasch, C. P., Bantscheff, M., Eberhard, D., Frenzel, L., Ghidelli, S., Maier, S. K., Werner, T., Wendtner, C. M., and Drewes, G. (2011) Chemoproteomics-based kinome profiling and target deconvolution of clinical multi-kinase inhibitors in primary chronic lymphocytic leukemia cells. *Leukemia* 25, 89-100.
87. 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.
88. Lolli, G., Thaler, F., Valsasina, B., Roletto, F., Knapp, S., Uggeri, M., Bachi, A., Matafora, V., Storici, P., Stewart, A., Kalisz, H. M., and Isacchi, A. (2003) Inhibitor affinity chromatography: profiling the specific reactivity of the proteome with immobilized molecules. *Proteomics* 3, 1287-1298.
89. Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19, 1720-1730.
90. Rosenberg, I. M. (2005) Protein structure. *Protein analysis and purification: benchtop techniques*, 20.
91. Carlson, C. B., Robers, M. B., Vogel, K. W., and Machleidt, T. (2009) Development of LanthaScreen cellular assays for key components within the PI3K/AKT/mTOR pathway. *J Biomol Screen* 14, 121-132.
92. Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* 253, 407-414.
93. Daub, H., Olsen, J. V., Bairlein, M., Gnad, F., Oppermann, F. S., Korner, R., Greff, Z., Keri, 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.
94. 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.

## Reference

---

95. 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.
96. Hopf, C., Eberhard, D., Boesche, M., Bastuck, S., Dumpelfeld, B., and Bantscheff, M. (2012) Determination of kinase inhibitor potencies in cell extracts by competition binding assays and isobaric mass tags. *Methods Mol Biol* 803, 141-155.
97. Patricelli, M. P., Szardenings, A. K., Liyanage, M., Nomanbhoy, T. K., Wu, M., Weissig, H., Aban, A., Chun, D., Tanner, S., and Kozarich, J. W. (2007) Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* 46, 350-358.
98. Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77, 383-414.
99. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal Biochem* 203, 173-179.
100. Hellman, U., Wernstedt, C., Gonez, J., and Heldin, C. H. (1995) Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal Biochem* 224, 451-455.
101. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* 4, 484-494.
102. Hauck, S. M., Dietter, J., Kramer, R. L., Hofmaier, F., Zipplies, J. K., Amann, B., Feuchtinger, A., Deeg, C. A., and Ueffing, M. (2010) Deciphering membrane-associated molecular processes in target tissue of autoimmune uveitis by label-free quantitative mass spectrometry. *Mol Cell Proteomics* 9, 2292-2305.
103. Smyth, G. K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3, Article3.
104. Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y., and Zhang, J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5, R80.
105. Zhang, B., VerBerkmoes, N. C., Langston, M. A., Uberbacher, E., Hettich, R. L., and Samatova, N. F. (2006) Detecting differential and correlated protein expression in label-free shotgun proteomics. *J Proteome Res* 5, 2909-2918.
106. Benjamini, Y., and Hochberg, Y. (1995) CONTROLLING THE FALSE DISCOVERY RATE - A PRACTICAL AND POWERFUL APPROACH TO MULTIPLE TESTING. *J. R. Stat. Soc. Ser. B-Methodol.* 57, 289-300.
107. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 4, 1265-1272.
108. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011) Global cancer statistics. *CA Cancer J Clin* 61, 69-90.
109. Leemans, C. R., Braakhuis, B. J., and Brakenhoff, R. H. (2011) The molecular biology of head and neck cancer. *Nat Rev Cancer* 11, 9-22.



110. Morse, D. E., Katz, R. V., Pendrys, D. G., Holford, T. R., Krutchkoff, D. J., Eisenberg, E., Kosis, D., and Mayne, S. T. (1996) Smoking and drinking in relation to oral epithelial dysplasia. *Cancer Epidemiol Biomarkers Prev* 5, 769-777.
111. Blot, W. J., McLaughlin, J. K., Winn, D. M., Austin, D. F., Greenberg, R. S., Preston-Martin, S., Bernstein, L., Schoenberg, J. B., Stemhagen, A., and Fraumeni, J. F., Jr. (1988) Smoking and drinking in relation to oral and pharyngeal cancer. *Cancer Res* 48, 3282-3287.
112. Franceschi, S., Talamini, R., Barra, S., Baron, A. E., Negri, E., Bidoli, E., Serraino, D., and La Vecchia, C. (1990) Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy. *Cancer Res* 50, 6502-6507.
113. Goldenberg, D., Lee, J., Koch, W. M., Kim, M. M., Trink, B., Sidransky, D., and Moon, C. S. (2004) Habitual risk factors for head and neck cancer. *Otolaryngol Head Neck Surg* 131, 986-993.
114. Kutler, D. I., Auerbach, A. D., Satagopan, J., Giampietro, P. F., Batish, S. D., Huvos, A. G., Goberdhan, A., Shah, J. P., and Singh, B. (2003) High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch Otolaryngol Head Neck Surg* 129, 106-112.
115. Pavia, M., Pileggi, C., Nobile, C. G., and Angelillo, I. F. (2006) Association between fruit and vegetable consumption and oral cancer: a meta-analysis of observational studies. *Am J Clin Nutr* 83, 1126-1134.
116. Agrawal, N., Frederick, M. J., Pickering, C. R., Bettegowda, C., Chang, K., Li, R. J., Fakhry, C., Xie, T. X., Zhang, J., Wang, J., Zhang, N., El-Naggar, A. K., Jasser, S. A., Weinstein, J. N., Trevino, L., Drummond, J. A., Muzny, D. M., Wu, Y., Wood, L. D., Hruban, R. H., Westra, W. H., Koch, W. M., Califano, J. A., Gibbs, R. A., Sidransky, D., Vogelstein, B., Velculescu, V. E., Papadopoulos, N., Wheeler, D. A., Kinzler, K. W., and Myers, J. N. (2011) Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 333, 1154-1157.
117. Stransky, N., Egloff, A. M., Tward, A. D., Kostic, A. D., Cibulskis, K., Sivachenko, A., Kryukov, G. V., Lawrence, M. S., Sougnez, C., McKenna, A., Shefler, E., Ramos, A. H., Stojanov, P., Carter, S. L., Voet, D., Cortes, M. L., Auclair, D., Berger, M. F., Saksena, G., Guiducci, C., Onofrio, R. C., Parkin, M., Romkes, M., Weissfeld, J. L., Seethala, R. R., Wang, L., Rangel-Escareno, C., Fernandez-Lopez, J. C., Hidalgo-Miranda, A., Melendez-Zajgla, J., Winckler, W., Ardlie, K., Gabriel, S. B., Meyerson, M., Lander, E. S., Getz, G., Golub, T. R., Garraway, L. A., and Grandis, J. R. (2011) The mutational landscape of head and neck squamous cell carcinoma. *Science* 333, 1157-1160.
118. Bornstein, S., White, R., Malkoski, S., Oka, M., Han, G., Cleaver, T., Reh, D., Andersen, P., Gross, N., Olson, S., Deng, C., Lu, S. L., and Wang, X. J. (2009) Smad4 loss in mice causes spontaneous head and neck cancer with increased genomic instability and inflammation. *J Clin Invest* 119, 3408-3419.
119. Hague, A., Eveson, J. W., MacFarlane, M., Huntley, S., Janghra, N., and Thavaraj, S. (2004) Caspase-3 expression is reduced, in the absence of cleavage, in terminally differentiated normal oral epithelium but is increased in oral squamous cell carcinomas and correlates with tumour stage. *J Pathol* 204, 175-182.
120. Wang, D., Song, H., Evans, J. A., Lang, J. C., Schuller, D. E., and Weghorst, C. M. (1997) Mutation and downregulation of the transforming growth factor beta type II receptor gene in primary squamous cell carcinomas of the head and neck. *Carcinogenesis* 18, 2285-2290.
121. Freudlsperger, C., Burnett, J. R., Friedman, J. A., Kannabiran, V. R., Chen, Z., and Van Waes, C. (2011) EGFR-PI3K-AKT-mTOR signaling in head and neck squamous cell carcinomas: attractive targets for molecular-oriented therapy. *Expert Opin Ther Targets* 15, 63-74.

## Reference

---

122. Sharafinski, M. E., Ferris, R. L., Ferrone, S., and Grandis, J. R. (2010) Epidermal growth factor receptor targeted therapy of squamous cell carcinoma of the head and neck. *Head Neck* 32, 1412-1421.
123. Sheu, J. J., Hua, C. H., Wan, L., Lin, Y. J., Lai, M. T., Tseng, H. C., Jinawath, N., Tsai, M. H., Chang, N. W., Lin, C. F., Lin, C. C., Hsieh, L. J., Wang, T. L., Shih le, M., and Tsai, F. J. (2009) Functional genomic analysis identified epidermal growth factor receptor activation as the most common genetic event in oral squamous cell carcinoma. *Cancer Res* 69, 2568-2576.
124. Astsaturov, I., Cohen, R. B., and Harari, P. M. (2008) Clinical application of EGFR inhibitors in head and neck squamous cell cancer. *Cancer Treat Res* 139, 135-152.
125. Song, J., Chen, C., and Raben, D. (2004) Emerging role of EGFR-targeted therapies and radiation in head and neck cancer. *Oncology (Williston Park)* 18, 1757-1767; discussion 1767, 1771-1752, 1777.
126. Zimmermann, M., Zouhair, A., Azria, D., and Ozsahin, M. (2006) The epidermal growth factor receptor (EGFR) in head and neck cancer: its role and treatment implications. *Radiat Oncol* 1, 11.
127. Huang, S. M., and Harari, P. M. (2000) Modulation of radiation response after epidermal growth factor receptor blockade in squamous cell carcinomas: inhibition of damage repair, cell cycle kinetics, and tumor angiogenesis. *Clin Cancer Res* 6, 2166-2174.
128. Mazumdar, A., Henderson, Y. C., El-Naggar, A. K., Sen, S., and Clayman, G. L. (2009) Aurora kinase A inhibition and paclitaxel as targeted combination therapy for head and neck squamous cell carcinoma. *Head Neck* 31, 625-634.
129. Gerster, K., Shi, W., Ng, B., Yue, S., Ito, E., Waldron, J., Gilbert, R., and Liu, F. F. (2010) Targeting polo-like kinase 1 enhances radiation efficacy for head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 77, 253-260.
130. Seiwert, T. Y., Jagadeeswaran, R., Faoro, L., Janamanchi, V., Nallasura, V., El Dinali, M., Yala, S., Kanteti, R., Cohen, E. E., Lingen, M. W., Martin, L., Krishnaswamy, S., Klein-Szanto, A., Christensen, J. G., Vokes, E. E., and Salgia, R. (2009) The MET receptor tyrosine kinase is a potential novel therapeutic target for head and neck squamous cell carcinoma. *Cancer Res* 69, 3021-3031.
131. Mallick, P., and Kuster, B. (2010) Proteomics: a pragmatic perspective. *Nat Biotechnol* 28, 695-709.
132. Bantscheff, M., Hopf, C., Kruse, U., and Drewes, G. (2007) Proteomics-based strategies in kinase drug discovery. *Ernst Schering Found Symp Proc*, 1-28.
133. Patel, V. J., Thalassinos, K., Slade, S. E., Connolly, J. B., Crombie, A., Murrell, J. C., and Scrivens, J. H. (2009) A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res* 8, 3752-3759.
134. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 18 Suppl 1, S96-104.
135. Karp, N. A., Huber, W., Sadowski, P. G., Charles, P. D., Hester, S. V., and Lilley, K. S. (2010) Addressing accuracy and precision issues in iTRAQ quantitation. *Mol Cell Proteomics* 9, 1885-1897.
136. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 4, 844-847.

137. Rauser, S., Langer, R., Tschernitz, S., Gais, P., Jutting, U., Feith, M., Hofler, H., and Walch, A. (2010) High number of CD45RO+ tumor infiltrating lymphocytes is an independent prognostic factor in non-metastasized (stage I-IIA) esophageal adenocarcinoma. *BMC Cancer* 10, 608.
138. Fang, R., Elias, D. A., Monroe, M. E., Shen, Y., McIntosh, M., Wang, P., Goddard, C. D., Callister, S. J., Moore, R. J., Gorby, Y. A., Adkins, J. N., Fredrickson, J. K., Lipton, M. S., and Smith, R. D. (2006) Differential label-free quantitative proteomic analysis of *Shewanella oneidensis* cultured under aerobic and suboxic conditions by accurate mass and time tag approach. *Mol Cell Proteomics* 5, 714-725.
139. Mann, M. (2008) Can proteomics retire the western blot? *J Proteome Res* 7, 3065.
140. Oh, P., Li, Y., Yu, J., Durr, E., Krasinska, K. M., Carver, L. A., Testa, J. E., and Schnitzer, J. E. (2004) Subtractive proteomic mapping of the endothelial surface in lung and solid tumours for tissue-specific therapy. *Nature* 429, 629-635.
141. Griffin, N. M., Yu, J., Long, F., Oh, P., Shore, S., Li, Y., Koziol, J. A., and Schnitzer, J. E. (2010) Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol* 28, 83-89.
142. Cox, J., and Mann, M. (2007) Is proteomics the new genomics? *Cell* 130, 395-398.
143. Caillaud, J. M., Mathieu, M. C., and Carlu, C. (1987) Immunohistochemistry in the diagnosis of difficult cases of prostatic cancer. *Prog Clin Biol Res* 243A, 477-483.
144. Shimosato, Y., Hirohashi, S., Nakajima, T., and Noguchi, M. (1989) Immunohistochemistry of lung cancer: cell differentiation and growth properties. *Cancer Treat Res* 45, 71-87.
145. Grandis, J. R., and Tweardy, D. J. (1993) Elevated levels of transforming growth factor alpha and epidermal growth factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer. *Cancer Res* 53, 3579-3584.
146. Tiemann, K., and Rossi, J. J. (2009) RNAi-based therapeutics-current status, challenges and prospects. *EMBO Mol Med* 1, 142-151.
147. Lee, S., and Wang, Q. (2007) Recent development of small molecular specific inhibitor of protein tyrosine phosphatase 1B. *Med Res Rev* 27, 553-573.
148. Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 21, 635-637.
149. Nyati, M. K., Morgan, M. A., Feng, F. Y., and Lawrence, T. S. (2006) Integration of EGFR inhibitors with radiochemotherapy. *Nat Rev Cancer* 6, 876-885.
150. Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., Mardis, E., Kupfer, D., Wilson, R., Kris, M., and Varmus, H. (2004) EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A* 101, 13306-13311.
151. Nagashima, K., Shumway, S. D., Sathyanarayanan, S., Chen, A. H., Dolinski, B., Xu, Y., Keilhack, H., Nguyen, T., Wiznerowicz, M., Li, L., Lutterbach, B. A., Chi, A., Paweletz, C., Allison, T., Yan, Y., Munshi, S. K., Klippel, A., Kraus, M., Bobkova, E. V., Deshmukh, S., Xu, Z., Mueller, U., Szewczak, A. A., Pan, B. S., Richon, V., Pollock, R., Blume-Jensen, P., Northrup, A., and Andersen, J. N. (2011) Genetic and pharmacological inhibition of PDK1 in cancer cells: characterization of a selective allosteric kinase inhibitor. *J Biol Chem* 286, 6433-6448.
152. Hidalgo-Carcedo, C., Hooper, S., Chaudhry, S. I., Williamson, P., Harrington, K., Leitinger, B., and Sahai, E. (2011) Collective cell migration requires suppression of actomyosin at cell-cell contacts mediated by DDR1 and the cell polarity regulators Par3 and Par6. *Nat Cell Biol* 13, 49-58.

## Reference

---

153. Gschwind, A., Fischer, O. M., and Ullrich, A. (2004) The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat Rev Cancer* 4, 361-370.
154. Pardani, A., Lasho, T., Smith, G., Burns, C. J., Fantino, E., and Tefferi, A. (2009) CYT387, a selective JAK1/JAK2 inhibitor: in vitro assessment of kinase selectivity and preclinical studies using cell lines and primary cells from polycythemia vera patients. *Leukemia* 23, 1441-1445.
155. Goldenberg-Furmanov, M., Stein, I., Pikarsky, E., Rubin, H., Kasem, S., Wygoda, M., Weinstein, I., Reuveni, H., and Ben-Sasson, S. A. (2004) Lyn is a target gene for prostate cancer: sequence-based inhibition induces regression of human tumor xenografts. *Cancer Res* 64, 1058-1066.
156. GXA Microarray data are available in the ArrayExpress database: <<<http://www.ebi.ac.uk/gxa>>> under accession number E-MTAB-62.
157. Jarvinen, A. K., Autio, R., Kilpinen, S., Saarela, M., Leivo, I., Grenman, R., Makitie, A. A., and Monni, O. (2008) High-resolution copy number and gene expression microarray analyses of head and neck squamous cell carcinoma cell lines of tongue and larynx. *Genes Chromosomes Cancer* 47, 500-509.
158. Turret, J., and McKeon, F. (1996) Tyrosine kinases wee1 and mik1 as effectors of DNA replication checkpoint control. *Prog Cell Cycle Res* 2, 91-97.
159. Posthumadeboer, J., Wurdinger, T., Graat, H. C., van Beusechem, V. W., Helder, M. N., van Royen, B. J., and Kaspers, G. J. (2011) WEE1 inhibition sensitizes Osteosarcoma to Radiotherapy. *BMC Cancer* 11, 156.
160. Schwartz, D. L., and Dong, L. (2011) Adaptive radiation therapy for head and neck cancer—can an old goal evolve into a new standard? *J Oncol* 2011.
161. Wykosky, J., and Debinski, W. (2008) The EphA2 receptor and ephrinA1 ligand in solid tumors: function and therapeutic targeting. *Mol Cancer Res* 6, 1795-1806.
162. Liu, D. P., Wang, Y., Koeffler, H. P., and Xie, D. (2007) Ephrin-A1 is a negative regulator in glioma through down-regulation of EphA2 and FAK. *Int J Oncol* 30, 865-871.
163. Yang, N. Y., Fernandez, C., Richter, M., Xiao, Z., Valencia, F., Tice, D. A., and Pasquale, E. B. Crosstalk of the EphA2 receptor with a serine/threonine phosphatase suppresses the Akt-mTORC1 pathway in cancer cells. *Cell Signal* 23, 201-212.
164. Wykosky, J., Gibo, D. M., and Debinski, W. (2007) A novel, potent, and specific ephrinA1-based cytotoxin against EphA2 receptor expressing tumor cells. *Mol Cancer Ther* 6, 3208-3218.
165. Shao, Z., Zhang, W. F., Chen, X. M., and Shang, Z. J. (2008) Expression of EphA2 and VEGF in squamous cell carcinoma of the tongue: correlation with the angiogenesis and clinical outcome. *Oral Oncol* 44, 1110-1117.
166. Liu, Y., Zhang, X., Qiu, Y., Huang, D., Zhang, S., Xie, L., Qi, L., Yu, C., Zhou, X., Hu, G., and Tian, Y. (2011) Clinical significance of EphA2 expression in squamous-cell carcinoma of the head and neck. *J Cancer Res Clin Oncol* 137, 761-769.
167. Gonzalez-Martin, A. (2005) Treatment of recurrent disease: randomized trials of monotherapy versus combination chemotherapy. *Int J Gynecol Cancer* 15 Suppl 3, 241-246.
168. Lane, D. (2006) Designer combination therapy for cancer. *Nat Biotechnol* 24, 163-164.
169. Casak, S. J., Lemery, S. J., Shen, Y. L., Rothmann, M. D., Khandelwal, A., Zhao, H., Davis, G., Jarral, V., Keegan, P., and Pazdur, R. (2011) U.S. Food and drug administration approval: rituximab in combination with fludarabine and cyclophosphamide for the treatment of patients with chronic lymphocytic leukemia. *Oncologist* 16, 97-104.
170. Larsen, A. B., Stockhausen, M. T., and Poulsen, H. S. (2010) Cell adhesion and EGFR activation regulate EphA2 expression in cancer. *Cell Signal* 22, 636-644.

171. Larsen, A. B., Pedersen, M. W., Stockhausen, M. T., Grandal, M. V., van Deurs, B., and Poulsen, H. S. (2007) Activation of the EGFR gene target EphA2 inhibits epidermal growth factor-induced cancer cell motility. *Mol Cancer Res* 5, 283-293.
172. Pratt, R. L., and Kinch, M. S. (2002) Activation of the EphA2 tyrosine kinase stimulates the MAP/ERK kinase signaling cascade. *Oncogene* 21, 7690-7699.
173. Zhang, Q., Thomas, S. M., Xi, S., Smithgall, T. E., Siegfried, J. M., Kamens, J., Gooding, W. E., and Grandis, J. R. (2004) SRC family kinases mediate epidermal growth factor receptor ligand cleavage, proliferation, and invasion of head and neck cancer cells. *Cancer Res* 64, 6166-6173.
174. Lombardo, L. J., Lee, F. Y., Chen, P., Norris, D., Barrish, J. C., Behnia, K., Castaneda, S., Cornelius, L. A., Das, J., Doweyko, A. M., Fairchild, C., Hunt, J. T., Inigo, I., Johnston, K., Kamath, A., Kan, D., Klei, H., Marathe, P., Pang, S., Peterson, R., Pitt, S., Schieven, G. L., Schmidt, R. J., Tokarski, J., Wen, M. L., Wityak, J., and Borzilleri, R. M. (2004) Discovery of N-(2-chloro-6-methyl-phenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J Med Chem* 47, 6658-6661.
175. Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., Chan, K. W., Ciceri, P., Davis, M. I., Edeen, P. T., Faraoni, R., Floyd, M., Hunt, J. P., Lockhart, D. J., Milanov, Z. V., Morrison, M. J., Pallares, G., Patel, H. K., Pritchard, S., Wodicka, L. M., and Zarrinkar, P. P. (2008) A quantitative analysis of kinase inhibitor selectivity. *Nat Biotechnol* 26, 127-132.
176. Kim, L. C., Song, L., and Haura, E. B. (2009) Src kinases as therapeutic targets for cancer. *Nat Rev Clin Oncol* 6, 587-595.
177. Argiris, A., Feinstein, T. M., Wang, L., Yang, T., Agrawal, S., Appleman, L. J., Stoller, R. G., Grandis, J. R., and Egloff, A. M. (2011) Phase I and pharmacokinetic study of dasatinib and cetuximab in patients with advanced solid malignancies. *Invest New Drugs*.
178. Christensen, J. G., Schreck, R., Burrows, J., Kuruganti, P., Chan, E., Le, P., Chen, J., Wang, X., Ruslim, L., Blake, R., Lipson, K. E., Ramphal, J., Do, S., Cui, J. J., Cherrington, J. M., and Mendel, D. B. (2003) A selective small molecule inhibitor of c-Met kinase inhibits c-Met-dependent phenotypes in vitro and exhibits cytoreductive antitumor activity in vivo. *Cancer Res* 63, 7345-7355.
179. De Herdt, M. J., and Baatenburg de Jong, R. J. (2008) HGF and c-MET as potential orchestrators of invasive growth in head and neck squamous cell carcinoma. *Front Biosci* 13, 2516-2526.
180. Cortesina, G., Martone, T., Galeazzi, E., Olivero, M., De Stefani, A., Bussi, M., Valente, G., Comoglio, P. M., and Di Renzo, M. F. (2000) Staging of head and neck squamous cell carcinoma using the MET oncogene product as marker of tumor cells in lymph node metastases. *Int J Cancer* 89, 286-292.
181. Di Renzo, M. F., Olivero, M., Martone, T., Maffe, A., Maggiora, P., Stefani, A. D., Valente, G., Giordano, S., Cortesina, G., and Comoglio, P. M. (2000) Somatic mutations of the MET oncogene are selected during metastatic spread of human HNSC carcinomas. *Oncogene* 19, 1547-1555.
182. Gray, P. J., Jr., Prince, T., Cheng, J., Stevenson, M. A., and Calderwood, S. K. (2008) Targeting the oncogene and kinome chaperone CDC37. *Nat Rev Cancer* 8, 491-495.
183. Whitesell, L., and Lindquist, S. L. (2005) HSP90 and the chaperoning of cancer. *Nat Rev Cancer* 5, 761-772.

## Reference

---

184. McClellan, A. J., Xia, Y., Deutschbauer, A. M., Davis, R. W., Gerstein, M., and Frydman, J. (2007) Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. *Cell* 131, 121-135.
185. Zhao, R., Davey, M., Hsu, Y. C., Kaplanek, P., Tong, A., Parsons, A. B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., Boone, C., Emili, A., and Houry, W. A. (2005) Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* 120, 715-727.
186. Millson, S. H., Truman, A. W., King, V., Prodromou, C., Pearl, L. H., and Piper, P. W. (2005) A two-hybrid screen of the yeast proteome for Hsp90 interactors uncovers a novel Hsp90 chaperone requirement in the activity of a stress-activated mitogen-activated protein kinase, Slt2p (Mpk1p). *Eukaryot Cell* 4, 849-860.
187. Citri, A., Harari, D., Shohat, G., Ramakrishnan, P., Gan, J., Lavi, S., Eisenstein, M., Kimchi, A., Wallach, D., Pietrokovski, S., and Yarden, Y. (2006) Hsp90 recognizes a common surface on client kinases. *J Biol Chem* 281, 14361-14369.
188. Tsaytler, P. A., Krijgsveld, J., Goerdayal, S. S., Rudiger, S., and Egmond, M. R. (2009) Novel Hsp90 partners discovered using complementary proteomic approaches. *Cell Stress Chaperones* 14, 629-638.
189. Sreedhar, A. S., Kalmar, E., Csermely, P., and Shen, Y. F. (2004) Hsp90 isoforms: functions, expression and clinical importance. *FEBS Lett* 562, 11-15.
190. Bagatell, R., and Whitesell, L. (2004) Altered Hsp90 function in cancer: a unique therapeutic opportunity. *Mol Cancer Ther* 3, 1021-1030.
191. Grbovic, O. M., Basso, A. D., Sawai, A., Ye, Q., Friedlander, P., Solit, D., and Rosen, N. (2006) V600E B-Raf requires the Hsp90 chaperone for stability and is degraded in response to Hsp90 inhibitors. *Proc Natl Acad Sci U S A* 103, 57-62.
192. Waza, M., Adachi, H., Katsuno, M., Minamiyama, M., Sang, C., Tanaka, F., Inukai, A., Doyu, M., and Sobue, G. (2005) 17-AAG, an Hsp90 inhibitor, ameliorates polyglutamine-mediated motor neuron degeneration. *Nat Med* 11, 1088-1095.
193. Hao, H., Naomoto, Y., Bao, X., Watanabe, N., Sakurama, K., Noma, K., Motoki, T., Tomono, Y., Fukazawa, T., Shirakawa, Y., Yamatsuji, T., Matsuoka, J., and Takaoka, M. (2010) HSP90 and its inhibitors. *Oncol Rep* 23, 1483-1492.
194. Arkin, M. R., and Wells, J. A. (2004) Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* 3, 301-317.
195. Zou, J., Guo, Y., Guettouche, T., Smith, D. F., and Voellmy, R. (1998) Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell* 94, 471-480.
196. Kim, H. R., Kang, H. S., and Kim, H. D. (1999) Geldanamycin induces heat shock protein expression through activation of HSF1 in K562 erythroleukemic cells. *IUBMB Life* 48, 429-433.
197. Bohonowych, J. E., Gopal, U., and Isaacs, J. S. (2010) Hsp90 as a gatekeeper of tumor angiogenesis: clinical promise and potential pitfalls. *J Oncol* 2010, 412985.
198. Hartson, S. D., and Matts, R. L. (2011) Approaches for defining the Hsp90-dependent proteome. *Biochim Biophys Acta*.
199. Castro, J. E., Prada, C. E., Loria, O., Kamal, A., Chen, L., Burrows, F. J., and Kipps, T. J. (2005) ZAP-70 is a novel conditional heat shock protein 90 (Hsp90) client: inhibition of Hsp90 leads to ZAP-70 degradation, apoptosis, and impaired signaling in chronic lymphocytic leukemia. *Blood* 106, 2506-2512.

200. Schumacher, J. A., Crockett, D. K., Elenitoba-Johnson, K. S., and Lim, M. S. (2007) Proteome-wide changes induced by the Hsp90 inhibitor, geldanamycin in anaplastic large cell lymphoma cells. *Proteomics* 7, 2603-2616.
201. Truman, A. W., Millson, S. H., Nuttall, J. M., King, V., Mollapour, M., Prodromou, C., Pearl, L. H., and Piper, P. W. (2006) Expressed in the yeast *Saccharomyces cerevisiae*, human ERK5 is a client of the Hsp90 chaperone that complements loss of the Slt2p (Mpk1p) cell integrity stress-activated protein kinase. *Eukaryot Cell* 5, 1914-1924.
202. Schwanhausser, B., Gossen, M., Dittmar, G., and Selbach, M. (2009) Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* 9, 205-209.
203. Immormino, R. M., Kang, Y., Chiosis, G., and Gewirth, D. T. (2006) Structural and quantum chemical studies of 8-aryl-sulfanyl adenine class Hsp90 inhibitors. *J Med Chem* 49, 4953-4960.
204. Ong, S. E., and Mann, M. (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc* 1, 2650-2660.
205. Team, R. D. C. (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
206. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 18 Suppl 1, S96--104.
207. Wu, Z., Doondeea, J. B., Moghaddas Gholami, A., 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*.
208. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289-300.
209. Huang da, W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., and Lempicki, R. A. (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 35, W169-175.
210. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44-57.
211. Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6, e21800.
212. Ruepp, A., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Stransky, M., Waegle, B., Schmidt, T., Doudieu, O. N., Stumpflen, V., and Mewes, H. W. (2008) CORUM: the comprehensive resource of mammalian protein complexes. *Nucleic Acids Res* 36, D646-650.
213. Lin, S. D., and Fann, M. J. (1998) Differential expression of protein kinases in cultured primary neurons derived from the cerebral cortex, hippocampus, and sympathetic ganglia. *J Biomed Sci* 5, 111-119.
214. Surawska, H., Ma, P. C., and Salgia, R. (2004) The role of ephrins and Eph receptors in cancer. *Cytokine Growth Factor Rev* 15, 419-433.
215. Burkard, T. R., Planyavsky, M., Kaupe, I., Breitwieser, F. P., Burckstummer, T., Bennett, K. L., Superti-Furga, G., and Colinge, J. (2011) Initial characterization of the human central proteome. *BMC Syst Biol* 5, 17.
216. Zhao, R., and Houry, W. A. (2007) Molecular interaction network of the Hsp90 chaperone system. *Adv Exp Med Biol* 594, 27-36.

## Reference

---

217. Sato, S., Fujita, N., and Tsuruo, T. (2000) Modulation of Akt kinase activity by binding to Hsp90. *Proc Natl Acad Sci U S A* 97, 10832-10837.
218. Annamalai, B., Liu, X., Gopal, U., and Isaacs, J. S. (2009) Hsp90 is an essential regulator of EphA2 receptor stability and signaling: implications for cancer cell migration and metastasis. *Mol Cancer Res* 7, 1021-1032.
219. Kim, H. R., Lee, C. H., Choi, Y. H., Kang, H. S., and Kim, H. D. (1999) Geldanamycin induces cell cycle arrest in K562 erythroleukemic cells. *IUBMB Life* 48, 425-428.
220. Jiang, S., Katayama, H., Wang, J., Li, S. A., Hong, Y., Radvanyi, L., Li, J. J., and Sen, S. (2010) Estrogen-induced aurora kinase-A (AURKA) gene expression is activated by GATA-3 in estrogen receptor-positive breast cancer cells. *Horm Cancer* 1, 11-20.
221. Engelbert, D., Schnerch, D., Baumgarten, A., and Wasch, R. (2008) The ubiquitin ligase APC(Cdh1) is required to maintain genome integrity in primary human cells. *Oncogene* 27, 907-917.
222. Sonenberg, N., and Hinnebusch, A. G. (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136, 731-745.
223. Jackson, R. J., Hellen, C. U., and Pestova, T. V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11, 113-127.
224. An, W. G., Schulte, T. W., and Neckers, L. M. (2000) The heat shock protein 90 antagonist geldanamycin alters chaperone association with p210bcr-abl and v-src proteins before their degradation by the proteasome. *Cell Growth Differ* 11, 355-360.
225. Theodoraki, M. A., and Caplan, A. J. (2011) Quality control and fate determination of Hsp90 client proteins. *Biochim Biophys Acta*.
226. Pollard, T. D. (2007) Regulation of actin filament assembly by Arp2/3 complex and formins. *Annu Rev Biophys Biomol Struct* 36, 451-477.
227. Makhnevych, T., and Houry, W. A. (2011) The role of Hsp90 in protein complex assembly. *Biochim Biophys Acta*.
228. Solit, D. B., and Chiosis, G. (2008) Development and application of Hsp90 inhibitors. *Drug Discov Today* 13, 38-43.
229. Shimamura, T., and Shapiro, G. I. (2008) Heat shock protein 90 inhibition in lung cancer. *J Thorac Oncol* 3, S152-159.
230. Tauchi, T., Okabe, S., Ashihara, E., Kimura, S., Maekawa, T., and Ohyashiki, K. (2011) Combined effects of novel heat shock protein 90 inhibitor NVP-AUY922 and nilotinib in a random mutagenesis screen. *Oncogene* 30, 2789-2797.
231. Wu, Z., Doondeea, J. B., Moghaddas Gholami, A., 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* 10, M111 011635.
232. LaBonte, M. J., Wilson, P. M., Fazzone, W., Russell, J., Louie, S. G., El-Khoueiry, A., Lenz, H. J., and Ladner, R. D. (2011) The dual EGFR/HER2 inhibitor lapatinib synergistically enhances the antitumor activity of the histone deacetylase inhibitor panobinostat in colorectal cancer models. *Cancer Res* 71, 3635-3648.
233. Liu, F. Y., Zhao, Z. J., Li, P., Ding, X., Zong, Z. H., and Sun, C. F. (2010) Mammalian target of rapamycin (mTOR) is involved in the survival of cells mediated by chemokine receptor 7 through PI3K/Akt in metastatic squamous cell carcinoma of the head and neck. *Br J Oral Maxillofac Surg* 48, 291-296.



234. Chang, K. Y., Tsai, S. Y., Wu, C. M., Yen, C. J., Chuang, B. F., and Chang, J. Y. (2011) Novel phosphoinositide 3-kinase/mTOR dual inhibitor, NVP-BGT226, displays potent growth-inhibitory activity against human head and neck cancer cells in vitro and in vivo. *Clin Cancer Res* 17, 7116-7126.
235. Fan, Q. W., Cheng, C., Knight, Z. A., Haas-Kogan, D., Stokoe, D., James, C. D., McCormick, F., Shokat, K. M., and Weiss, W. A. (2009) EGFR signals to mTOR through PKC and independently of Akt in glioma. *Sci Signal* 2, ra4.
236. Cambridge, S. B., Gnad, F., Nguyen, C., Bermejo, J. L., Kruger, M., and Mann, M. (2011) Systems-wide Proteomic Analysis in Mammalian Cells Reveals Conserved, Functional Protein Turnover. *J Proteome Res* 10, 5275-5284.
237. Shinozaki, F., Minami, M., Chiba, T., Suzuki, M., Yoshimatsu, K., Ichikawa, Y., Terasawa, K., Emori, Y., Matsumoto, K., Kurosaki, T., Nakai, A., Tanaka, K., and Minami, Y. (2006) Depletion of hsp90beta induces multiple defects in B cell receptor signaling. *J Biol Chem* 281, 16361-16369.
238. Yorgin, P. D., Hartson, S. D., Fella, A. M., Scroggins, B. T., Huang, W., Katsanis, E., Couchman, J. M., Matts, R. L., and Whitesell, L. (2000) Effects of geldanamycin, a heat-shock protein 90-binding agent, on T cell function and T cell nonreceptor protein tyrosine kinases. *J Immunol* 164, 2915-2923.
239. Belle, A., Tanay, A., Bitincka, L., Shamir, R., and O'Shea, E. K. (2006) Quantification of protein half-lives in the budding yeast proteome. *Proc Natl Acad Sci U S A* 103, 13004-13009.
240. Sawai, A., Chandarlapaty, S., Greulich, H., Gonen, M., Ye, Q., Arteaga, C. L., Sellers, W., Rosen, N., and Solit, D. B. (2008) Inhibition of Hsp90 down-regulates mutant epidermal growth factor receptor (EGFR) expression and sensitizes EGFR mutant tumors to paclitaxel. *Cancer Res* 68, 589-596.
241. Senju, M., Sueoka, N., Sato, A., Iwanaga, K., Sakao, Y., Tomimitsu, S., Tominaga, M., Irie, K., Hayashi, S., and Sueoka, E. (2006) Hsp90 inhibitors cause G2/M arrest associated with the reduction of Cdc25C and Cdc2 in lung cancer cell lines. *J Cancer Res Clin Oncol* 132, 150-158.
242. O'Keeffe, B., Fong, Y., Chen, D., Zhou, S., and Zhou, Q. (2000) Requirement for a kinase-specific chaperone pathway in the production of a Cdk9/cyclin T1 heterodimer responsible for P-TEFb-mediated tat stimulation of HIV-1 transcription. *J Biol Chem* 275, 279-287.
243. Lange, B. M., Rebollo, E., Herold, A., and Gonzalez, C. (2002) Cdc37 is essential for chromosome segregation and cytokinesis in higher eukaryotes. *EMBO J* 21, 5364-5374.
244. Hauser, H. P., Bardroff, M., Pyrowolakis, G., and Jentsch, S. (1998) A giant ubiquitin-conjugating enzyme related to IAP apoptosis inhibitors. *J Cell Biol* 141, 1415-1422.
245. Van Houdt, W., Emmink, B., Pham, T., Piersma, S., Verheem, A., Vries, R., Fratantoni, S., Pronk, A., Clevers, H., Borel Rinkes, I., Jimenez, C., and Kranenburg, O. (2011) Comparative Proteomics of Colon Cancer Stem Cells and Differentiated Tumor Cells Identifies BIRC6 as a Potential Therapeutic Target. *Mol Cell Proteomics* 10, M111 011353.
246. Ford, C. E., Lau, S. K., Zhu, C. Q., Andersson, T., Tsao, M. S., and Vogel, W. F. (2007) Expression and mutation analysis of the discoidin domain receptors 1 and 2 in non-small cell lung carcinoma. *Br J Cancer* 96, 808-814.
247. Vogel, W. F., Abdulhussein, R., and Ford, C. E. (2006) Sensing extracellular matrix: an update on discoidin domain receptor function. *Cell Signal* 18, 1108-1116.

## ***Abbreviations***

---

### **Abbreviations**

ACN: acetonitrile

ANOVA: analysis of variance

AQUA: absolute quantification

BP: biological process

CC: cellular component

CCCP: compound-centric chemical proteomics

CMC: critical micelle concentration

CML: chronic myelogenous leukemia

CORUM: Comprehensive Resource of Mammalian protein complexes

CP: compound pulldown

CTD: carboxy terminal domain

CVs: coefficient of variations

DAVID: Database for Annotation: Visualization and Integrated Discovery

DMEM: dulbecco's modified eagle medium

DMSO: dimethyl sulfoxide

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

emPAI: exponentially modified protein abundance index

ESI: electrospray ionization

ETD: electron transfer dissociation

FA: fomic acid

FBS: fetal bovine serum

FDR: false discovery rate

FTICR: Fourier-transform ion cyclotron resonance

GA: geldanamycin

GO: Gene Ontology

HNSCC: head and neck squamous cell carcinoma

HPLC: high-pressure liquid chromatography

HPV: human papillomavirus

HSP90: heat shock protein 90

IAA: iodoacetamide  
ICAT: Isotope-coded affinity tag  
IHC: Immunohistochemistry  
IP: immuno-precipitation  
IPA: Ingenuity pathway analysis  
iTRAQ: Isobaric tag for relative and absolute quantitation  
LIT: linear ion trap  
MALDI: Matrix-assisted laser desorption/ionization  
MD: middle domain  
Mgf: Mascot generic format  
MS: mass spectrometry  
NEAA: non-essential amino acids  
NTD: amino terminal domain  
PBS: phosphate buffered saline  
PCR: polymerase chain reaction  
PH: pleckstrin domain  
PTM: post-translational modification  
REVIGO: Reduce and Visualize Gene Ontology  
RF: radio frequency  
RIPA: Radioimmunoprecipitation assay buffer  
RP-HPLC: reversed phase high-pressure liquid chromatography  
RPMI: Roswell Park Memorial Institute medium  
SDS: sodium dodecyl sulphate  
SFKs: Src family kinases  
SILAC: stable Isotope Labeling with Amino acids in Cell culture  
siRNA: small interference RNA  
TBS: TRIS-buffered saline  
TEAB: Triethylammonium bicarbonate  
TMAs: tissue microarrays  
TOF: time-of-flight  
TRIS: tris(hydroxymethyl)aminomethane:

## ***Abbreviations***

---

UPS2: Universal Proteomics Standard set 2

VSN: variance stabilization normalization

XIC: extracted ion chromatogra

## Appendix

Table S1. List of 34 head and neck cell lines

Name	Provided by	Biopsie taken from
UTSCC10	University of Turku, Finland	Primary tumor
UTSCC14	University of Turku, Finland	Primary tumor
UTSCC16A	University of Turku, Finland	Primary tumor
UTSCC21	University of Turku, Finland	Primary tumor
UTSCC24A	University of Turku, Finland	Primary tumor
UTSCC30	University of Turku, Finland	Primary tumor
UTSCC40	University of Turku, Finland	Primary tumor
UTSCC67	University of Turku, Finland	Primary tumor
UTSCC73	University of Turku, Finland	Primary tumor
UTSCC74A	University of Turku, Finland	Primary tumor
UTSCC76A	University of Turku, Finland	Primary tumor
UTSCC87	University of Turku, Finland	Primary tumor
UPCI:SCC016	University of Pittsburgh, USA	Primary tumor
UPCI:SCC040	University of Pittsburgh, USA	Primary tumor
UPCI:SCC056	University of Pittsburgh, USA	Primary tumor
UPCI:SCC075	University of Pittsburgh, USA	Primary tumor
UPCI:SCC103	University of Pittsburgh, USA	Primary tumor
UPCI:SCC122	University of Pittsburgh, USA	Primary tumor
UPCI:SCC154	University of Pittsburgh, USA	Primary tumor
BICR16	ECACC, Salisbury, UK	Primary tumor
BICR56	ECACC, Salisbury, UK	Primary tumor
SCC-4	ATCC	Primary tumor
SCC-9	ATCC	Primary tumor
SCC-15	ATCC	Primary tumor
SCC-25	ATCC	Primary tumor
CAL27	ATCC	Primary tumor
CAL33	DSMZ, Braunschweig Germany	Primary tumor
HSC-3	Health Sciences Research Resources Bank, Japan	Primary tumor
HSC-4	Health Sciences Research Resources Bank, Japan	Primary tumor
OSC-19	Health Sciences Research Resources Bank, Japan	Metastasis(neck lymph node)
OSC-20	Health Sciences Research Resources Bank, Japan	Metastasis(cervical lymph node)
SAS	Health Sciences Research Resources Bank, Japan	Primary tumor
SIHN-005A	S.A. Eccles, University of Surrey, Guildford, UK	Primary tumor
SIHN-006	S.A. Eccles, University of Surrey, Guildford, UK	Primary tumor

## Appendix

Table S2. ATP mimetics immobilized on the agarose beads

Compound Identifier	Composition	Primary Targets	Structure	coupling density (mmol/L)
TUM0001	C <sub>21</sub> H <sub>17</sub> Cl <sub>2</sub> N <sub>5</sub> O	multiple kinases, ABL		1
TUM0002	C <sub>17</sub> H <sub>16</sub> FN <sub>5</sub>	none		2
TUM0003	C <sub>22</sub> H <sub>27</sub> FN <sub>4</sub> O <sub>2</sub>	multiple kinases, PDGFR, VEGFR, c-kit, Flt, CSF, RET		2
TUM0004	C <sub>28</sub> H <sub>26</sub> N <sub>4</sub> O <sub>3</sub>	multiple kinases, GSK3a		4
TUM0005	C <sub>19</sub> H <sub>20</sub> BrFN <sub>4</sub> O <sub>2</sub>	multiple kinases, VEGFR, EGFR, KDR		2
TUM0006	C <sub>23</sub> H <sub>20</sub> N <sub>4</sub> O <sub>2</sub>	multiple kinases, GSK3a		2
TUM0007	C <sub>20</sub> H <sub>25</sub> ClN <sub>6</sub> O <sub>3</sub>	multiple kinases, CDKs		2
TUM0008	C <sub>21</sub> H <sub>27</sub> N <sub>7</sub> O <sub>3</sub>	AKT1/-2/-3, PKC, PKA		2

Table S3: Cell viability data (in %) following siRNA mediated protein knock-down

Cell line	Gene name													
	ADK	AURKA	BMP2K	CAMK2 D	CAMK2 G	CDK5	CSK	DDR1	EPHA1	EPHA2	EPHA4	EPHB2	EPHB4	FGFR2
	siRNA concentration used in experiment													
	10nM	50nM	10nM	50nM	10nM	10nM	50nM	50nM	10nM	50nM	50nM	50nM	50nM	50nM
BICR16	89.9 ±7.4				107.4 ±9.5							42.7 ±19.8		
CAL27			79.3 ±4.5								51.7 ±10.3			
Cal33			70.1 ±8.4		94.15 ±9.6									
HSC3	89.71 ±21.8													
SCC15							83.2± 27.9							96.2 ±24.2
SCC25		49.2± 20.3												
SCC4										82.2± 15.4				
SCC9	85.2± 19.7													
SIHN 005A					100.6 ±7.5		112.2 ±17.5			118.3 ±7.3	122.5 ±16.8	97.02 ±2.6	112 ±4.4	
UPCI: SCC016								91.3± 9.8						
UPCI: SCC056				109.2 ±16.2										
UPCI: SCC075			101.9 ±1.2			98.1± 6.9					107.3 ±1.6			
UPCI: SCC103												95.2± 9.4		
UPCI: SCC154			99.3 ±13.3								93.9± 8.1			
UTSCC 16A				94.3 ±15.6	96.5 ±12.2									
UTSCC 21							79± 25.3							
UTSCC 24A													106.3 ±12.2	
UTSCC 40										35.7 ±19.6				72.4 ±15.1
UTSCC 73		38.0 ±13.9											93.9± 8.4	
UTSCC 74A	92.3 ±0.2													
UTSCC 76A										93.6 ±6.2				
UTSCC 87									78.94 ±6.0				47.2 ±5.0	

## Appendix

Table S3. Continued

Cell line	Gene name													
	JAK1	LIMK2	LYN	MAPK1	MAPK3	MAPK8	MET	NEK9	PTK2B	RIPK2	SRC	TBK1	WEE1	ZAK
	siRNA concentration used in experiment													
	50nM	10nM	10nM	50nM	50nM	10nM	100nM	50nM	10nM	100nM	10nM	10nM	10nM	50nM
BICR16						93.5± 32.9			92.5± 11.7	29.75± 5.4				
BICR56										62.0± 12.1				
SAS							91.6± 5.8							
SCC15									102.7± 15.8					
SCC25	94.3± 4.5						109.4± 2.4			64.5± 12.1				103.7± 2.8
SIHN 005A			102.4± 5.9				108.4± 5.5	104.1± 24.8	104.2± 3.9		98.3± 4	101.1± 6.4		
SIHN 006								83.4± 2.1						
UPCI: SCC016														92.8± 7.2
UPCI: SCC040						96.8±7 .3								
UPCI: SCC056	58.0± 0.5			92.2± 7.3	97.5± 3.1			47.9± 16.6			103± 14.1		61.9± 20.9	
UPCI: SCC075											69.3± 25		94.7± 10.8	
UPCI: SCC103		87.6± 7.1												
UPCI: SCC154													44.5± 9.7	
UTSCC 10		97± 9.2								37.5± 10.9				
UTSCC 14														
UTSCC 21		92.3± 17.5												95.5± 0.5
UTSCC 24A								56.33± 1.3						
UTSCC 67		86.8± 2.8												
UTSCC 74A						88.7± 8.9	93.0± 1.2					88.6± 2.1		
UTSCC 87		34.2± 19.2												

Columns: cell viability (in %) of cell lines following siRNA mediated knock down of the specified protein and using the specified concentration of siRNA. Values correspond to the average (+/- standard deviation) of three separate experiments. Coloring: green cells indicated cell lines that express high levels of the respective protein





Figure S1. Molecular function GO on the datasets from CAL27 cell line. Upper panel) GO analysis was performed on the full proteome dataset only. Lower panel) GO analysis with the whole dataset (full proteome and kinome) Most of the supercluster remained in the list except the protein kinase activity group suggesting the rationality of GO analysis on combined data. The layout follows figure 48.

## Appendix



Figure S2 Top functional gene ontology categories of the K562 expression profiling data. Upper panel) Go terms associated to up-regulated proteins. Lower panel) Go terms associated to down-regulated proteins. The layout follows figure 48.



Figure S3. Top functional gene ontology categories of the MDAMB231 expression profiling data. Upper panel) Go terms associated to up-regulated proteins. Lower panel) Go terms associated to down-regulated proteins. The layout follows figure 48.

## Appendix



Figure S4 Top functional gene ontology categories of the Colo205 expression profiling data. Upper panel) Go terms associated to up-regulated proteins. Lower panel) Go terms associated to down-regulated proteins. The layout follows figure 48.

### **Acknowledgments**

During the last three and half years of my PhD study and research, I have received extensive support both professionally and personally. Therefore I would like to take this opportunity to express the sincere gratitude to my colleagues and friends:

My deepest gratitude goes first and foremost to my supervisor, Prof. Dr. Bernhard Küster for providing me the opportunity to do my research in his lab. This thesis would not have been possible without his expert supervision and constant support during the entire course of my PhD. Meanwhile I would like to thank Prof. Dr. Bernhard Küster, Prof. Dr. Dieter Langosch and Prof. Dr. Johannes Buchner for being part of my thesis committee.

I would like to express my gratitude and appreciation to Dr. Amin Moghaddas Gholami, Dr. Stephan M Feller, Dr. Karl Kramer, Dr. Simone Lemeer and Prof. Dr. med. Axel Walch for the fruitful collaboration on Head and neck project and Dr. Amin Moghaddas Gholami again for the HSP90 proteome project.

Furthermore, I am most grateful to Dr. Karl Kramer for the guidance and discussion for the biology and Dr. Simone Lemeer and Hannes Hanne (ABD) for the patiently explanation of mass spectrometry and proteomics and extensive support during the experiments.

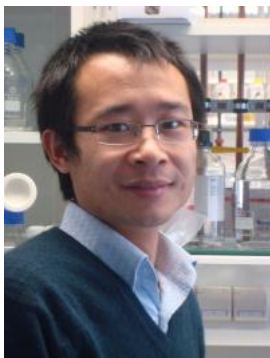
A big thanks to all of my past and current lab colleagues for an extremely friendly environment and helpful gestures that kept me going in the lab. Special thanks to Antje Pfeiffer for the helps of starting my time in Germany and Michaela Krötz-Fahning and Andrea Hubauer for the help of starting the work in the lab.

The generous financial support from the China Scholarship Council (CSC) and the Faculty Graduate Center Weihenstephan of the TUM Graduate School is gratefully acknowledged.

Last thanks would go to my beloved family for the loving considerations and great confidence in me through these years.

Many thanks to all of you!

## Curriculum Vitae – Zhixiang Wu



### Personal Information:

Date of Birth: 12.11.1981

Place of Birth: Jiangsu China

Citizenship: Chinese

### Contact Information:

Lehrstuhl fuer Proteomik und Bioanalytik

Technische Universitaet Muenchen

Emil Erlenmeyer Forum 5

85354 Freising, Germany

Phone: +49-8161712262

Mobile:+49-176-78232025

Fax: 49-8161715931

Email: [wu@wzw.tum.de](mailto:wu@wzw.tum.de)

### Education Background:

#### PhD student in Bioanalytic and proteomics

Technische Universität München (Germany)

Sep. 2008 -present

*Project title:* Quantitative chemical proteomics for cancer characterization

*Supervisor:* Prof. Bernhard Kuster

#### Master in biomedicine engineering. Southeast University (Nanjing, China)

Sep. 2005-Jul.2008

*Project title:* Microarray-based high-throughput DNA methylation analysis for cancer biomarker hunting

*Supervisor:* Prof. Zuhong Lu

#### BSc in clinical medicine. Southeast University (Nanjing, China)

Sep.2000- Jul.2005

## List of Publications

### Publications:

- **Wu, Z.**, Moghaddas Gholami, A and Kuster Global proteome profiling of the NCI-60 cell line panel using label-free quantitative mass spectrometry. *Manuscript in preparation*.
- **Wu, Z.**, Moghaddas Gholami, A., and Kuster, B. (2012) Systematic identification of the HSP90 regulated proteome. *Mol Cell Proteomics*.
- Lemeer, S., Bluwstein, A., **Wu, Z.**, Leberfinger, J., Muller, K., Kramer, K., and Kuster, B. (2011) Phosphotyrosine mediated protein interactions of the discoidin domain receptor 1. *J Proteomics*.
- **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* 10, M111 011635.
- Luo, J., Zheng, W., Wang, Y., **Wu, Z.**, Bai, Y., and Lu, Z. Detection method for methylation density on microarray using methyl-CpG-binding domain protein. *Anal Biochem*. 2009 Apr 15; 387(2):143-9.
- **Wu, Z.**, Luo, J., Ge, Q., and Lu, Z. Microarray-based Ms-SNuPE: near-quantitative analysis for a high-throughput DNA methylation. *Biosens Bioelectron*. 2008 Apr 15; 23(9):1333-9.
- **Wu, Z.**, Luo, J., Ge, Q., Zhang, D., Wang, Y., Jia, C., and Lu, Z. Allele-specific extension on microarray for DNA methylation analysis. *Anal Chim Acta*. 2007 Nov 12; 603(2):199-204.

### Conference:

- The 43<sup>rd</sup> Deutsche Gesellschaft für Massenspektrometrie (DGMS) **Wu Z**, and Kuster B, Comparison of label-free protein quantification approaches for chemical proteomics (Poster). Halle, Germany, March 2010
- The 58<sup>th</sup> American Society for Mass Spectrometry (58<sup>th</sup> ASMS), **Wu Z**, Fellenberg K , Lemeer S and Kuster B, Comparison of label-free protein quantification approaches for chemical proteomics (Poster). Salt lake city, UT, May 2010
- The 59<sup>th</sup> American Society for Mass Spectrometry (59<sup>th</sup> ASMS), **Wu Z**; Pachi F; Kuster B. Performance evaluation of a label-free quantitative chemical proteomics assay (Poster). Denver, CO, May, 2011.
- The 2<sup>nd</sup> Department conference, Wu, Z and Kuster B. Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer. (Talk), Freising, Germany, October, 2011